J. Synchrotron Rad. (1999). **6**, 834–844

Better structures from better data through better methods: a review of developments in *de novo* macromolecular phasing techniques and associated instrumentation at LURE

R. Fourme,[a]* W. Shepard,[a] M. Schiltz,[a] T. Prangé,[a,b] M. Ramin,[a,c] R. Kahn,[d] E. de la Fortelle[e] and G. Bricogne[a,e]

[a]LURE (CNRS, CEA, MENRT), Bâtiment 209D, Université Paris-Sud, 91898-Orsay CEDEX, France, [b]Chimie Structurale Biomoléculaire (URA 1430 CNRS), 93017 Bobigny CEDEX, France, [c]Université d'Evry, Bld François Mitterrand, 91025 Evry CEDEX, France, [d]IBS (CNRS-CEA), 41 Rue des Martyrs, 38027 Grenoble CEDEX, France, and [e]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. E-mail: fourme@lure.u-psud.fr

(*Received 31 March 1999; accepted 6 May 1999*)

This paper presents a survey of developments in *de novo* phasing methods and instrumentation in protein crystallography that have been carried out over the past 20 years at the French synchrotron radiation facility LURE. This includes progress in detector technology, particularly with multiwire proportional chambers, contributions to the development of the MAD and MASC methods for experimental phase-determination *via* anomalous dispersion, the exploration of the use of xenon and krypton as heavy atoms and anomalous scatterers, as well as the substantialization of parts of the 'Bayesian programme' for structure determination in highly efficient and user-friendly software. It is shown how the conjunction of high-quality data collection with novel phasing methods and with optimized data-processing schemes can bring about major improvements, even when the signal is very weak, in the accuracy of structural determinations.

Keywords: protein crystallography; MAD; MASC; xenon; X-ray detectors; maximum-likelihood refinement.

1. Introduction

A common theme in the scientific activity of the co-authors of this contribution has been the simultaneous improvement of experimental and computational techniques for the solution of the phase problem in macromolecular crystallography. This article gives a survey of contributions on *de novo*† phasing. They involve an interplay between new instrumentation and methodologies, theoretical advances and their implementation in new software. The presentation of the 'Bayesian programme' (Bricogne, 1988, 1993), which tackles the phase problem in a more general context, has been restricted to the part which is directly relevant to *de novo* phasing, namely the maximum-likelihood refinement of heavy-atom parameters.

† Here, *de novo* means based on measurements of diffracted intensities, which may include anomalous dispersion or heavy-atom substitution effects, without reference to an already known structure. This is to be distinguished from *ab initio*, a term which we feel should be reserved for computational phasing from native amplitudes only.

2. Advanced data acquisition and processing systems, and first MAD results

2.1. *Spherical-drift multiwire proportional chambers*

The use of synchrotron radiation in the X-ray range became possible at the french synchrotron radiation facility LURE in 1976 when DCI, a 1.85 GeV electron–positron storage ring designed for high-energy physics, became operational. The exploitation of synchrotron radiation in a parasitic mode began immediately with six instruments installed on the first beamline D1. One of these instruments, on station D12, was dedicated to macromolecular crystallography. The set-up was equipped with a purposely modified Enraf–Nonius rotation camera. The monochromator introduced a crystal bender based on elastic properties of a triangular-shaped crystal (R. Fourme, unpublished; Lemonnier *et al.*, 1978). The polarization correction to be applied with the highly polarized synchrotron radiation was derived (Kahn, Fourme, Gadet *et al.*, 1982). This instrument was the first rotation camera used to collect diffraction data with synchrotron radiation

(Fourme, 1978; Kahn, Fourme, Gadet *et al.*, 1982; Wilson *et al.*, 1983; Fourme & Kahn, 1985). A typical example of protein structure solved from data collected at D12 is deoxy hemoglobin at 1.74 Å resolution (Fermi *et al.*, 1984). Modifications in the collimation system, suggested by P. B. Sigler, allowed recording of spatially resolved diffraction patterns from crystals with unit-cell parameters in the range 300–1000 Å (Usha *et al.*, 1984). Experience obtained during these early years was essential for preparing a prospective report on the use of ESRF for protein crystallography (Helliwell & Fourme, 1983).

While D12 was under construction we realised that, in a set-up designed for protein crystallography with synchrotron radiation, the area detector was both the most crucial and the weakest component. Based on the pioneering work of G. Charpak at CERN (Charpak *et al.*, 1968), a multiwire proportional chamber (MWPC) for X-ray diffraction applications had been developed by N. H. Xuong and coworkers at San Diego in collaboration with V. Perez-Mendez at Berkeley (Cork *et al.*, 1974; Xuong *et al.*, 1978). This device had attractive characteristics, especially the high accuracy inherent to photon counters and low instrumental noise. But it suffered from parallax effects (*i.e.* spot elongation observed for beams inclined with respect to the detector axis) and the count-rate capability of its delay-line position encoder was limited. During a meeting held in 1973 with G. Charpak, R. Mössbauer, W. Parak and R. Fourme, it was decided to build two detectors, one for Mössbauer experiments and the other for synchrotron radiation applications. This was the starting point of a long-term collaboration between G. Charpak and his collaborators at CERN (R. Bouclier, R. Benoit, G. Million, J. C. Santiard and F. Sauli) and the LURE team.

At LURE, R. Kahn, the main investigator of this project, and R. Fourme were joined by B. Caudron, then R. Bosshard and later by A. Bahri. Parallax was suppressed by coupling the MWPC to a drift space with radial field lines (Charpak *et al.*, 1974). This drift space is defined by two spherical electrodes and a conical edge with ring-shaped electrodes. Due to the geometry of the drift space and the associated mapping function, the spherical drift MWPC performs a stereographic projection of the reciprocal space (Bricogne, 1987*a*); this projection is conformal, *i.e.* angle preserving, transforming circles on the Ewald sphere into circles on the detector plane. The position encoder was of the wire-by-wire type, with one amplifier per wire and priority encoders. These technical solutions were investigated on a prototype detector, called Mark I (Kahn *et al.*, 1980; Kahn, Fourme, Bosshard *et al.*, 1982), and equipped with a short-bending-radius Al entrance window. This detector was developed with a Cu X-ray tube, then installed on the D25 station. The software for data acquisition and partial on-line data reduction was written by R. Kahn, using the limited resources of a PDP 11/34 minicomputer with a 128 K 16-bit memory. An improved and completely redesigned version of this first detector, called Mark II, was built and installed at the D23 station of LURE-DCI. The Mark II was equipped with a wire-by-wire digital encoder allowing a total count rate in excess of 300 kHz. With a large disc-shaped sensitive area (diameter 0.486 m, 185 500 1×1 mm pixels), a negligible instrumental noise, a detective quantum efficiency close to unity, and a very short dead time ($\sim$1 s) obtained by double-buffering, this detector was a milestone in the development of gas chambers for macromolecular crystallography (Kahn, Fourme, Bosshard, Chiadmi *et al.*, 1986).

At this stage it became clear that a new software package was needed both to drive data acquisition and to process diffraction images. This was provided by the EEC initiative described below, which resulted in the *EEC MADNES* software package. The implementation of *EEC MADNES* on the Mark II diffractometer was a crucial step forward, as it allowed the full exploitation of the intrinsic qualities of the detector [reviewed by Kahn & Fourme (1997)]. The final evolution of the spherical drift MWPC was the Mark III. With respect to the Mark II, the new feature was a position encoder designed for fast determination of the centroid of electron clusters resulting from the conversion of X-ray photons. This encoder improved both spatial resolution and uniformity (Bahri *et al.*, 1992). The Mark II was installed on the D41 station at LURE-DCI and open to users under the supervision of A. Bentley. It was used for data collection at an essentially fixed wavelength ($\sim$1.4 Å). Then, image-plate detectors which were simpler to operate and maintain became available, and the Mark III operation was stopped in 1993.

2.2. *The EEC detector software initiative*

The *EEC MADNES* software package was the product of an EEC initiative organized in 1986 by G. Bricogne (Bricogne, 1986*a*,*b*, 1987*b*) for the purpose of implementing, through a series of cooperative workshops, a single program structure providing device independence, computer independence without loss of efficiency, and method independence in the data reduction. This implementation was carried out starting from the *MADNES* package (Messerschmidt & Pflugrath, 1987) which was at the time specific for the Enraf–Nonius FAST system. Many specialists participated in the various workshops. This EEC initiative had a substantial and irreversible impact on the evolution of position-sensitive detector software: before the workshops, each package was tied to a particular combination of detector and goniometers; after the workshops, many of the packages were upgraded to process data from many different such combinations. Several original concepts emerged during these workshops, notably the idea of three-dimensional auto-indexing by Fourier transform methods (Bricogne, 1986*b*, p. 28) which was subsequently implemented in the *DENZO* package (Otwinowski & Minor, 1997).

The scientific manifesto of the workshops (Bricogne, 1986*a*) proposed a long-term approach to data acquisition in terms of Bayesian statistics, whose methods were invoked in particular to estimate and update three-

dimensional reflection profiles and to carry out reflection integration itself. This Bayesian approach was viewed as seamlessly connected to the subsequent steps of phase determination and structure refinement which were later described under the name of 'Bayesian programme' (Bricogne, 1988, 1993). This unified picture of the statistical underpinnings of crystal structure determination, from raw data acquisition through phasing to the final structure refinement, still holds much untapped potential.

2.3. *First results on the MAD method*

One of the goals of the MWPC project was the implementation of the multiwavelength anomalous diffraction method (MAD). A research project on MAD began at LURE in the late 1970s. This development was triggered by the pioneering work at SSRL (Stanford) in several fields, including EXAFS instrumentation; first experiments on anomalous diffraction performed with a precession camera (Phillips *et al.*, 1977); the discovery of large resonances at the *L*-absorption edges of cesium and the associated large magnitudes and changes in f' and f'' which were potentially very attractive for phasing applications (Phillips *et al.*, 1978); and the determination of MAD phases by a MIR-like approach (Phillips & Hodgson, 1980). We established a collaboration with O. Dideberg at Liege for the determination of the unknown structure of parvalbumin IIIf from the superfast swimbladder muscle of a fish, *Opsanus tau*. This protein was chosen for several reasons: it is a small molecule (10.1 kDa), which binds two Ca^{2+} ions which can be substituted by lanthanides; the space group is $P2_12_12$ with one molecule in the asymmetric unit and good crystals of the terbium complex, diffracting to at least 1.7 Å, were available. Diffraction experiments were performed at the terminal station, D15, of the bending-magnet beamline D1. The set-up included a Ge(220) channel-cut monochromator, a single-axis goniometer from an Enraf–Nonius camera and the Mark I detector. A crystal was accurately aligned along the *b*-axis in order to measure Bijvoet mates at the same time. Three data sets at a resolution of 2.3 Å were measured in 36 h at wavelengths close to the L_{III}-absorption edge of terbium. These wavelengths were selected in order to minimize average phase errors (Narayan & Ramaseshan, 1981): at the peak (f'' maximum), at the inflection point of the f'' curve ($|f'|$ maximum) and 0.0021 Å below the peak wavelength. Diffraction images were recorded over small angular intervals (20 images per degree), with the wavelength kept constant through 15° sectors. Coordinates of lanthanide ions were determined from Harker sections of an anomalous Patterson-difference map at the peak wavelength and refined against the strongest anomalous differences (Hendrickson & Teeter, 1981). One fully occupied and one minor site were found. Phasing of acentric reflections was performed in a pseudo-MIR mode with a program written by R. Fourme. The (*hkl*) set at the peak wavelength was selected as reference, and the total phase probability distribution $P(\varphi)$ was taken as the product of distributions formed from the five pairings of the reference set with the remaining sets (*hkl* and $h\bar{k}l$) at the various wavelengths. Centric reflections were determined by a method where the various data sets are treated on an equal footing and which considers phases and amplitudes of the wavelength-independent part of structure factors as random variables. This method was later extended to acentric reflections (Chiadmi *et al.*, 1993). In a MAD experiment it is essential to derive accurate values for f' and f'' at the various wavelengths. In the parvalbumin study, initial f'' values were determined by measuring, in the wavelength range of interest, the difference in intensities of mates in a selected Bijvoet pair; f' values were obtained by numerical integration *via* the Kramers–Kronig dispersion relation. These values were refined by minimization of the total lack of closure over all reflections. The MAD electron density based on 2439 phased reflections was calculated for the two enantiomorphs and one of these maps allowed an unambiguous chain tracing. Independently, O. Dideberg solved the structure at 3.2 Å resolution by molecular replacement, using the carp parvalbumin molecule (Moews & Kretsinger, 1975) as model. The average phase discrepancy for reflections common to both determinations is 54°, prior to any refinement of either structure. The structure of *Opsanus tau* parvalbumin (Kahn *et al.*, 1985; Kahn, Fourme, Bosshard, Chiadmi *et al.*, 1986) is the first unknown structure determined by MAD phasing (Fourme & Hendrickson, 1990), although the folding was of a known type. The use of an electronic area detector with a short dead time allowing a fine slicing of reciprocal space was very innovative. Nevertheless, the accuracy and completeness of data collected during this experiment were not optimal, due to inhomogeneities in the response of the various detector channels and limitations in both detector hardware and data analysis software. These problems were essentially solved with the installation in 1984 of the Mark II detector, coupled to a multicircle diffractometer, on the D23 station (Kahn, Fourme, Bosshard, Chiadmi *et al.*, 1986). After the implementation of *EEC MADNES*, the progress with respect to the pioneer parvalbumin experiments can be best appreciated with the solution by MAD of the structure of the carbohydrate recognition domain of a rat mannose-binding protein (Weis *et al.*, 1991). The asymmetric unit contains a dimer in which the four native Ca^{2+} ions can be substituted by lanthanides. A crystal from the holmium complex was used for a three-wavelength MAD experiment at the Ho L_{III}-absorption edge. Values for the observed anomalous diffraction ratios $\langle\Delta|F|^2\rangle/\langle|F|^2\rangle$ (where $\Delta|F|$ is the Bijvoet difference at each wavelength) for centric reflections can serve as an estimate of the noise in the anomalous signals. These values (1.9, 2.1 and 1.7% at the remote, peak and edge wavelengths, respectively) as well as the merging Rsym value on intensities for the remote wavelength data set (3.1%) are very low, even according to current standards, especially as the whole data were obtained from a single sample kept at ∼277 K (not cryocooled) for several days. Monochromator instability

during data collection produced a variability of edge and peak Ho^{3+} scattering factors. This problem was overcome by a least-squares refinement of scattering factors by blocks of data. It drew our attention to the fact that small changes in f'' and f' values due to wavelength drifts produced subtle non-isomorphism; accordingly, the refinement of f'' and f' values as parameters of the anomalous structure had to be included in a fully fledged treatment, which was achieved a few years later with the program *SHARP* as discussed in §5. Finally, a unique set of 7195 reflections (completeness 89.3%) was obtained. The phasing procedure was performed by the algebraic method (Karle, 1980; Hendrickson, 1985) using *MADLSQ* (Hendrickson, 1985). The $|^0F_A|$ Patterson map was extremely clear, and the quality of the MAD-phased electron density map allowed the construction of a molecular model with no significant deviations from the ideal geometry which was refined to a crystallographic *R*-factor of 0.176. This example demonstrated the feasibility of obtaining very accurate MAD phases.

W. Shepard and R. Fourme installed in 1994 a new MAD station, DW21, on the wiggler beamline of the storage ring DCI. The equipment includes double-focusing optics with a curved mirror and a sagittal-focusing monochromator, a commercial image-plate detector, a cryocooling system and a system for the detection of fluorescence based on an Si diode. With typical exposure times in the range 15–300 s per degree and the fairly long dead time of the image-plate detector, DW21 cannot compete with CCD-based instruments at third-generation synchrotron sources for data-collection rates. In contrast, the accuracy of data collected at DW21 may be excellent. This is due to the stability of the operating conditions of DCI (broad positron beam, beam lifetime in excess of 220 h) and to the care taken at the various steps of data collection and analysis, as illustrated in §6 (Schiltz *et al.*, 1997*a*). Representative examples of structures solved by MAD using data collected at DW21 and which led to high-quality electron density maps are: a 'zipper' DNA (Shepard *et al.*, 1998), a double-stranded DNA containing a cisplatin interstrand cross-link at 1.63 Å resolution (Coste *et al.*, 1999), and cytochrome 552 from *Pseudomonas nautica* with eight molecules in the asymmetric unit (Brown *et al.*, 1999); the anomalous scattering species were Br, Pt and Fe, respectively.

3. From MAD to MASC

In the MAD method the unknown structure contains many ordered atoms which behave essentially as normal scatterers and a few ordered atoms with significant anomalous scattering (the partial structure *A*) which are, or can be, inserted into the total structure. Instead of being ordered, the partial structure *A* can be disordered atoms dispersed in the solvent which, in macromolecular crystals, occupies a large fraction of the unit-cell volume. This method is a way of producing contrast variation (Stuhrmann & Kirste, 1965). It is different from other contrast-variation methods (such as chemical or isotopic substitution, solvent exchange and nuclear spin-dependent scattering), since it contains an imaginary part of the scattering amplitude. There were few references concerning the measurement and the exploitation of Bijvoet differences due to anomalous scattering effects of the solvent in macromolecular crystals and they refer essentially to single-wavelength cases. Such differences have been observed by H. W. Wyckoff (unpublished), Dumas (1988), Crumley (1989) and Carter *et al.* (1990). Bricogne and co-workers (Dumas, 1988; Carter *et al.*, 1990) have pointed out that it is possible to take advantage of Bijvoet differences to supplement chemical standard contrast-variation measurements. The potential usefulness of multiwavelength anomalous scattering of the solvent phase was mentioned in the context of the 'Bayesian programme' (Bricogne, 1993). A MAD-like analysis was presented (Fourme, 1993; Fourme *et al.*, 1995) using, as a starting point, a formalism which separated the diffraction effects of the molecular envelope and the internal fluctuations (G. Bricogne, unpublished).

Define:

$$\Gamma(\mathbf{h}) = -{}^0\rho_{sA}\,G(\mathbf{h})\exp(-BS^2/4), \tag{1}$$

where ${}^0\rho_{sA}$ is the 'normal' electron density of anomalous scatterers, $G(\mathbf{h})$ is the structure factor value at $\mathbf{h}$ of the molecular envelope, B is a parameter reflecting the disorder in the solvent phase and S is the modulus of the scattering vector. One can derive expressions for the overall structure factor $F(\mathbf{h})$ and the complex conjugate $F(-\mathbf{h})^*$ of the structure factor of the anomalous mate similar to the ones in the algebraic MAD method (Hendrickson, 1985), where $\Gamma(\mathbf{h})$ replaces ${}^0F_A(\mathbf{h})$:

$$F(\mathbf{h}) = {}^0F(\mathbf{h}) + ({}^\lambda f'/f^0 + i\,{}^\lambda f''/f^0)\Gamma(\mathbf{h}), \tag{2}$$

$$F(-\mathbf{h})^* = {}^0F(\mathbf{h}) + ({}^\lambda f'/f^0 - i\,{}^\lambda f''/f^0)\Gamma(\mathbf{h}). \tag{3}$$

The separation of the effects of the anomalous partial structure *A* from the overall diffraction effects can be applied using a set of equations equivalent to the algebraic MAD equations. Using *MADLSQ* (Hendrickson, 1985), these equations can be solved for $|\Gamma(\mathbf{h})|$, $|^0F_T(\mathbf{h})|$ and the phase difference $\Delta\varphi$ between ${}^0F_T(\mathbf{h})$ and $\Gamma(\mathbf{h})$. Accordingly, the prime information which can be derived from MASC data is the set of amplitudes of envelope structure factors. The calculation of the Fourier transform (the molecular envelope) requires the derivation of phases associated with these amplitudes. This is a non-trivial problem, which we are currently addressing by several methods (Ramin, 1999; R. Kahn, unpublished), including the use of maximum entropy and likelihood ranking to test envelope hypotheses (Bricogne, 1993).

This method is called MASC (an acronym for multiwavelength anomalous scattering contrast, suggested by W. Shepard). It has an advantage over other contrast-variation methods since the contrast variation is generated by inducing a physical change, and thus preserving strict isomorphism. A variety of anomalous scatterers may be used in a MASC experiment, and the most suitable ones

will depend on the crystallization conditions of the macromolecule. Analogues of the precipitating agent are good choices since such compounds are less likely to perturb the crystalline lattice (*e.g.* selenate for sulfate, bromide for chloride, tribromoacetate for acetate) although changes of pK_a should be allowed for. Pioneer MASC experiments were performed at LURE-DCI (Fourme *et al.*, 1995) on crystals of two proteins with known structures. Further experiments were conducted at LURE-DCI and at the ESRF (Grenoble), using cryocooled crystals of three proteins with very different molecular weights and various anomalous scatterers [ytterbium, bromine and selenium in $YbCl_3$, NaBr and $(NH_4)_2SeO_4$, respectively]. For the three structures, a few ordered sites were found on phased anomalous Fourier difference maps. These sites are located close to the protein surface. In such cases the total 'normal' structure factors of the partial structure incorporate a contribution from ordered anomalous scatterers. Respective contributions from ordered and disordered anomalous scatterers vary according to resolution. A good agreement between experimental and model amplitudes was found taking into account MASC effects only at very low resolution ($d > 20$ Å), a superposition of MASC and MAD effects in the range 20–8 Å, and MAD effects only at higher resolution. Accordingly, MAD and MASC effects may be jointly used to derive phase information in different resolution ranges. Furthermore, the addition of a high concentration of anomalous scatterers in the solvent may be purposely and systematically exploited to obtain derivatives suitable for the SAD or MAD methods. In other experiments one might wish to obtain a pure MASC effect. We have found that binding the anomalous scatterer to a zwitterion was effective in this respect, as significant ordered sites could not be detected.

Experimental aspects of a MASC experiment are now essentially in hand. The comparison between experimental and model values has shown that amplitudes from envelope structure factors $G(\mathbf{h})$ derived from MASC data are meaningful (Ramin *et al.*, 1999). The future of this method is strongly dependent on the availability of an efficient procedure for phasing these coefficients. If this phasing step can be dealt with, then the combination of anomalous diffraction and contrast variation techniques can lead to a general method for low-resolution phasing of diffraction patterns from very large macromolecules. A different approach to low-resolution phasing, involving the location in the crystallographic unit of the molecular shape determined from solution scattering, has been introduced recently (Hao *et al.*, 1999).

4. Xenon and krypton as heavy-atom labels and anomalous scattering centres

4.1. *Binding of xenon and krypton to proteins*

Early experiments with xenon gas in protein crystallography go back to the work of Schoenborn *et al.* (1965) and Tilton *et al.* (1984) on myoglobin–xenon complexes. The general principles of xenon binding to proteins are reviewed by Schiltz *et al.* (1994) but a brief outline is given here. Xenon complexes with proteins are obtained by subjecting a native protein crystal to a xenon-gas atmosphere pressurized in the range 1×10^5–50×10^5 Pa. Xe atoms are able to diffuse rapidly towards potential interaction sites in protein molecules *via* the solvent channels that are always present in protein crystals. The number and the occupancies of xenon binding sites vary with the applied pressure. The interaction of xenon with proteins is the result of non-covalent weak-energy van der Waals forces and therefore the process of xenon binding is completely reversible. It also implies that xenon binding only induces very marginal perturbations to the surrounding molecular structure. As a consequence, these xenon complexes are almost always highly isomorphous to the native crystals. Xenon is able to bind to a large variety of sites in proteins (Prangé *et al.*, 1998), including closed intramolecular hydrophobic cavities, accessible active sites, intermolecular cavities and channel-pores. Essentially the same interactions as for xenon also exist with krypton, and protein-krypton complexes can be prepared in the same way (Schiltz *et al.*, 1997*a*).

4.2. *A brief historical survey*

The use of xenon as a heavy atom for phase determination was first suggested by Schoenborn & Featherstone (1967) who immediately realised its potential benefits by stating that "Xenon is a little 'lighter' than desirable for a heavy atom, but this is counteracted by the fact that xenon protein complexes show a high degree of isomorphism with the native crystals – a fact often not true with most of the commonly used 'heavy atoms' which are generally ionic groups capable of inducing some disorder into the native structure." However, it was not until 1991 that Vitali *et al.* (1991) demonstrated that SIRAS-phases computed from the xenon complex of myoglobin yielded an interpretable electron density map for that protein. Following a suggestion by T. Prangé, a comprehensive research project on xenon derivatives was started at LURE in 1993 with three specific goals in mind: (i) devising a simple and generally applicable method for the preparation and X-ray data collection of isomorphous xenon derivatives; (ii) testing this method on proteins other than myoglobin and haemoglobin; and (iii) using xenon derivatives for determining phases of unknown protein structures. A first outcome of this work was the design of a high-pressure xenon cell and the setting up of a standard protocol for data collection under xenon gas pressure (Schiltz *et al.*, 1994). Furthermore, it was shown that, apart from the well known cases of the haem-proteins, xenon is able to bind to a large number of different proteins (Schiltz *et al.*, 1995; Prangé *et al.*, 1998).

The first xenon derivative on a protein of unknown structure was obtained on urate oxidase in 1993 [the structural study was published in 1997 by Colloc'h *et al.* (1997)]. A highly substituted xenon derivative was

obtained at a gas pressure of 8×10^5 Pa. Mercury and lead derivatives had been obtained previously for this protein, but the binding sites in both derivatives are at almost identical locations, thus limiting their phasing power. The contribution of the xenon derivative was crucial for the successful determination of this structure by the MIR method: the single xenon binding site is at a large distance from the heavy-atom sites in the two classical derivatives. This emphasizes another advantage of xenon derivatives; because xenon tends to fix into hydrophobic cavities, these binding sites are very likely to be different from those of classical heavy-atom compounds which predominantly bind to specific functional groups.

The real breakthrough came in 1994, when the structure of the ligand-binding domain of the human nuclear receptor RXR-α was solved with a xenon derivative prepared at LURE at a gas pressure of 20×10^5 Pa (Bourguet *et al.*, 1995). This structure was solved with an initial 5 Å SIR map, based solely on the xenon derivative. The α-helices of the structure were clearly visible in this initial map and a discontinuous polyalanine model could be traced. Density modification *via* solvent-flattening and recombination with the SIR phases yielded an interpretable electron density map for the whole molecule. At a latter stage, a classical mercury derivative was obtained and these data were added to the SIR phases to improve the map. Because of its high isomorphism, the xenon derivative has a significantly better phasing power than the mercury derivative. This was the first published example where a xenon derivative had been used to determine the phases of an unknown protein structure. It established the credentials of the method and it immediately triggered the attention of the crystallographic community.

4.3. *Recent developments and current status*

Since then, xenon derivatives have been successfully used for the structure determination of a number of other proteins, including Photosystem I (Krauss *et al.*, 1996), cartilage oligomeric matrix protein (Malashkevich *et al.*, 1996), DMSO-reductase (Schindelin *et al.*, 1996), the outer-surface antigen P64k from *Neisseria meningitidis* (Li de la Sierra *et al.*, 1997), the CheY-binding domain of histidine kinase CheA (Welch *et al.*, 1998), *N*-myristoyl transferase (Weston *et al.*, 1998), Arcelin-1 (Mourey *et al.*, 1998) and thermostable β-mannanase (Hilge *et al.*, 1998).

Meanwhile, alternative designs for room-temperature pressure cells had been proposed (Stowell *et al.*, 1996) and the problem of flash-freezing xenon derivatives was addressed by Sauer *et al.* (1997), Schiltz, Shepard *et al.* (1997) and Soltis *et al.* (1997) who showed that the pressurization and flash-freezing steps can be separated in time. Sauer *et al.* (1997) also developed two elegant devices for the production of freeze-trapped xenon derivatives. Alternative designs of such devices were presented by P. Mancia & R. P. Evans (unpublished), Soltis *et al.* (1997) and Djinovic-Carugo *et al.* (1998).

Xenon and krypton also display interesting anomalous scattering properties (Schiltz, Shepard *et al.*, 1997; Schiltz, Kvick *et al.*, 1997). The *K*-absorption edge of krypton is situated at a wavelength (0.866 Å) that is readily accessible on synchrotron radiation sources and that is near the selenium and bromine edges. The absorption edges of xenon are situated either at long or very short wavelengths (2.6 Å for the $L_{\rm III}$ edge and 0.36 Å for the *K* edge), which are more difficult to harness technically, but at the Cu $K\alpha$ wavelength the residual anomalous signal from the $L_{\rm III}$ edge is still significant: $f'' = 7.4$. Thus, SIRAS experiments on xenon derivatives with Cu $K\alpha$ radiation were successfully carried out by Vitali *et al.* (1991) and Weston *et al.* (1998).

In two SIRAS test studies (Schiltz, Shepard *et al.*, 1997; Schiltz, Kvick *et al.*, 1997), carried out respectively on krypton and xenon derivatives of the protein porcine pancreatic elastase (PPE), the potentials of highly isomorphous noble-gas derivatives were evidenced. Perhaps even more significant, these studies strikingly demonstrated the crucial importance of the coupling between careful data-collection schemes designed to accurately harvest weak signals with the optimal statistical treatment of these data, as will be discussed in more detail in §6.

For many MIR projects, testing for xenon binding is now included as a routine part in the screening for derivatives. From the experience gathered at LURE and at SSRL (Stowell *et al.*, 1996), one can estimate conservatively that the success rate for xenon binding is higher than 30%. Many protein crystallography groups have their own in-house pressurization apparatus and equipment for the preparation and X-ray data collection of xenon derivatives is provided to users on the crystallography beamlines at a number of synchrotron centres including LURE Orsay,† SRS Daresbury‡ and SSRL Stanford.§ As a latest development in the field, xenon-derivatization flash-cooling devices are being commercialized by companies, including the Molecular Structure Corporation and Oxford Cryosystems (in collaboration with the University of Oxford).

Apart from their use in phase determination, xenon derivatives have also been used or proposed for other interesting applications such as the mapping of hydrophobic cavities in proteins (Tilton *et al.*, 1984; Quillin *et al.*, 1996; Prangé *et al.*, 1998), the exploration of pathways for gas access in hydrogenases (Montet *et al.*, 1997) or for X-ray contrast-variation imaging of lipids and detergents in crystals of membrane proteins (Sauer *et al.*, 1997).

5. *De novo* phasing using maximum-likelihood refinement

De novo phasing methods rely entirely on measurements of Bragg intensities. They are used to solve the phase

† http://www.lure.u-psud.fr/sections/Xenon/XENON.HTM.
‡ http://www.dl.ac.uk/SRS/PX/xenon cell/xenon notes 1.html.
§ http://biosg1.slac.stanford.edu/research/workshop97/xenon.

problem in macromolecular crystallography when molecular replacement is not applicable. For each diffracted beam **h**, the phase of the structure factor of the unknown structure is estimated with respect to the phase of the structure factor of a subset of atoms (reference structure). In order to use the *calculated* phase of the wave diffracted by the subset structure as reference, its structure must be solved and refined prior to the solution of the main structure. This can be achieved if atoms in the subset have distinct features – a high Z value and/or anomalous scattering – with respect to all other atoms, so that they can be located using Patterson or direct methods. In order to determine unambiguously the unknown phase, at least two partial structures and three intensity measurements are required. Several variants of the same concept can be used for the generation of reference waves: the isomorphous replacement (MIR), the wavelength-dependence of scattering from anomalous scatterers (MAD) or a blend of two (SIRAS, MIRAS). With two intensity measurements only (SIR, SAD) the phase determination is ambiguous; additional physical information, such as solvent flattening or non-crystallographic symmetry, is necessary to break the ambiguity.

Bias-free refinement of partial structures in *de novo* methods, which is an essential step towards obtaining the best possible electron-density maps given the available data, has remained for a long time a troublesome issue in macromolecular crystallography. The origin of difficulties was identified in the use of least squares in a situation where several rules which are cardinal for their application are not respected, the cure being the use of maximum-likelihood refinement as advocated by Bricogne (1985, 1988, 1991*a*,*b*). The first rule is that any quantity involved in the observational least-squares equations should be either a model parameter or an observation. Treating the native phase as a known constant within each refinement cycle but recalculating it after each refinement step introduces bias on the parameters, especially in the case of mostly bimodal distributions. The second rule is that the inverse-variance 'weights' in the expression for the least-squares residual should be kept fixed as if they were part of the observed data. Since the method of least squares is a special case of the maximum-likelihood methods when errors are normally distributed with fixed (co)variances, it was clear that the problem of properly estimating the lack-of-isomorphism parameters demanded a full-fledged maximum-likelihood treatment rather than least squares. The outline of maximum-likelihood formalism for the isomorphous replacement (MIR and SIR case) was first given then extended to probability distributions incorporating anomalous diffraction effects as well as measurement error and non-isomorphism (Bricogne, 1991*a*,*b*). Integrating the distributions in the whole complex plane leads to likelihood functions that can be used for heavy-atom detection and refinement, and for producing phase probability distributions encoded with the '*ABCD*' coefficients (Hendrickson & Lattman, 1970).

This formalism has been implemented in a computer program called *SHARP*, an acronym for statistical heavy-atom refinement and phasing (de la Fortelle & Bricogne, 1997). In *SHARP*, the parametrization amounts to a physical description of diffraction properties, involving heavy-atom coordinates, occupancies, isotropic and (if need be) anisotropic temperature factors, as well as normal and anomalous scattering factors. A parametrization of the anisotropy of anomalous scattering has recently been implemented, allowing a refinement of the corresponding parameters from unmerged data carrying suitable goniometric information for each measurement (M. Schiltz & G. Bricogne, unpublished). The implementation of *SHARP* uses a hierarchical organization for the various parameters, that enables common attributes to be shared appropriately. A list of site coordinates is determined that contains all known sites in all derivatives and, for each level of the hierarchy, these sites are 'qualified' (by a chemical identity, by an occupancy *etc.*). In this way the long-standing problem of the same site being refined independently at each wavelength of a MAD experiment cannot occur, and common sites in a MIR experiment are parametrized correctly. As *SHARP* can accomodate data from different experimental procedures (MIR, MAD or a blend of the two), the user has to be guided during the build up of a hierarchical parameters file describing this experiment. This is achieved by means of an HTML browser-based graphical user interface. An extensive on-line help is also available.

We shall point out the way in which some important issues are adressed by *SHARP*. The use of two-dimensional integration over the full (amplitude and phase) native structure factor in the likelihood function removes, in the case of SIR and MIR, the assumption that the native amplitude measurement is error-free and, in the case of MAD, respects the statistical equivalence between the data attached to the various wavelengths. The randomness caused by lack-of-isomorphism is treated as part of the substitution model. As a result, it does not attempt to match a calculated value to an observed value (as in the case of the least-squares method) but to match the probability distributions of these two quantities. Finally, the likelihood formalism provides the opportunity of checking for significant disagreement between the data and the substitution model. For each reflection **h**, the gradients of the log-likelihood function with respect to the real and imaginary parts of the various heavy-atom structure factors $F_{Aj}(\mathbf{h})$ are calculated. These numbers are then used in Fourier syntheses to produce residual maps, that have the symmetry of the crystal. Similarly, in the case where there is significant anomalous diffraction, the gradients with respect to $(F_{Aj+} + F_{Aj-})$ become coefficients for isomorphous residual maps and those with respect to $(F_{Aj+} - F_{Aj-})$ for anomalous residual maps. These residuals maps are essentially Fourier syntheses calculated from inverse-variance weighted difference coefficients between the derivative and native data. Their enhanced sensitivity to

any departure from the current heavy-atom model makes them an instrument of choice to detect minor sites, structural disorder at certain sites, or anisotropy in the heavy-atom temperature factors.

SHARP has been extensively tested in order to make sure that the refinement procedure converges to an unbiased solution. For this purpose, simulated data files were created in which the 'perfect' heavy-atom parameters are specified, and the lack-of-isomorphism and measurement errors have ideal distributions with known parameters. The refinement is initiated from perturbed values of all these parameters, and the accuracy of their final refined values is measured by a suitable 'distance' to the perfect distance in parameter space. Numerical tests for SIR, SIRAS and MAD have shown, for the first time with a heavy-atom refinement procedure, that the estimated standard errors on the refined parameters are quantitatively reliable. MIRAS tests have also shown that several weak heavy-atom derivatives could be refined with final unbiased parameter values, while each of them was too weak for a pure Patterson-based refinement. First results on measured crystallographic data compared favourably with those of other programs, giving indications of significant overall improvement with *SHARP* (Ramakrishnan & Biou, 1997). *SHARP* has now been installed and effectively used at several hundred sites throughout the world. Close interaction between *SHARP* developers and users around a variety of difficult problems has been an effective driving force for the continued improvement of the program.

The formalism used by *SHARP* applies to the various *de novo* phasing methods and underlines their common conceptual background, while being flexible enough to take into account their specificities (for instance, the site invariance in the MAD method). Furthermore, *SHARP* uses and produces phase information in a form which is most suitable for the communication with, in particular, the *BUSTER* suite which implements the 'Bayesian program' (Bricogne, 1993). The accurate description of reference structures and the production of minimally biased phases are especially useful for difficult cases with small signals, noisy or impoverished data. *SHARP*, in its standard implementation, is interfaced to the density-modification program *SOLOMON* (Abrahams & Leslie, 1996) which was instrumental in the solution of the F1-ATPase structure. There is ample evidence that minimally biased initial phases produced by maximum likelihood improve the efficiency of density modification algorithms. Even in cases of complete bimodality, *SOLOMON* selects the correct mode for most phases and is able to produce high-quality phases.

6. Selected examples

6.1. *Test SIRAS phasing of a 26 kDa protein with half a krypton atom*

X-ray diffraction data at the high-energy side of the krypton *K*-edge (0.86 Å) were collected on a crystal of pork pancreatic elastase (PPE) put under a krypton gas pressure of 56×10^5 Pa (Schiltz, Shepard *et al.*, 1997). The occupancy of the single krypton atom is approximately 0.5, giving isomorphous and anomalous scattering strengths of 15.2 and 1.9 electrons, respectively. This derivative was used successfully for phase determination with the SIRAS method. After phase improvement by solvent flattening, the resulting 1.87 Å electron density map is of exceptionally high quality, and has a correlation coefficient of 0.90 with a map calculated from the refined native structure. Over 50 water molecules appear as spherical peaks in this map, which is based purely on experimental data. That the equivalent of half a krypton atom is sufficient for the structure determination of a 26 kDa protein is in itself a remarkable outcome. Apart from the near-perfect isomorphism between the krypton derivative and the native structure, the key steps in the achievement of this result were as follows:

(i) High-quality data. Data were collected on the DW21 beamline at LURE on an imaging plate. Parameters were carefully chosen so as to collect highly redundant data sets with minimal noise and without detector saturation effects, even at low resolution.

(ii) Reducing systematic errors in the isomorphous differences. This was accomplished by collecting native and derivative data on the same crystal (before and after pressurization, respectively), in the same orientation. Residual systematic errors were reduced by parametrized local scaling of derivative to native data which also proved to enhance the accuracy of the anomalous differences.

(iii) Optimization of the anomalous signal. This was achieved by tuning the energy of the X-rays to 14.361 keV, just past the krypton *K*-edge (on the high-energy side), to optimize the anomalous signal (*i.e.* to maximize f''). Also, the crystal was pre-oriented in such a way that Bijvoet mates were collected on the same frame. Even though the anomalous signal is quite weak, the SIRAS phases are of significantly higher quality than phases computed in SIR mode (these give a final map correlation coefficient of 0.72).

(iv) Optimal statistical treatment of isomorphous and anomalous differences. The use of the program *SHARP* for heavy-atom refinement and phasing played an essential role in the successful determination of the high-quality electron density map. Initial attempts to use the program *MLPHARE* gave rise to problems with the treatment of the weak anomalous signals. The final solvent-flattened SIRAS map which was calculated from *MLPHARE* phases is of lower quality and has a correlation coefficient of 0.69 with a map computed from the refined native structure. This study was one of the very first applications of *SHARP* to real data and it gave the first indications of the superiority of a complete implementation of the maximum-likelihood theory for heavy-atom refinement over previous approaches. Thus, the use of xenon and krypton derivatives, when they can be obtained, associated with statistical heavy-atom refinement allows one to overcome the major limitations of

the isomorphous replacement method, *i.e.* non-isomorphism and the problem of optimal estimation of heavy-atom parameters.

A similar experiment, which was carried out at the ESRF, was also performed on a xenon derivative of PPE, with data collected at the high-energy side of the xenon *K* edge ($\lambda = 0.36$ Å) (Schiltz, Kvick *et al.*, 1997). It essentially confirms the conclusions drawn in the study with krypton. This is the first fully documented report on a complete protein crystallography experiment, from data collection up to phase determination and calculation of an electron density map, carried out with data obtained at ultrashort (∼0.3 Å) wavelengths. Potential advantages of the use of very short (∼0.5 Å) or ultrashort wavelengths in macromolecular crystallography have already been addressed by Helliwell & Fourme (1983), Helliwell (1992) and Helliwell *et al.* (1993).

6.2. *N-myristoyl transferase*

In many respects the structure determination of *N*-myristoyl transferase (Weston *et al.*, 1998) also proves to be a typical illustration of the advantages that highly isomorphous xenon derivatives, combined with mindful data-collection schemes and with a proper statistical heavy-atom refinement and phasing, can offer. A three-wavelength MAD dataset of standard quality was collected on an Se-Met protein at the ESRF, but the determination of the positions of the 12 Se atoms proved intractable. At a latter stage a flash-frozen xenon derivative (pressure 10^6 Pa) was prepared and data were collected on an in-house X-ray generator with Cu $K\alpha$ radiation. A second crystal was subjected to exactly the same treatment, but the xenon gas was allowed to diffuse out of the crystal before cryocooling, thus giving a native dataset. The authors report that the relative weakness of the isomorphous differences ($R_{\rm iso}$ = 0.091) almost led them to discard the xenon derivative as being native. However, the isomorphous difference Patterson map showed clear cross-vectors and ten xenon sites could be refined in *SHARP*. The resulting SIRAS phases allowed for the straightforward detection of the positions of the 12 Se atoms by difference Fourier analysis. Most surprisingly, the *SHARP*-refined xenon SIRAS phases alone (after solvent-flattening) were of better quality than both the combined Xe/Se phases and the pure Se-MAD phases (after solvent-flattening). Eventually, the structure was solved from the solvent-flattened xenon SIRAS map. The authors also report that various attempts to derive SIRAS phases with the program *MLPHARE* (combined with solvent-flattening) failed to produce an interpretable electron density map.

6.3. *SAD phasing with weak anomalous signals*

The possibility of solving a protein structure from anomalous scattering at a single wavelength (SAD) was first achieved on crambin (Hendrickson & Teeter, 1981). Rusticyanin is the largest structure solved from SAD data to date (Harvey *et al.*, 1998). The lack of information inherent to SAD alone was completed by information coming either from a Sim distribution (crambin) or direct methods (rusticyanin). A test SAD experiment on native tetragonal crystals of hen egg-white lysozyme (HEWL) was performed by Dauter *et al.* (1999). The data collection with an image-plate detector at the NSLS (Brookhaven) was standard, except for the high multiplicity and high completion of observations resulting mainly from multiple measurement passes with different limits and exposure times. The structure was solved from the weak anomalous signal of ten native S atoms and seven chlorine ions from the crystallization solution loosely bound to the protein surface (expected Bijvoet ratio at 1.54 Å, ∼1.5%). The partial structure of anomalous scatterers was solved by direct methods on the basis of Bijvoet differences with the program *SHELXM* (Sheldrick, 1998). From the initial bimodal phase distribution, the application of *SHARP* and *SOLOMON* produced an electron density map highly correlated (correlation factor 0.80) to the electron density of the refined structure.

7. Conclusions

The developments presented in this contribution are representative of the general evolution which has taken place in macromolecular crystallography over the past two decades. One major aspect of this evolution has been the drive towards ever higher standards of data quality. The combination of cryocooling techniques, of third-generation synchrotron sources and of advanced detector systems has recently achieved astounding data-collection rates, especially in the context of experimental protocols developed for MAD data collection. It is now relatively easy and no longer so time-consuming to collect accurate highly redundant and complete data (including very low-resolution reflections) up to the highest possible resolution. We believe, however, that much work remains to be performed in the design of optimal data-collection strategies capable of delivering not only completeness but also a spatially homogeneous degree of redundancy.

Another aspect of the recent evolution of methods is that data collection and subsequent computational treatment must be considered as two complementary contributors to the phasing process. The MAD method is an obvious example of this interaction, but it is by no means an immutable paradigm. Besides classical heavy-atom compounds and seleno-methionine incorporation, xenon and krypton have added to the range of possibilities available to crystallographers to create a partial structure for phasing purposes; and the systematic implementation of a more rigorous treatment (by maximum-likelihood) of the small signals emanating from such partial structures has allowed a qualitative jump in the quality of maps which can be obtained for *de novo* structure determination. We feel that these developments will encourage greater diversity in data-collection strategies than the standard MAD paradigm does: for instance, if a full and accurate MAD

experiment cannot be performed for practical reasons, it may be advisable to measure only a single dataset, at the wavelength for which the anomalous signal is largest, with particular care, in view of the mounting evidence that SAD may be sufficient to solve structures, in combination with appropriate computational procedures such as *SHARP* and *SOLOMON*. In the perspective of structural genomics, where the average time required to solve a structure will have to be drastically reduced, such considerations are likely to play a key role in streamlining the time-consuming steps of structure determination such as tracing the map and refining the molecular model.

The synergy between the respective contributions of the various co-authors of this article has been sustained by a shared belief that progress in the development of new mathematical approaches to the phase problem, on the one hand, and in the design of new experimental protocols, on the other, occurs most effectively when these two activities are guided by each other.

All co-authors are, or have been, members of the LURE team. Writing this article, dedicated to the Nobel Prize of John Walker, was also a friendly opportunity to evoke some memorable episodes. It is also a pleasure to acknowledge the collaboration of A. Bentley at LURE and of G. Charpak, R. Bouclier, R. Benoit, G. Million, J. C. Santiard and F. Sauli at CERN, who were faithful companions along the MWPC adventure.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.

Bahri, A., Bosshard, R., Santiard, J. C., Kahn, R. & Fourme, R. (1992). *Rev. Sci. Instrum.* **63**(1), 655–658.

Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. (1995). *Nature (London)*, **375**, 377–382.

Bricogne, G. (1985). *Bischenberg Conference on the Crystallography of Molecular Biology.* Unpublished.

Bricogne, G. (1986*a*). Editor. *Proceedings of the EEC Cooperative Workshop on Position-Sensitive Detector Software – Phase 1*, 26 May–6 June 1986, LURE, Orsay, France.

Bricogne, G. (1986*b*). Editor. *Proceedings of the EEC Cooperative Workshop on Position-Sensitive Detector Software – Phase 2*, 12–15 November 1986, LURE, Orsay, France..

Bricogne, G. (1987*a*). *Computational Aspects of Protein Crystal Data Analysis*, edited by J. R. Helliwell. pp. 107–119. Warrington: Daresbury Laboratory.

Bricogne, G. (1987*b*). *Computational Aspects of Protein Crystal Data Analysis*, edited by J. R. Helliwell. pp. 120–145. Warrington: Daresbury Laboratory.

Bricogne, G. (1988). *Acta Cryst.* A**44**, 517–545.

Bricogne, G. (1991*a*). *Proceedings of the CCP4 Study Weekend*, 25–26 January 1991, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 60–68. Warrington: Daresbury Laboratory.

Bricogne, G. (1991*b*). *Crystallographic Computing 5*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 257–297. Oxford: Clarendon Press.

Bricogne, G. (1993). *Acta Cryst.* D**49**, 37–60.

Brown, K., Nurizzo, D., Besson, S., Shepard, W., Moura, J., Moura, I., Tegoni, M. & Cambillau, C. (1999). *J. Mol. Biol.* In the press.

Carter, C. W., Crumley, K. V., Coleman, D. E., Hage, F. & Bricogne, G. (1990). *Acta Cryst.* A**46**, 57–58.

Charpak, G., Bouclier, R., Bressani, T., Favier, J. & Zupancic, C. (1968). *Nucl. Instrum. Methods*, **62**, 262–268.

Charpak, G., Hajduk, Z., Jeavons, A., Stubbs, R. & Kahn, R. (1974). *Nucl. Instrum. Methods*, **122**, 307–312.

Chiadmi, M., Kahn, R., de la Fortelle, E. & Fourme, R. (1993). *Acta Cryst.* D**49**, 522–529.

Colloc'h, N., El Hajji, M., Bachet, B., L'Hermite, G., Schiltz, M., Prangé, T., Castro, B. & Mornon, J. P. (1997). *Nat. Struct. Biol.* **4**, 947–952.

Cork, C., Hamlin, R., Vernon, W., Xuong, N. H. & Perez-Mendez, V. (1974). *J. Appl. Cryst.* **7**, 319–323.

Coste, F., Malinge, J. M., Serre, L., Shepard, W., Roth, M., Leng, M. & Zelwer, C. (1999). *Nucleic Acids Res.* **27**, 1837–1846.

Crumley, K. V. (1989). MSc thesis, University of North Carolina, USA.

Dauter, Z., Dauter, M., de la Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* In the press.

Djinovic-Carugo, K., Everitt, P. & Tucker, P. A. (1998). *J. Appl. Cryst.* **31**, 812–814.

Dumas, C. (1988). PhD thesis, Université Paris-Sud, France.

Fermi, G., Perutz, M., Shaanan, B. & Fourme, R. (1984). *J. Mol. Biol.* **175**, 159–174.

Fortelle, E. de la & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Fourme, R. (1978). *Synchrotron Radiation Applied to Biophysical and Biochemical Research*, edited by A. Castellani & I. F. Quercia, pp. 349–375. London: Plenum Press.

Fourme, R. (1993). *Collected Abstracts of Third European Workshop on Crystallography of Biological Macromolecules*, edited by K. Wilson, 24–28 May 1993, Como, Italy. Hamburg: EMBL Outstation.

Fourme, R. & Hendrickson, W. A (1990). *Synchrotron Radiation in Biophysics*, edited by S. Hasnain, p. 151. Chichester: Ellis Horwood.

Fourme, R. & Kahn, R. (1985). *Methods Enzymol.* **114**, 281–299.

Fourme, R., Shepard, W., Kahn, R., L'Hermite, G. & de La Sierra, I. L. (1995). *J. Synchrotron Rad.* **2**, 36–48.

Hao, Q., Dodd, F. E., Grossmann, J. G. & Hasnain, S. S. (1999). *Acta Cryst.* D**55**, 243–246.

Harvey, I., Hao, Q., Duke, E. M. H., Ingledew, W. J. & Hasnain, S. S. (1998). *Acta Cryst.* D**54**, 629–635.

Helliwell, J. R. (1992). *Macromolecular Crystallography with Synchrotron Radiation.* Cambridge University Press.

Helliwell, J. R., Ealick, S., Doing, P., Irving, T. & Szebenyi, M. (1993). *Acta Cryst.* D**49**, 120–128.

Helliwell, J. R. & Fourme, R. (1983). *The ESRF as a Facility for Protein Crystallography: Report and Design Study.* ESRP Report IRI-4/83, pp. 1–36. European Science Foundation, Strasbourg.

Hendrickson, W. A. (1985). *Trans. Am. Crystallogr. Assoc.* **21**, 11–21.

Hendrickson, W. A. & Lattman, E. E. (1970). *Acta Cryst.* B**26**, 136–143.

Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.

Hilge, M., Gloor, S. M., Rypniewski, W., Sauer, O., Heightman, T. D., Zimmermann, W., Winterthaler, K. & Piontek, K. (1998). *Structure*, **6**, 1433–1444.

Kahn, R. & Fourme, R. (1997). *Methods Enzymol.* **276**, 268–286.

Kahn, R., Fourme, R., Bosshard, R., Caudron, B., Santiard, J. C. & Charpak, G. (1982). *Nucl. Instrum. Methods*, **201**, 203–208.

Kahn, R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J. L., Brunie, S., Wery, J. P., Dideberg, O. & Janin, J. (1986). *Structural Biological Applications of X-ray Absorption, Scattering and Diffraction*, edited by H. D. Bartunik & B. Chance, pp. 297–304. New York: Academic Press.

Kahn, R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J. L., Dideberg, O. & Wery, J. P. (1985). *FEBS Lett.* **179**(1), 133–137.

Kahn, R., Fourme, R., Bosshard, R. & Saintagne, V. (1986). *Nucl. Instrum. Methods*, A**246**, 596–603.
Kahn, R., Fourme, R., Caudron, B., Bosshard, R., Benoit, R., Santiard, J. C., Charpak, G. & Sauli, F. (1980). *Nucl. Instrum. Methods*, **172**, 337–344.
Kahn, R., Fourme, R., Gadet, A., Dumas, C., Janin, J. & André, D. (1982). *J. Appl. Cryst.* **15**, 330–337.
Karle, J. (1980). *Int. J. Quantum Chem. Quantum Biol. Symp.* **7**, 357–367.
Krauss, N., Schubert, W.-D., Klukas, O., Fromme, P., Witt, H. T. & Saenger, W. (1996). *Nat. Struct. Biol.* **3**, 965–973.
Lemonnier, M., Fourme, R., Rousseaux, F. & Kahn, R. (1978). *Nucl. Instrum. Methods*, **152**, 173–177.
Li de la Sierra, I., Pernot, L., Prangé, T., Saludjian, P., Schiltz, M., Fourme, R. & Padrón, G. (1997). *J. Mol. Biol.* **269**, 129–141.
Malashkevich, V. N., Kammerer, R. A., Efimov, V. P., Schulthess, T. & Engel, J. (1996). *Science*, **274**, 761–765.
Messerschmidt, A. & Pflugrath, J. W. (1987). *J. Appl. Cryst.* **20**, 306–315.
Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol.* **91**, 201–228.
Montet, Y., Amara, P., Volbeda, A., Vernede, X., Hatchikian, C. E., Field, M. J., Frey, M. & Fontecilla-Camps, J. C. (1997). *Nat. Struct. Biol.* **4**, 523–526.
Mourey, L., Pédelacq, J.-D., Birck, C., Fabre, C., Rougé, P. & Samama, J.-P. (1998). *J. Biol. Chem.* **273**, 12914–12922.
Narayan, R. & Ramaseshan, S. (1981). *Acta Cryst.* A**37**, 636–641.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 472–494.
Phillips, J. C. & Hodgson, K. O. (1980). *Acta Cryst.* A**36**, 856–864.
Phillips, J. C., Templeton, D. H., Templeton, L. K. & Hodgson, K. O. (1978). *Science*, **201**, 257–259.
Phillips, J. C., Wlodawer, A., Goodfellow, J. M., Watenpaugh, K. D., Sieker, L. C., Jensen, L. H. & Hodgson, K. O. (1977). *Acta Cryst.* A**33**, 445–455.
Prangé, T., Schiltz, M., Pernot, L., Colloc'h, N., Longhi, S., Bourguet, W. & Fourme, R. (1998). *Proteins Struct. Funct.* **30**, 71–73.
Quillin, M. L., Baase, W. A. & Matthews, B. W. (1996). *Acta Cryst.* A**52**, C-215.
Ramakrishnan, V. & Biou, V. (1997). *Methods Enzymol.* **276**, 472–494.
Ramin, M. (1999). PhD thesis, Université Paris-Sud, France.
Ramin, M., Shepard, W., Fourme, R. & Kahn, R. (1999). *Acta Cryst.* D**55**, 157–167.
Sauer, O., Schmidt, A. & Kratky, C. (1997). *J. Appl. Cryst.* **30**, 476–486.
Schiltz, M., Fourme, R., Broutin, I. & Prangé, T. (1995). *Structure*, **3**, 309–316.
Schiltz, M., Kvick, Å., Svensson, O., Shepard, W., de la Fortelle, E., Prangé, T., Kahn, R., Bricogne, G. & Fourme, R. (1997). *J. Synchrotron Rad.* **4**, 287–297.
Schiltz, M., Prangé, T. & Fourme, R. (1994). *J. Appl. Cryst.* **27**, 950–960.
Schiltz, M., Shepard, W., Fourme, R., Prangé, T., de La Fortelle, E. & Bricogne, G. (1997). *Acta Cryst.* D**53**, 78–92.
Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K. V. & Rees, D. C. (1996). *Science*, **272**, 1615–1621.
Schoenborn, B. P. & Featherstone, R. M. (1967). *Adv. Pharmacol.* **5**, 1–17.
Schoenborn, B. P., Watson, H. C. & Kendrew, J. C. (1965). *Nature (London)*, **207**, 28–30.
Sheldrick, G. M. (1998). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, pp. 401–411. Dordrecht: Kluwer.
Shepard, W., Cruse, W., Fourme, R., de la Fortelle, E. & Prangé, T. (1998). *Structure*, **6**, 849–861.
Shepard, W., Ramin, M., Kahn, R. & Fourme, R. (1997). *Recent Advances in Phasing*, edited by K. S. Wilson, G. Davies, A. W. Ashton & S. Bailey, pp. 103–118. Warrington: Daresbury Laboratory.
Soltis, M. S., Stowell, M. H. B., Wiener, M. C., Phillips, G. N. Jr & Rees, D. C. (1997). *J. Appl. Cryst.* **30**, 190–194.
Stowell, M. H. B., Soltis, S. M., Kisker, C., Peters, J. W., Schindelin, H., Rees, D. C., Cascio, D., Beamer, L., Hart, P. J., Wiener, M. C. & Whitby, F. G. (1996). *J. Appl. Cryst.* **29**, 608–613.
Stuhrmann, H. B. & Kirste, R. (1965). *Z. Phys. Chem.* **46**, 247–250.
Tilton, R. F. Jr, Kuntz, I. D. Jr & Petsko, G. (1984). *Biochemistry*, **23**, 2849–2857.
Usha, R., Johnson, J. E., Moras, D., Thierry, J. C., Fourme, R. & Kahn, R. (1984). *J. Appl. Cryst.* **17**, 147–153.
Vitali, J., Robbins, A. H., Almo, S. C. & Tilton, R. F. (1991). *J. Appl. Cryst.* **24**, 931–935.
Weis, W. I., Kahn, R., Fourme, R., Drickamer, K. & Hendrickson, W. A. (1991). *Science*, **254**, 1608–1615.
Welch, M., Chinardet, N., Mourey, L., Birck, C. & Samama, J.-P. (1998). *Nat. Struct. Biol.* **5**, 25–29.
Weston, S. A., Camble, R., Colls, J., Rosenbrock, G., Taylor, I., Egerton, M., Tucker, A. D., Tunnicliffe, A., Mistry, A., Mancia, P., de la Fortelle, E., Irwin, J., Bricogne, G. & Pauptit, R. A. (1998). *Nat. Struct. Biol.* **5**, 213–221.
Wilson, K. S., Stura, E. A., Wild, D. L., Todd, R. J., Stuart, D. I., Babu, Y. S., Jenkins, J. A. Standing, T. S., Johnson, L. N., Kahn, R., Fourme, R., Gadet, A., Bartels, K. S. & Bartunik, H. D. (1983). *J. Appl. Cryst.* **16**, 28–41.
Xuong, N. H., Freer, S. T., Hamlin, R., Nielsen, C. & Vernon, W. (1978). *Acta Cryst.* A**34**, 289–296.