Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Determination of the structure of seleno-methionine-labelled hydroxymethylbilane synthase in its active form by multi-wavelength anomalous dispersion

A. Hädener,[a]* P. K. Matzinger,[b] A. R. Battersby,[c] S. McSweeney,[d]† A. W. Thompson,[e] A. P. Hammersley,[e] S. J. Harrop,[f]‡ A. Cassetta,[f]§ A. Deacon,[f]¶ W. N. Hunter,[f]†† Y. P. Nieh,[f]‡‡ J. Raftery,[f] N. Hunter[f] and J. R. Helliwell[f]

[a]Department of Chemistry, University of Basle, Switzerland, [b]F. Hoffmann-La Roche Ltd, Basle, Switzerland, [c]University Chemical Laboratory, University of Cambridge, England, [d]Daresbury Laboratory, Warrington, England, [e]ESRF/EMBL, Grenoble CEDEX, France, and [f]Department of Chemistry, University of Manchester, England

† Present address: ESRF/EMBL, Grenoble CEDEX, France.
‡ Present address: School of Physics, University of New South Wales, Australia.
§ Present address: CNR-Istituto di Strutturistica Chimica 'G. Giacomello,' Basovizza, Trieste, Italy.
¶ Present address: Department of Chemistry and Chemical Biology, Cornell University, Ithaca, USA.
†† Present address: Wellcome Trust Building, Department of Biochemistry, University of Dundee, Scotland.
‡‡ Present address: Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.

Correspondence e-mail: haedener@ubaclu.unibas.ch

© 1999 International Union of Crystallography
Printed in Denmark – all rights reserved

Received 13 May 1998
Accepted 9 November 1998

PDB Reference: seleno-methionine-labelled hydroxy-methylbilane synthase, 1ah5.

The enzyme hydroxymethylbilane synthase (HMBS, E.C. 4.3.1.8) catalyzes the conversion of porphobilinogen into hydroxymethylbilane, a key intermediate for the biosynthesis of heme, chlorophylls, vitamin B_{12} and related macrocycles. The enzyme is found in all organisms, except viruses. The crystal structure of the selenomethionine-labelled enzyme ([SeMet]HMBS) from *Escherichia coli* has been solved by the multi-wavelength anomalous dispersion (MAD) experimental method using the Daresbury SRS station 9.5. In addition, [SeMet]HMBS has been studied by MAD at the Grenoble ESRF MAD beamline BM14 (BL19) and this work is described especially with respect to the use of the ESRF CCD detector. The structure at ambient temperature has been refined, the *R* factor being 16.8% at 2.4 Å resolution. The dipyrromethane cofactor of the enzyme is preserved in its reduced form in the crystal and its geometrical shape is in full agreement with the crystal structures of authentic dipyrromethanes. Proximal to the reactive C atom of the reduced cofactor, spherical density is seen consistent with there being a water molecule ideally placed to take part in the final step of the enzyme reaction cycle. Intriguingly, the loop with residues 47–58 is not ordered in the structure of this form of the enzyme, which carries no substrate. Direct experimental study of the active enzyme is now feasible using time-resolved Laue diffraction and freeze-trapping, building on the structural work described here as the foundation.

1. Introduction

The biosyntheses of the tetrapyrrolic ring systems of heme, chlorophylls, vitamin B_{12} and related pigments proceed *via* a common tetrapyrrolic intermediate, uroporphyrinogen III (Battersby *et al.*, 1980; Eschenmoser, 1988; Battersby & Leeper, 1990). The latter compound is built from eight molecules of 5-aminolevulinic acid in a convergent manner in which three enzymes are involved (Fig. 1). In a reaction catalyzed by porphobilinogen synthase, pairs of 5-aminolevulinic acid molecules are first combined to form porphobilinogen (PBG) with release of two equivalents of water (Mitchell & Jaffe, 1993). Then, an unstable oligomeric compound, hydroxymethylbilane (HMB), is built from four molecules of PBG with release of four equivalents of ammonia and the consumption of one water molecule. This reaction is catalyzed by hydroxymethylbilane synthase (HMBS; E.C. 4.3.1.8; also known as porphobilinogen deaminase). A remarkable cyclization involving a rearrangement with loss of water, catalyzed by uroporphyrinogen III synthase, then leads to uroporphyrinogen III (Fig. 1). Considerable worldwide effort is put into the investigation of these enzymes (Chadwick

& Ackrill, 1994) not only because they are of vital importance for all organisms, except viruses, but also because defects in the biosynthesis of uroporphyrinogen III are associated with specific manifestations of the inherited disease porphyria in humans (Kappas *et al.*, 1995).

HMBS has been purified from many sources, including human erythrocytes, rat liver, spinach, *Euglena gracilis* and *E. coli* (Jordan, 1994). Both chemical and biochemical experimental evidence shows that, in order to construct HMB from PBG, the HMBS holoenzyme uses an unusual dipyrromethane cofactor (Hart *et al.*, 1987; Jordan & Warren, 1987) which is itself derived from two PBG building blocks and is covalently attached to the apoprotein *via* the S atom of a cysteine residue (Fig. 2*a*) (Jordan *et al.*, 1988; Miller *et al.*, 1988; Scott *et al.*, 1988; Hart *et al.*, 1990). The cofactor serves as an anchor to which the growing oligopeptide attached until a hexapyrrol chain has been built. It was established by ^{13}C labelling that the free α-position of the cofactor's PBG unit C2 (Fig. 2*a*) is the site to which the first incoming PBG unit is covalently bound. Cleavage of the hexapyrrolic chain at that position releases the product and leaves the cofactor in place, ready for another reaction cycle (Hart *et al.*, 1990). Hence, the enzyme is a polymerase with an inbuilt mechanism that somehow restricts the number of polymerization steps to four.

The enzyme from *E. coli* (Hart *et al.*, 1986; Jordan *et al.*, 1988; Hädener *et al.*, 1990) is a 34 268 Da monomer of 313 amino-acid residues, six of which are methionines. The dipyrromethane cofactor is attached to Cys242. There are three further cysteine residues but none of them is involved with disulfide bridges. The pH optimum for HMBS from *E. coli* is between 7.4 and 8.0, and the isoelectric point is 4.5. Within the optimal pH interval and for the overall reaction, the Michaelis constant is between 5 and 20 μM, and the turnover constant is of the order of 0.1 s^{-1} with respect to the formation of HMB.

With the aim of determining the three-dimensional structure of the enzyme, crystals of a variant of HMBS having seleno-L-methionine (SeMet) in place of methionine ([SeMet]HMBS) were produced for multi-wavelength anomalous dispersion (MAD) (Hädener *et al.*, 1993). Another effort was directed towards the elucidation of the three-dimensional structure of HMBS by crystallography and used multiple isomorphous replacement (MIR) (Louie *et al.*, 1992).

For the MAD approach, Hädener *et al.* (1993) optimized protocols for the expression, purification and crystallization of catalytically active [SeMet]HMBS. High levels of selenium occupancies were obtained. Strongly reducing conditions were used to prevent both the cofactor and the side chains of the SeMet residues from being oxidized. Freshly grown crystals obtained in this way are colourless.

The MIR structure determination (Jordan *et al.*, 1992) used data collected from yellow crystals that had been grown at pH 5.0–5.5 in the presence of polyethylene glycol 6000. Louie *et al.* (1992) found that the polypeptide chain of the enzyme is folded into three α/β domains of roughly equal size (Protein Data Bank entry 1PDA). Two of the domains were found to have an overall topology similar to that of the group-II periplasmic binding proteins (Spurlino *et al.*, 1991) and of a duplicated lobe of the bilobal transferrins (Baker *et al.*, 1987; Sarra *et al.*, 1990). Electron density for the cofactor is located in a cleft at the interface between these domains. It is connected to the S atom of the side chain of Cys242 in agreement with the earlier biochemical evidence mentioned above. Both the shape of the electron density for the cofactor and the yellow colour of the crystals indicated that the cofactor was present in an oxidized, inactive state. The oxidized state can most probably be interpreted by dipyrromethene or dipyrromethenone structures or mixtures thereof (Figs. 2*b* and 2*c*). Proximal to the active-site region, electron density for a stretch of polypeptide (residues 49–57) was found to be missing. Louie *et al.* (1992) proposed that this loop may constitute a flexible 'lid' with a role in transiently blocking the access of solvent to the active site during catalysis. Short stretches of polypeptide have also been found to be mobile at both the N- and the C-terminus (missing residues 1–2 and 308–313, respectively).

For [SeMet]HMBS Hädener *et al.* (1993), in a preliminary X-ray analysis with a Cu $K\alpha$ source in Basle, located the selenium sites in a Fourier difference map that was calculated using the difference $F_{\rm SeMet}$ (colourless crystal, reduced [SeMet]HMBS) minus $F_{\rm wild\text{-}type}$

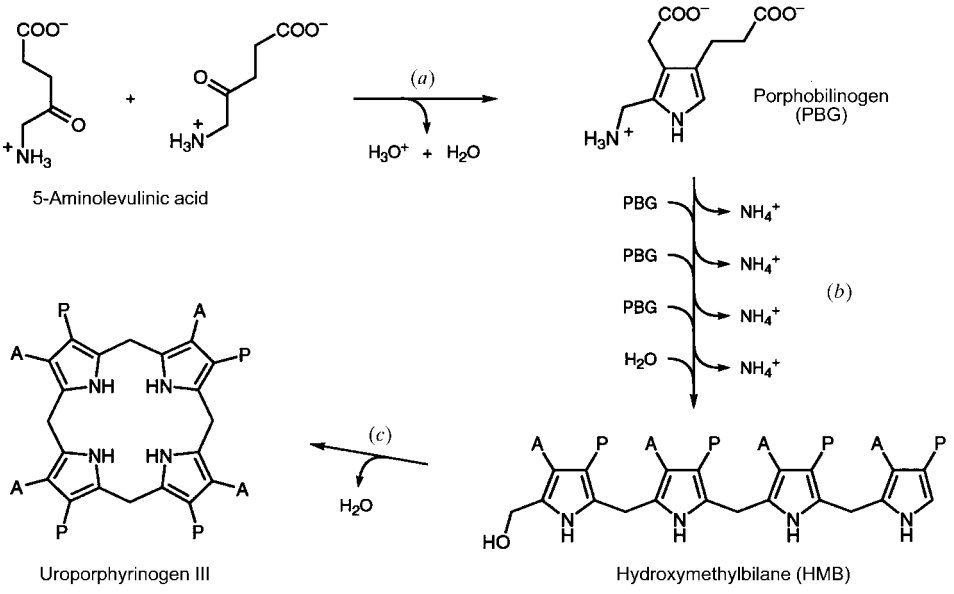

Figure 1
The biosynthesis of uroporphyrinogen III from 5-aminolevulinic acid. The enzymes involved are (*a*) 5-aminolaevulinic acid dehydratase; (*b*) hydroxymethylbilane synthase (HMBS); and (*c*) uroporphyrinogen III synthase. A = CH_2COO^-; P = $CH_2CH_2COO^-$.

(yellow crystal, oxidized HMBS) and where phases were calculated from the crystal structure model of oxidized HMBS, determined by MIR (Louie *et al.*, 1992). One of the six Se sites was found to be of low peak height indicating increased mobility at that position. Five of the six sites were, however, well ordered and well suited therefore for subsequent MAD data collection. In addition to difference peaks at the selenium sites, the difference map revealed significant difference density near the active site, where there are no Se atoms present, indicating that the colourless crystals of [SeMet]HMBS may indeed represent the catalytically active form of the enzyme carrying the cofactor in its original, reduced state and thus be structurally distinct from the oxidized, yellow, crystals (Hädener *et al.*, 1993). Subsequently the reduced, active, form of HMBS was refined using the model of the oxidized form of HMBS as the starting point (Louie *et al.*, 1996). This revealed in particular that the position of the C2 pyrrole unit of the cofactor (Fig. 2*a*) is distinct for the two oxidation states of the cofactor. The position of the C2 pyrrole unit of the oxidized cofactor has been proposed then to represent a binding pocket for substrate [termed the S site (Louie *et al.*, 1996)] in each of the chain elongation steps involved in the catalytic cycle (Louie *et al.*, 1992).

A MAD structure elucidation of the active, reduced, form of the enzyme was nevertheless feasible, and would be independent of the oxidized structure (Louie *et al.*, 1992, 1996). Moreover, the lack of visibility of the 49–57 loop in the two earlier structures, most likely due to its being associated with the catalytic mechanism in some currently unknown way, certainly was worthy of further investigation with reduced enzyme and freshly prepared crystals. The possible movement of this loop, very close to which is a conserved Asp residue (Asp46), may have a critical bearing on the 'proton shuttle' mechanism proposed by Louie *et al.* (1992, 1996) involving solely Asp84.

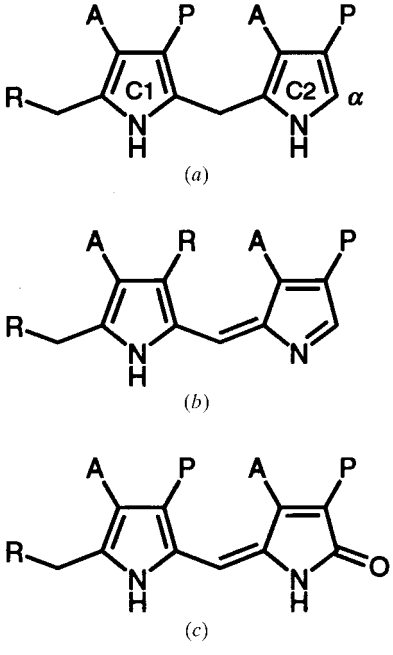

Figure 2
Oxidation states of the cofactor of the HMBS holoenzyme (R represents the apoenzyme moiety). (*a*) Dipyrromethane structure (catalytically active); for labelling conventions, see text; (*b*) dipyrromethene structure (inactive); (*c*) dipyrromethenone structure (inactive). A = CH_2COO^-; P = $CH_2CH_2COO^-$.

Two sets of MAD data have been collected, one at the Synchrotron Radiation Source (SRS) in Daresbury, the other at the European Synchrotron Radiation Facility (ESRF) in Grenoble. The purpose of the SRS experiment was to solve the structure of the active crystal form *de novo*. At SRS, [SeMet]HMBS was the first SeMet-labelled protein that was subjected to a MAD experiment in the course of the commissioning of station 9.5 (Brammer *et al.*, 1988). At ESRF this was the first project to test the ERSF image-intensifier detector during a commissioning run (Biou *et al.*, 1997; Helliwell *et al.*, 1996). Each of these MAD technical studies, as well as the SRS HMBS MAD structure, are reported in this paper.

2. Data collection and processing

2.1. Expression, purification and crystallization of [SeMet]HMBS

[SeMet]HMBS was expressed by growing the methionine-requiring mutant *E. coli* PO1562, carrying the plasmid pPA410, in a medium containing 50 mg l^{-1} SeMet as the sole source of methionine, and purified as previously described (Hädener *et al.*, 1993). Dithiothreitol was included with the buffers at up to 10 m*M* during purification and at 30 m*M* for storage so as to obstruct the oxidation of the cofactor. The enzyme was crystallized at pH 5.3 by equilibrating sitting drops of 50 µl with 6–7 mg ml^{-1} of protein, 0.3 m*M* EDTA, 15 m*M* dithiothreitol, 10%(*w*/*v*) polyethylene glycol 6000 and 0.01% NaN_3 in 0.1 *M* sodium acetate against a reservoir of 10–20 mg solid dithiothreitol. For the SRS experiments the crystals were used within 2 weeks, and at ESRF within 8 weeks, of setting up the crystallizations with the enzyme in its reduced state. In both cases the crystallization experiments were set up near the beamline to avoid the need for excessive transportation.

2.2. MAD data collection at SRS station 9.5 and diffraction data reduction

Station 9.5 at the Daresbury SRS consists of a rapidly tunable, water-cooled, double crystal monochromator [Si(111)] and a toroidal mirror set to reflect and produce a point focus of X-rays for X-ray wavelengths greater than 0.5 Å (Brammer *et al.*, 1988; Thompson *et al.*, 1992; Deacon *et al.*, 1995). An extended description of the station data acquisition software is given in Kinder *et al.* (1996). The SRS Wiggler I has a λ_c of 0.89 Å set by the 5T magnet field and the SRS machine energy of 2 GeV.

A single crystal of [SeMet]HMBS was used to measure the X-ray fluorescence spectrum at the selenium *K* absorption edge by tuning the wavelength of the incident X-ray beam. The positioning of the fluorescence detector, a small proportional chamber, was found to be critical to record a good

Table 1
MAD data collection at ambient temperature and merging statistics at the SRS station 9.5 IP.

λ_2, λ_3, λ_4 nomenclature follows that of Peterson *et al.* (1996) *i.e.* on edge, white line and short wavelength reference. λ_1, a pre-edge wavelength data set, was not measured. Resolution limit 2.6 Å. *R* values and completeness values in parentheses are for the outer resolution annulus 2.60–2.67 Å.

	λ_2	λ_3	λ_4
Wavelength (Å)	0.9799	0.9795	0.9505
R_{merge} (%)	4.9 (15.7)	4.9 (15.5)	5.2 (17.8)
R_{anom} (%)†	5.0	5.0	5.0
Total No. of reflections	28798	28932	27774
No. of independent reflections	9110	9101	8808
Completeness (%)	83.6 (82.7)	83.6 (83.2)	80.8 (81.0)
Multiplicity	3.2	3.2	3.2
Percentage of reflections with $\Delta_{ANOM} > 3\sigma(\Delta_{ANOM})$ on *I*†	11.9	12.8	8.8

† The R_{anom} values at 6 Å resolution for λ_2, λ_3, λ_4 are respectively 5.5, 5.8 and 4.5% with R_{merge} values here of 2.4, 3.2 and 3.5%. Correspondingly the percentage of reflections with $\Delta_{ANOM} > 3\sigma(\Delta_{ANOM})$ at 6 Å, on intensity, are 29.7, 37.1 and 21.3% for λ_2, λ_3 and λ_4.

quality spectrum largely free of scattered radiation. The entire sensitive area of the proportional counter was placed, relative to the direction of the incident X-ray beam, in (slightly behind in fact) a plane going through the crystal and perpendicular to the beam. Moreover a thin piece of aluminium (400 µm) was used to attenuate the beam so as to prevent detector saturation. The measured spectrum (Fig. 3) showed a sharp peak and white-line enhancement. The spectral bandpass (discussed in Thompson *et al.*, 1992) was ~5 eV. The three wavelengths at which the X-ray crystallographic data were measured were the inflection point of the edge (0.9799 Å; λ_2), the peak of the white-line absorption (0.9795 Å; λ_3) and a short wavelength reference (0.9505 Å; λ_4). A pre-edge wavelength crystallographic data set (λ_1) was not measured [see Peterson *et al.* (1996) for the nomenclature of wavelength labelling and for an evaluation of data collection and phasing strategies at 4, 3 and 2 such wavelengths]. One wavelength (λ_3) was therefore chosen to maximize the anomalous differences (proportional to $2f''_{max}$), the others (*i.e.* λ_2 and λ_4) to maximize the isomorphous differences (*i.e.* $\Delta f'$), although two wavelengths are in principle sufficient (based on F^+ and F^- at one and F^+ or F^- at a second wavelength) for phasing (Helliwell, 1984).

The crystallographic data were recorded from aligned crystals at ambient temperature in the dark on an image plate, MAR Research, detector (180 mm diameter) in 3° rotations (ten oscillations) per image with total exposure times ranging from 90 to 120 s per image yielding good diffraction data to 2.4 Å resolution. The SRS beam current ranged from 170 to 111 mA. To minimize any data errors due to radiation damage, the Friedel anomalous differences and Δf were measured close together in time. This was accomplished by switching the wavelength after wedges of 6° (*i.e.* two images were recorded per wavelength). A set of six 3° images, two for each wavelength, was thus collected in sequence in an iterative manner. The time taken to change the wavelength was *ca* 8 s. The time taken to read out each diffraction image was *ca* 90 s (*i.e.* the data acquisition efficiency was ~50%). Overall the time for the MAD data collection was 5 × 8 h shifts (*i.e.* 40 continuous h). This involved use of two separate [SeMet]HMBS crystals. One was set to yield Friedel anomalous differences on the same image, thus avoiding image-to-image errors. The other crystal orientation was across the (110) direction rather than (100) or (010) due to the fact that the crystal habit is like a parallelepiped thin plate (size ~0.5 × 0.5 × 0.075 mm). The total angular coverage of the data at three wavelengths was 78°, collected in one shift, from the set crystal. In a separate run more data were measured on another crystal at a single wavelength (0.9 Å) involving 58 × 2° images; this 'third-crystal' data set was the one used in the final refinement of the model, which was built into the MAD electron-density map based on the 3λ data from the set crystal.

All the diffraction data were integrated using *MOSFLM* and merged with *ROTAVATA*/*AGROVATA* (Collaborative Computational Project, Number 4, 1994). The crystals, in space group $P2_12_12$, at ambient temperature had cell parameters a = 88.47, b = 76.15, c = 50.79 Å with one molecule of 34 550 Da, corresponding to the mass of [SeMet]HMBS fully substituted with six SeMet residues, in the asymmetric unit. The data statistics are given in Table 1.

2.3. MAD data collection at ESRF BM14 (BL19) and diffraction data reduction

ESRF BM14 (BL19) in Grenoble consists of a collimating mirror, double-crystal monochromator and toroid focusing mirror whereby the useful wavelength range of this bending magnet station is 0.6–2 Å (Thompson, 1997; Biou *et al.*, 1997). The addition of the collimating mirror allows fine $\delta\lambda/\lambda$ ($\simeq 2 \times 10^{-4}$) to be preserved without any slitting down

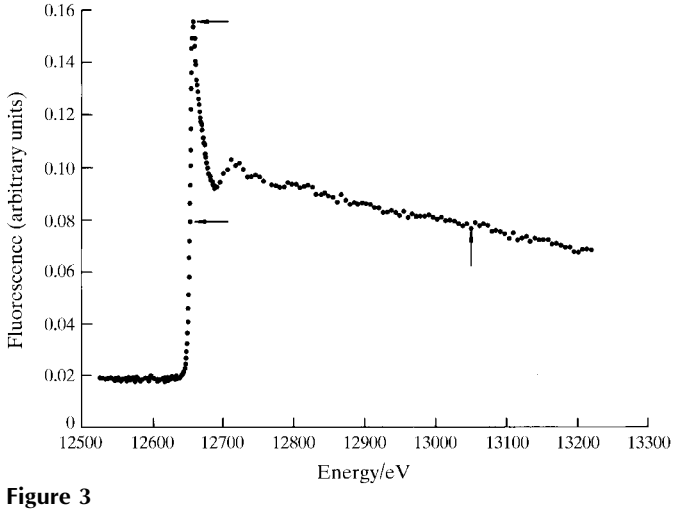

Figure 3
The selenium fluorescence spectrum measured directly off a single crystal at the Daresbury SRS station 9.5. The wavelengths used for subsequent data collection are those that correspond to the points marked with an arrow.

Table 2
Cryo temperature MAD data collection and merging statistics at ESRF BM14 CCD.

The values for λ_{2a} and λ_{2b} refer to the two options for the position of the inflection point arising because of the white line (see text). Resolution limit 2.70 Å. *R* values and completeness values in parentheses are in the outer resolution annulus 2.84–2.70 Å. The ESRF R_{merge} values are larger than the SRS IP R_{merge} values. The ESRF data have a higher redundancy (~9:1 *versus* 3:1) and it is known that this inflates R_{merge} values even though the accuracy of the data is improved, *via* data averaging (Weiss & Hilgenfeld, 1997; Diederichs & Karplus, 1997; Harrop *et al.*, 1996).

	λ_{2a}	λ_{2b}	λ_3	λ_4
Wavelength (Å)	0.9800	0.9801	0.9798	0.90
R_{merge} (%)	8.4 (24.4)	9.1 (29.7)	9.2 (28.0)	7.4 (18.7)
R_{anom} (%)†	3.7	3.7	6.1	5.0
Total No. of reflections	78099	77313	76339	33841
No. of independent reflections	8945	9031	8822	9071
Completeness (%)	95.2 (97.9)	95.7 (98.4)	95.7 (97.8)	97.6 (99.9)
Multiplicity	8.7 (8.4)	8.6 (8.2)	8.6 (8.3)	3.7 (3.8)
Percentage of reflections with $\Delta_{ANOM} > 3\sigma$ (Δ_{ANOM}) on *I* (%)†	3.7	2.9	12.7	2.8

† The R_{anom} values at 6 Å resolution for λ_{2a}, λ_{2b}, λ_3 and λ_4 are, respectively, 3.3, 2.9, 7.0 and 4.8% with R_{merge} values here of 5.2, 5.2, 6.0 and 4.4%. The percentage of reflections with $\Delta_{ANOM} > 3\sigma$ (Δ_{ANOM}) at 6 Å, on intensity are 8.5, 4.9, 42.5 and 13.0%.

vertically. Such a mirror could not be accommodated on station 9.5 on the SRS wiggler beamline owing to the congestion of six other stations and beamlines being accommodated there, unlike ESRF BM14 which is a single-purpose beamline.

A crystal sample (~0.3 × 0.5 × 0.07 mm) was chosen for flash freezing. The cryoprotectant was prepared by mixing a solution of polyethylene glycol 6000 (25%) in 0.1 *M* sodium acetate buffer (pH 5.3, 75%) with glycerol in a ratio of 90 to 10%. The X-ray fluorescence spectrum was measured in a similar manner to that described in the previous section. The spectra measured at ESRF BM14 and SRS 9.5 were of similar sharpness, whereby the collimating mirror at ESRF (upstream of the monochromator) allowed the full divergence of the ESRF beam to be collected whilst at SRS vertical slits were employed (as referred to above).

Crystallographic data were recorded on the 'ESRF II CCD' (Moy, 1994) previously calibrated for non-uniformity and spatial distortion as described by Moy *et al.* (1996). Four data sets were collected at different wavelengths around the edge (Helliwell *et al.*, 1996). The protocol described in the previous section was modified in two ways as follows. Firstly, the inflection point data (λ_2) were actually recorded twice at very slightly different wavelengths, namely half-way up between the pre-edge floor and the post-edge plateau (λ_{2a}) and also half-way between the pre-edge and the top of the white line spike (λ_{2b}). λ_{2b} is equivalent to λ_2 in the previous section. Secondly, full data sets at each wavelength were recorded, rather than by 6° wedges, because, with the sample being frozen, differences due to time-dependent radiation damage were eliminated and also because the long-term positional stability of the X-ray beam of ESRF was better than that of SRS at that time. The actual wavelengths targeted were λ_{2a} = 0.9800, λ_{2b} = 0.9801, λ_3 = 0.9798 and λ_4 = 0.90 Å. However, for λ_{2a} the set up wavelength in the instrument configuration file was logged as 0.9800 Å, in fact 0.97996 Å, but at the end of data collection it was 0.980145 Å *i.e.* λ_{2b}'s value. Hence, the planned evaluation of the effect of this small difference was not possible and λ_{2a} and λ_{2b} have to be considered identical in this report although, since they represent slightly different parts of the steeply rising part of the absorption edge, the two respective data sets were not merged.

Diffraction images were recorded in 1° intervals with an exposure of 90 s. Each image was read out in 3.5 s. Hence, the data acquisition efficiency was 96% with the CCD and this exposure time.

The data collection protocols followed at SRS 9.5 (§2.2) and ESRF BM14 (this section) were compared by taking into account various parameters that can affect the choice of exposure times such as beam currents, bandpasses, beam sizes, wavelengths, slit settings, the sizes of the crystal samples and the states of the detectors. The recorded exposure times for the BM14 data are, after correction for differences in sample and experimental conditions, a factor of only ~5 times higher than would be recorded on SRS 9.5. There remains a missing factor whereby ESRF BM14's beam intensity is in fact higher than this indicates. It is evident from subsequent information (Thompson, personal communication) that the detector gain, and probably alignment of the beam with respect to sample slits, were not optimal during this commissioning experiment.

The sequence of whole data set measurements was λ_{2b}, λ_3, λ_4 and finally λ_{2a}. The first three data sets were produced in 'inverse beam' *i.e.* φ and φ + 180° so that a total angular coverage of data collection from the one single crystal comprised (3 × 180 + 90) = 630°. This was completed in a total of 21 h of continuous beamtime with the 6 GeV ESRF circulating current commencing at 127 mA with a lifetime of 30 h in $^1/_3$rd fill mode.

The ESRF diffraction data were integrated and merged using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The refined unit-cell parameters of the frozen crystal were *a* = 86.01, *b* = 76.00, *c* = 49.14 Å (compared with those at room temperature of *a* = 88.47, *b* = 76.15, *c* = 50.79 Å). The cryo data reduction details are given in Table 2.

2.4. Ambient-temperature SRS MAD data: Scaling of data, location of Se atoms, *f*′ and *f*″ refinement, MAD phase calculation, electron-density map

The data sets were put on a common scale using *SCALEIT* (Collaborative Computational Project, Number 4, 1994) (using λ_2 as the reference data set); for λ_3 to λ_2 k_{emp} = 1.22 and for λ_4 to λ_2 k_{emp} = 1.45 where k_{theor} would be 0.75 and 1.95, respectively (Table 3). The location of the selenium sites was attempted *via* the anomalous difference Patterson map Harker sections calculated from λ_3. This yielded only two of the five well ordered sites leading to the later experiment at

ESRF (see below). Subsequently use of averaged anomalous differences over the three wavelengths measured at SRS (which are each of equivalent quality, see Table 1) yielded a high-quality Patterson map (Fig. 4). In any case, the selenium positions could be derived from calculated phases based on the oxidized HMBS structure (Louie *et al.*, 1992), *i.e.* as had been reported by Hädener *et al.* (1993); such anomalous difference Fourier maps using the Δ_{ANOM} data sets for each of the three SRS wavelengths were indeed calculated. These various Fourier maps were all in agreement with each other, and also are in agreement with the anomalous difference Patterson map shown in Fig. 4. The coordinates of the Se atoms in amino-acid residues 28, 41, 82, 234 and 286 were then refined with *MLPHARE*. The Se atom of residue Met1 was not sufficiently ordered. The *MLPHARE* overall figure of merit for 8875 reflections for the MAD phasing was 0.29 for data to 2.6 Å (up to 6 Å it was 0.53) and comprised 7788 acentric and 1087 centric reflections. These phases and the resultant MAD electron-density map were then subjected to the density-modification phase-improvement procedure *DM* (Cowtan & Main, 1993) assuming 40% solvent content. The phase improvement was applied in gradually increasing resolution shells commencing at 3.8 Å ending up at a final resolution limit for the phasing of 2.6 Å. Table 4 summarizes the substantial improvement in the figure of merit of the MAD three-wavelength phases after *DM*. These *DM*-improved three-wavelength phases were used to calculate an electron-density map (Fig. 5), using the f'-dip data as native (λ_2), from which a model was constructed using the *BONES* option of *O* (Jones *et al.*, 1991).

A two-wavelength calculation in *MLPHARE* was also performed using the f' dip (λ_2) as native and the f'' reference data set as the second wavelength (λ_4). This was also subjected to phase improvement with *DM*. Table 5 shows the comparison of these two-wavelength experimental phases and after *DM* phase improvement (a comparison of a slab of the electron-density maps for the 2-λ after *DM* phases *versus* the 3-λ MAD experimental phases was made by Chayen *et al.*, 1996).

Table 3
Estimates of $\Delta f'$ and f'' (*MLPHARE* for Daresbury 9.5 data).

$k_{theor} = \Delta f'_{(\lambda_2-\lambda_j)}/f''_j$. $k_{emp} = \langle\Delta_{ISO}\rangle/\langle|\Delta_{ANO}|\rangle/2$; *i.e. SCALEIT* output with λ_2 as 'native'.

	$\Delta f'$	f''	k_{theor}	k_{emp}
λ_2	—	4.69	—	—
λ_3	3.63	4.87	0.75	1.22
λ_4	7.05†	3.61‡	1.95	1.45

† Value derived from Sasaki (1989) used to convert refined *MLPHARE* occupancy into absolute $\Delta f'$ value for (λ_3–λ_2). ‡ Value from Sasaki (1989) used to convert refined *MLPHARE* occupancies, on an arbitrary scale, into absolute f'' values for λ_2 and λ_3.

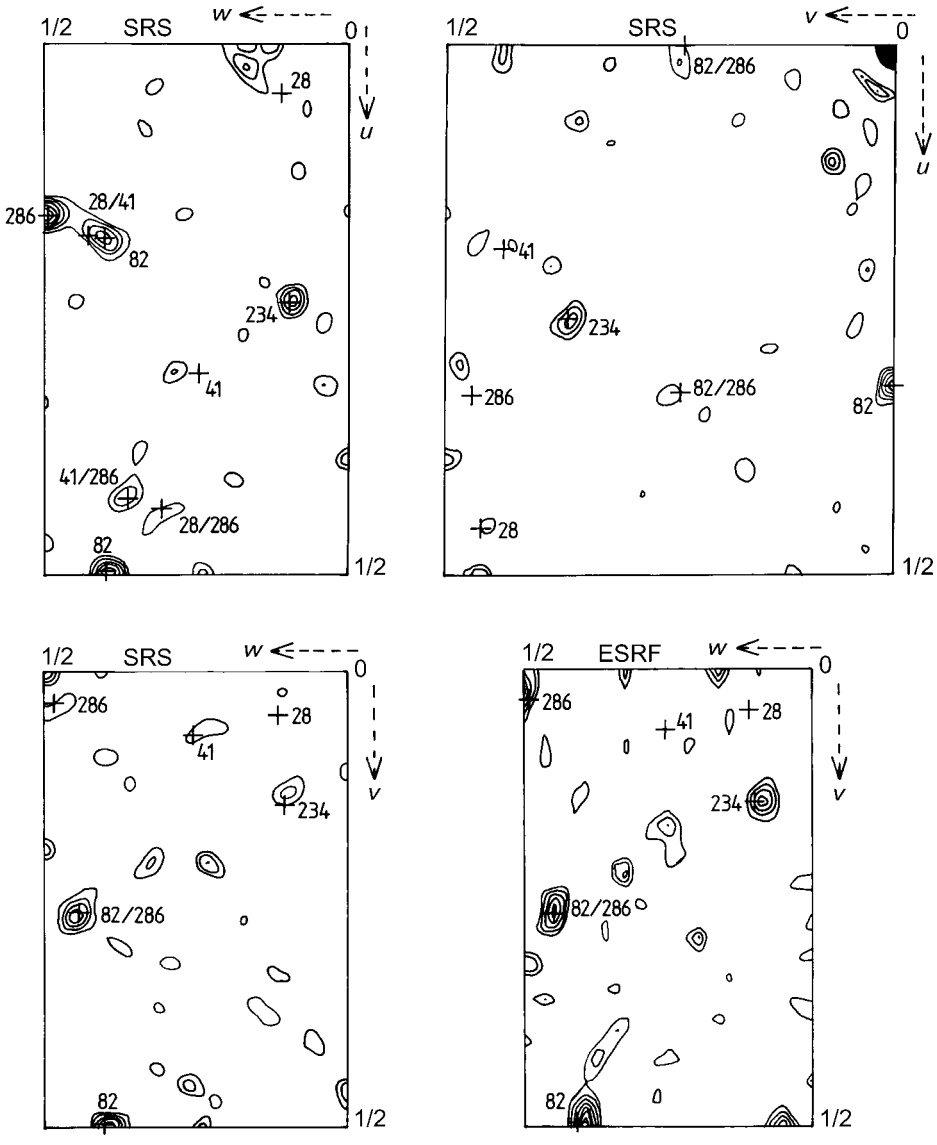

Figure 4
Anomalous difference Patterson map (Harker sections $u = \frac{1}{2}$, bottom left; $v = \frac{1}{2}$, top left; $w = 0$, top right) for the SRS data (with the anomalous differences averaged across the three wavelength data sets measured). The anomalous difference Patterson map ($u = \frac{1}{2}$ Harker section) for the ESRF data (again averaged across three wavelengths) is also shown (at bottom right). The labels are SeMet residue numbers. Some peaks which appear close to the Harker sections are cross-vectors between two Se atoms; these are labelled with the numbers of the residues involved.

2.5. Cryo-temperature ESRF MAD data: anomalous Patterson maps

With the more complete and highly redundant ESRF data an

Table 4
Three-wavelength phase calculation using *MLPHARE* and *DM* phase improvement.

$\Delta\varphi^\circ$ = average phase change of the MAD phases compared with those after DM. $\langle m\rangle$ = average figure of merit. The value for the average figure of merit for the MIR phase determination by Louie *et al.* (1992) of the oxidized, inactive form of the protein, based on heavy-atom derivatives, was 0.73 (∞ to 3 Å). From this the mean phase error can be estimated, from $\cos^{-1}\langle m\rangle$, as 43°. The MIR map had subsequently been subjected to solvent flattening/phase improvement

$(4\sin^2\theta)/\lambda^2$	$\langle m_{3\lambda\mathrm{MAD}}\rangle$	$\langle m_{\mathrm{DM}_{3\lambda}}\rangle$	$\Delta\varphi^\circ$
0.035	0.449	0.752	30.01
0.052	0.408	0.727	30.89
0.070	0.393	0.685	27.84
0.084	0.353	0.662	31.78
0.098	0.252	0.626	42.37
0.111	0.198	0.611	50.97
0.123	0.198	0.658	58.55
0.133	0.174	0.635	57.06
0.148	0.158	0.574	52.64

Table 5
Two-wavelength phase calculation using *MLPHARE* and *DM* phase improvement.

For definition of symbols see Table 4.

$(4\sin^2\theta)/\lambda^2$	$\langle m_{2\lambda\mathrm{MAD}}\rangle$	$\langle m_{\mathrm{DM}_{2\lambda}}\rangle$	$\Delta\varphi^\circ$
0.035	0.298	0.703	40.47
0.052	0.259	0.718	45.40
0.070	0.245	0.667	43.38
0.084	0.211	0.671	52.24
0.098	0.150	0.668	62.28
0.111	0.118	0.672	71.76
0.123	0.113	0.750	76.44
0.133	0.102	0.672	72.35
0.148	0.091	0.610	72.76

improved Patterson map (for λ_3) was achieved (Helliwell *et al.*, 1996). Both the better completeness and the fact that the data were recorded at low temperature may have significantly contributed to this result. This work showed that the CCD detector developed at ESRF, with its rapid readout and accurate data, is useful for protein crystallography, including MAD. Subsequent use of the detector for solving new protein structures by MAD has indeed confirmed the high quality of the data obtainable and its utility for MAD experiments (Biou *et al.*, 1997). Fig. 4 (at bottom right) shows, for comparison with the SRS IP Patterson calculation, one of the Harker sections (u = 1/2); this is also calculated from the averaged anomalous differences across λ_{2b}, λ_3, and λ_4 as was carried out for the SRS Harker sections shown in Fig. 4.

3. The ambient-temperature structure of [SeMet]HMBS

3.1. Molecular geometry parameters for the refinement of the reduced dipyrromethane cofactor

To obtain a valid geometry for the reduced dipyrromethane cofactor, the Cambridge Structural Database (CSD) (Allen *et al.*, 1979) was searched for matches containing either an unsubstituted dipyrromethane moiety or an appropriately furnished thioether substructure (Fig. 6*a*). Seven of the matches, *i.e.* BILRUB10 (Bonnett *et al.*, 1978), DAPPIT (Stark *et al.*, 1985), ECTEPM (Bonnett *et al.*, 1972), EXBILE (Sheldrick & Becker, 1979), JOSMUZ (Sessler *et al.*, 1992), LAZMOO (Engel *et al.*, 1993), MESBIL (Becker & Sheldrick, 1978), as well as JAMRUK (Reagen & Radonovich, 1989), containing an unsubstituted pyrrole molecule, were found useful to define the dipyrromethane geometry. Of the large number of matches for the thioether moiety, 14 were used to define the geometry of the thioether linkage. A dictionary of bond distances and bond angles for the reduced cofactor and its attachment to the Cys side chain was constructed from these entries (Fig. 6*b*, Table 6).

Table 6
Bond distances and bond angles for the reduced dipyrromethane cofactor and its linkage to the side chain of a Cys residue.

For the arrangements of atoms and their names, see Fig. 6(*b*). The following bond types were considered equivalent with pre-defined bond types: *f* and *h* with C—CH2E; *g* with CH2E—CH2E; *i* with C—CO; *o* with H—NH1. The following angle types were considered equivalent with pre-defined angle types: 9 with C—CH2E—CH2E; 10 with CH2E—C—OC; 15 with CR1E—NH1—H. The geometries of the dipyrromethane and of the thioether substructures were derived from eight and 14 crystal structure models, respectively (for details see text).

Bond type	Distance (Å)	s.u. (Å)	Angle type	Angle (°)	s.u. (°)
a	1.414	0.046	1	108.76	1.90
b	1.361	0.026	2	108.46	1.65
c	1.364	0.055	3	107.16	2.39
d	1.511	0.046	4	120.07	2.75
e	1.505	0.041	5	125.66	1.62
f	1.516	0.025	6	115.37	1.24
g	1.520	0.030	7	113.46	1.53
h	1.516	0.025	8	113.32	2.80
i	1.249	0.019	9	112.60	1.70
k	1.824	0.016	10	118.40	2.30
l	1.810	0.007	11	113.96	1.56
m	1.518	0.010	12	101.93	1.39
n	1.483	0.011	13	114.40	2.30
o	0.980	—	14	121.22	1.83
			15	125.62	—

Corresponding topology and parameter files for use within *X-PLOR* (Brünger, 1992) were then generated from the dictionary and the self consistency of the parameter file was confirmed by *X-PLOR* energy minimization. Appropriate improper constraints were included to restrain the carboxylate groups of the acetate and propionate side chains of the cofactor as well as the atoms of each of the pyrrole rings, including the atoms directly attached to the rings, to planarity. Appropriate dihedral constants were included in analogy to those found in the topology and parameter files for the parameterization of protein residues (Engh & Huber, 1991), in particular of Phe, Asp, Glu, and Met. Nonbonded constants were defined in analogy to those of the pre-defined atoms CH2E, C, OC, NH1, and some of the atoms of the indole moiety of Trp. The O atoms of the carboxylate groups were specified as hydrogen-bond acceptors and the pyrrole H—N groups were defined as hydrogen-bond donors. The PATCH statement of *X-PLOR* was used to attach the CHA atom of the cofactor (residue DPM) to the SG atom of the Cys side chain.

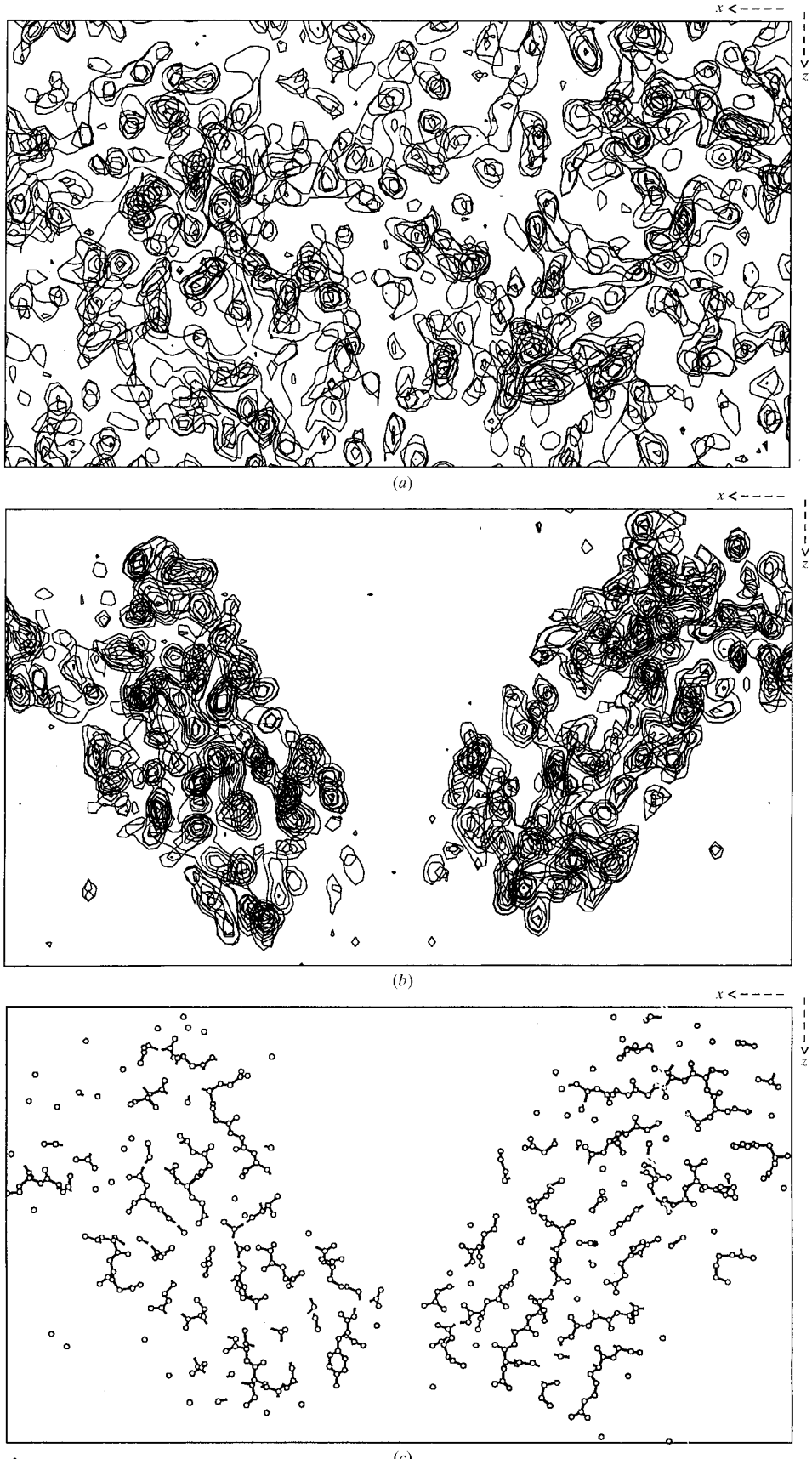

Figure 5
Representative MAD electron-density map slab for (*a*) three wavelengths, (*b*) three wavelengths plus *DM* map/phase-improvement. Both maps were calculated from the SRS data measured at ambient temperature. *(c)* The MAD model for this slab.

3.2. Molecular geometry parameters for the refinement of the SeMet residue

The topology and parameter files for the SeMet residue for use within *X-PLOR* were constructed in analogy to those for the Met residue. Appropriate bond distances and bond angles were adapted from a crystal structure of seleno-D,L-methionine (Rajeswaran & Parthasarathy, 1984). In particular, the selenoether substructure was defined by specifying 1.91 and 1.94 Å for the Se—CH_3 and Se—CH_2 bond distances, and by setting the angles CH_3—Se—CH_2 and Se—CH_2—CH_2 to 98.91 and 114.16°, respectively.

3.3. Map interpretation and refinement

The SRS 3-λ MAD+*DM* phased electron-density map (Fig. 5*b*) was interpreted on both a Silicon Graphics and an Evans & Sutherland molecular graphics workstation using the automatic *BONES* option of *O*. This was readily performed as polyalanine in five polypeptide segments corresponding to the following regions within the amino-acid sequence, namely, residues 6–41; 63–110; 113–219; 221–236 and 244–302. Initially the loop carrying the cofactor was not visible and density for the cofactor was not clear. Alternating cycles of simulated annealing with *X-PLOR* and model building within *O* was then conducted. At each refinement cycle both the *R* and the R_{free} (Brünger, 1993) factors (based on 10% of the data not included in the refinement) were calculated. The data set used for the refinement was that collected at 0.9 Å wavelength to 2.4 Å resolution (Table 7).

After the final cycle, the r.m.s. deviations of geometry from ideal values were 0.009 Å (bonds), 1.751° (angles), 23.46° (dihedrals), and 1.428° (impropers).

The outcome from this refinement was the structure of [SeMet]HMBS in its active reduced

form illustrated in Fig. 7. The final model comprises 2298 atoms, representing 299 residues, now including the C terminus up to residue 313 (*i.e.* beyond the 308 position of the MIR model), and the dipyrromethane cofactor of the [SeMet]HMBS holoenzyme, and 177 water molecules. The coordinates of the refined structure model have been deposited with the Protein Data Bank (Bernstein *et al.*, 1977).

3.4. The structure of the active site

The dipyrromethane structure model of the reduced cofactor, as built according to mean values for bond lengths and angles taken from independent representative dipyrromethanes in the CSD, readily fitted into the electron density that was available in the active site after refinement cycle 5 of the ambient-temperature structure (Table 7). This is noteworthy because the cofactor is conformationally restricted due to the small number of bonds that allow rotation and due to the restriction to planarity of the atoms of each of the pyrrole rings including the atoms directly attached to the rings. The most critical bond angle at the methylene bridge C4A—CHB—C1B (Fig. 6) between the pyrrole rings is required to be $115.37 \pm 1.24^\circ$ ($n = 9$) based on the collection of independent dipyrromethanes reported above. This bond angle in the refined final model is 114.0° in agreement with this value.

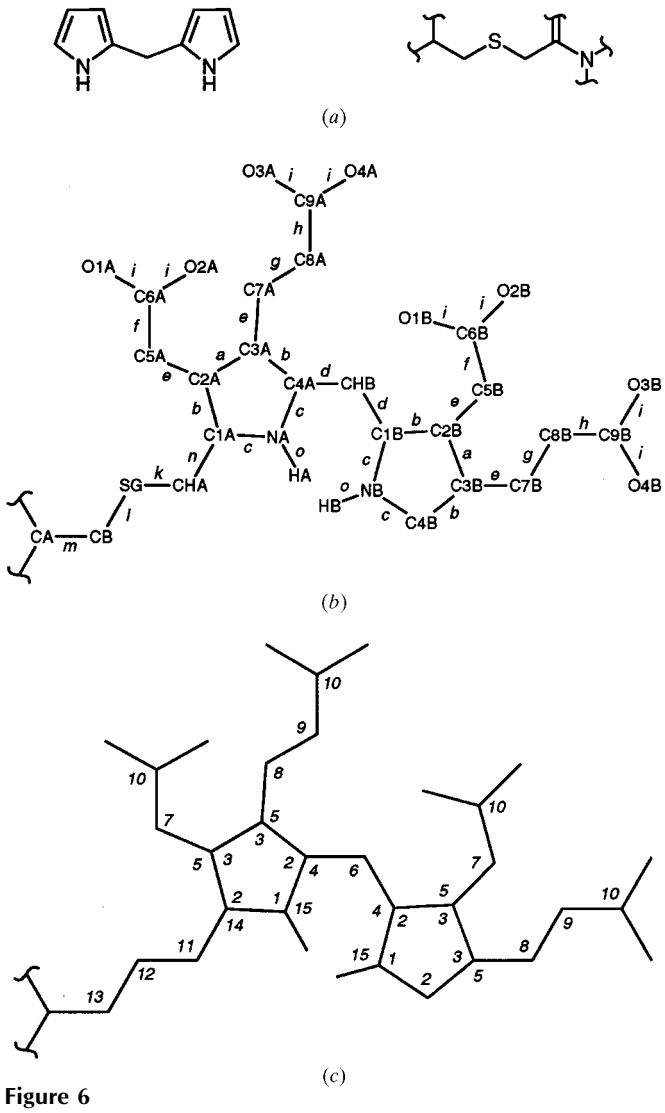

Figure 6
Geometry of the reduced dipyrromethane cofactor. For values of bond distances and bond angles, see Table 6. (*a*) Dipyrromethane core structure and thioether moiety used to search the CSD; (*b*) atom names and bond types, and (*c*) angle types used for the parameterization of the cofactor and the Cys residue to which it is attached.

This evidence shows that this MAD-derived structure, which is independent, *i.e.* derived from a completely new set of phases, electron-density map and subsequent model refinement, represents the enzyme in its catalytically active state with its cofactor in the reduced form. The MIR-derived model (Louie *et al.*, 1992) however, shows the cofactor in an oxidized, catalytically inactive state, probably as a dipyrromethenone (Fig. 2*c*) possibly in admixture with the corresponding dipyrromethene (Fig. 2*b*). This view is supported by the findings summarized below.

(i) The two pyrrole rings in the MIR-derived model are almost coplanar in agreement with the π-conjugation being extended across both rings of a dipyrromethenone or dipyrromethene.

(ii) The C4A—CHB—C1B bond angle of 124.9° in the MIR-derived model (PDB entry code 1PDA) agrees with CHB being in a planar sp^2-hybridized state.

(iii) Whereas the MIR-derived electron-density map shows a lobe of electron density emanating from C4B of the cofactor's second pyrrole ring (in keeping with a dipyrromethenone, Fig. 2*c*), the MAD-derived map does not show this feature, again supporting an active, dipyrromethane structure for the cofactor.

(iv) Pyrromethanes are colourless and so active HMBS should also be colourless; this was true for the crystals used for the MAD studies. Those for the work based on MIR were yellow.

Fig. 8 compares the two cofactor states and the reduced cofactor is highlighted by the $(F_o - F_c)$ difference electron density. It is interesting that adjacent to the atom C4B, which represents the catalytically important free α-position of the cofactor's second pyrrole ring C2, electron density is present which is spherical in shape (at 3.5σ, see Fig. 8). Since this density is closer to C4B than to any other atom of the holoenzyme or ordered water molecule, its molecular origin deserves further comment. Among possible candidates are an ordered water molecule or an ion. Electrospray mass spectrometry of [SeMet]HMBS did not reveal the presence of any strongly bound material (Hädener *et al.*, 1993) and HMBS is known not to depend on any essential ions for catalysis. Furthermore, the only spherical ion in the mother liquor that was used to crystallize [SeMet]HMBS is Na^+. The shape of the density and the fact that there is a hydrogen bond donor/acceptor in its vicinity (the OH group of Ser81 at 3.3 Å) support the view that the density represents an ordered water molecule. A number of other groups are located within a sphere of 4.0 Å around the putative water molecule, *i.e.* the water molecules 386 (3.5 Å), 385 (3.6 Å), 381 (4.0 Å), the N atom NB of the cofactor (3.8 Å), the O atom O2D of Asp84 (3.8 Å), and the N atom NH1 of Arg11 (3.9 Å).

Table 7
Structural refinement of the model of [SeMet]HMBS against the data collected at SRS station 9.5, fitted into the 3-λ MAD+*DM* phased electron-density map (Fig. 5*b*).

Data were measured on station 9.5 at λ = 0.9 Å from the third crystal (for details see text) and at ambient temperature. The resolution range 8.0–2.4 Å comprised 13 486 reflections (99.63% complete) with 13 285 reflections having $|F_o| > 2\sigma(|F_o|)$. The model refinement working set comprised 12 112 reflections (89% completeness) and the R_{free} test set comprised 1374 reflections (10% completeness). Simulated-annealing refinement (4000 K) was used in cycles 1–7. Positional refinement was used in cycles 8–13. Of all 299 residues of the final model, 254 are neither glycine, proline, nor end residues. Of these 254, 234 (92.1%), 17 (6.7%) and 3 (1.2%) are in the most favoured, the additional allowed, and the generously allowed regions of the Ramachandran plot, respectively, according to a *PROCHECK* analysis (Laskowski *et al.*, 1993). The bulk of the main-chain and side-chain parameters are better than the average of structures at this resolution.

Cycle	Main actions taken (apart from regularization)	R_{free}	R
1	Built residues 6–41, 63–110, 113–219, 221–236, 244–302	n.d.	n.d.
2	Added residues 4–5, 111–112, 220, 237–243, 303–307	0.374	0.282
3	Added residues 3, 42–43, adjusted the main chain at residue 243, and introduced one *B* factor per residue	0.363	0.273
4	Added residue 44, adjusted the main chain at residue 199	0.343	0.254
5	Deleted residues 43–44, adjusted the main chain at residue 243, included parameters for Se atoms	0.324	0.243
6	Added the cofactor according to a model in agreement with authentic dipyrromethanes	0.313	0.249
7	Deleted residues 3 and 42, added 34 water molecules	0.305	0.237
8	Introduced individual atomic *B* factors, deleted 4 water molecules	0.299	0.226
9	Added residues 3, 42, 62, 308, added a further 63 water molecules	0.283	0.215
10	Added residues 43–46, 59–61, deleted 23 water molecules, added 28 new water molecules	0.262	0.206
11	Added residues 309–313, added 50 water molecules	0.255	0.189
12	Deleted 7 and added 39 new water molecules	0.245	0.177
13	Deleted 3 water molecules	0.235	0.168

The refined position of this putative water molecule, the *B* factor of which is 26.2 Å^2, at a distance of 3.2 Å from the atom C4B (coordinates x = 19.16, y = 7.03, z = 23.03 Å) is very interesting in the context of the catalytic reaction cycle. The last step of the cycle was proposed to be initiated by protonation of the enzyme-bound hexapyrrole at C4B (Hart *et al.*, 1990). An azafulvene can then be formed by elimination and this reactive intermediate – still bound to the enzyme non-covalently – can react with water to give the product HMB (Battersby *et al.*, 1982).

A water molecule at the position of the spherical density shown in Fig. 8 would be an ideal candidate for being involved in this catalytic mechanism. Since it is proximal to both the electron-rich C4B and the carboxylate group of Asp84, it could, being protonated, initiate the last step of the catalytic cycle by providing a proton to C4B. After elimination of the azafulvene, the now deprotonated water molecule would be ready to capture the reactive intermediate to produce HMB.

Further corroboration of this 'catalytic' water molecule can come from a higher resolution refinement, which we are now undertaking at 1.7 Å. The fact that this is cryo data, in combination with theoretical modelling, may also assist the definition of the missing loop (residues 47–58).

A least-squares superposition of the structures of the reduced, active form (1AH5, MAD-phased) and the oxidized form (1PDA, MIR-phased) of the enzyme was carried out. The different cofactors were obviously removed from the calculation, as were residues 308–313 of the MAD model which were not available in the MIR model. The r.m.s. coordinate displacement on Cα atoms is 0.28 Å and on all atoms is 0.68 Å. Since it is clear in this analysis that residues 40, 60–61, 149 and 241–243 show substantial differences due to the flexible loop and the alteration of the C2 ring position, these were then removed and the r.m.s. values recalculated. These were 0.22 Å on Cα atoms and 0.53 Å on all atoms.

4. Discussion and concluding remarks

This MAD study has yielded an independently derived structure of the HMBS enzyme in its active form at room temperature (PDB deposited coordinate file 1AH5), to 2.4 Å resolution. The reduced, active form of the cofactor takes up a conformation consistent with known dipyrromethane crystal structures. The loop of residues 47–58, proximal to the active site, is still not visible. The positions of the C-terminal residues 308–313 are now determined.

The MAD structure determination, based on the SRS data, proceeded smoothly except for the fact that the anomalous Patterson map quality, based on the λ_3 (white-line f'' maximum) data set collected at ambient temperature (84% completeness), did not reveal all five expected (ordered) seleno sites. Subsequently use of averaged anomalous differences, across all three wavelengths, for the SRS data showed that the calculated features of the selenium positions could be replicated. [SeMet]HMBS also served as a valuable test case to commission the ESRF BM14 CCD detector. Certainly the fact that the CCD detector could be used successfully here encouraged the subsequent wider use of the CCD for MAD projects at ESRF BM14. Indeed the data-collection efficiency of the CCD was ~100% *versus* that of the image-plate system of ~50%; obviously the values depend on the exposure times as well as detector readout times but the trend to quicker readout, whilst preserving accuracy, is clearly advantageous (Deacon *et al.*, 1995; Thompson, 1996).

The structure of HMBS in its active form, presented here, provides a platform for further studies of the enzyme binding with substrate and/or inhibitors. HMBS exhibits several features that make this enzyme an ideal candidate for time-resolved Laue crystallography. First of all, it is a rather slow enzyme [overall k_{cat} = 0.1 s^{-1} at 310 K and pH 7.5 for the saturated enzyme from *E. coli* (Hädener *et al.*, 1993)]. It is possible to slow down the reaction velocity further by using, possibly in combination, mutant variants that show an altered

kinetic behaviour and can accumulate particular ES_i complexes (Jordan & Woodcock, 1991; Lander *et al.*, 1991), lower pH values like those needed to grow crystals of the enzyme, lower temperatures, and substrate concentrations below K_m [for a detailed discussion, see Niemann *et al.* (1994) and Helliwell *et al.* (1998)]. Secondly, crystals of HMBS are relatively radiation-resistant to the monochromatic X-ray beam (as reported here). Also they are resistant to the 'white' X-ray beam, illustrated by the fact that at ESRF ID09 (BL3) a total of 47 Laue exposures could be recorded from one HMBS crystal, each of 1 ms exposure time (Chayen *et al.*, 1996). Furthermore, the enzyme active site is accessible to the solvent channels of the crystal and the time-dependent changes of electron density are expected to be substantial, as a number of aromatic rings should appear during a single catalytic cycle.

We have applied Laue diffraction to establish the time-dependent diffraction 'titration' behaviour of crystals of [Lys59Gln]HMBS (Niemann *et al.*, 1994) as they are exposed to substrate in a flow cell (Helliwell *et al.*, 1997). The refined HMBS MAD active structure has been used as the basis for a refined time-resolved structure of the mutant (Helliwell *et al.*, 1998). Freeze trapping of the enzyme after soaking the crystals in substrate solution, guided by the Laue time-resolved sequence, are now being undertaken on this the active form of the enzyme.

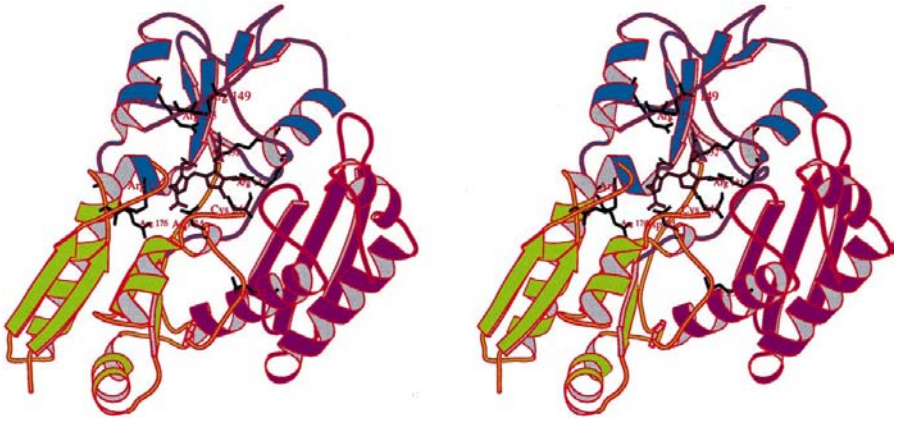

Figure 7
Three-dimensional structure of active, reduced HMBS, solved using the map depicted in Fig. 5(*b*), after refinement (PDB code 1AH5). The dipyrromethane cofactor, attached to the S atom of Cys242, and the side chains of some of the catalytically important residues, are also shown. The ribbon representation, in stereo, was generated by *MOLSCRIPT* (Kraulis, 1991).

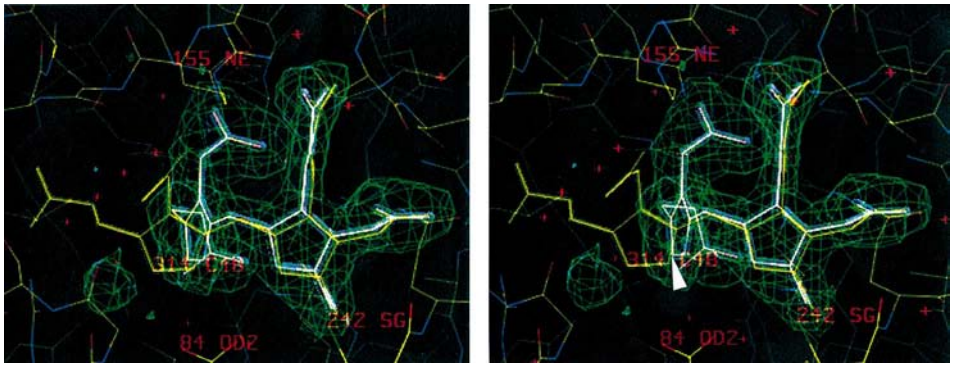

Figure 8
The active site of HMBS, shown in stereo, with the reduced cofactor in white and, superimposed on it, the 3.5σ $F_o - F_c$ difference electron density. The latter was generated by omitting the atoms of the cofactor from the calculation and subjecting the protein model to additional cycles of refinement. In yellow is the model of the oxidized cofactor for comparison. Proximal to the C2 ring atom C4B of the reduced cofactor (labelled with an arrow) is, at left and slightly below, spherical density probably representing a water molecule. This putative water molecule had not been included at all in the refinements. The conserved, catalytic Asp84 is labelled, as are the Cys242 'anchor' to the cofactor C1 ring and Arg155 (whose mutation affects later stages of the catalytic reaction). Note that there is no difference electron density where ring C2 of the oxidized cofactor would be, if the crystal was partly or wholly oxidized. This also confirms that the cofactor, and crystal, is indeed in the reduced form.

AH thanks the Swiss National Science Foundation and the Treubel Fonds, Basle, Switzerland, for financial support of this work. The Science and Engineering Research Council [SERC, UK, now the Engineering and Physical Sciences Research Council (EPSRC)] are thanked for the award of a grant to WNH and JRH for the MAD work in July 1993 at the Synchrotron Radiation Source (SRS) in Daresbury, UK, which supported SJH. ARB thanks the EPSRC for financial support. JRH thanks the Swedish Natural Sciences Research Council (NFR), SERC Daresbury Laboratory and the Medical Research Council (UK) for the funding which allowed SRS station 9.5 to be built. JRH and AWT thank the European Synchrotron Radiation Facility (ESRF), Grenoble, France, for the award of beamtime on BM14 for MAD during part of the commissioning period of the BM14 station and to Jean-Pierre Moy for the development of the CCD detector used in the MAD run in January 1996 as well as EMBL for travel support of the experiment. JRH thanks the European Union for a Host Institute Fellowship Grant for AC and an Overseas Research Scholarship for YPN. The Wellcome Trust (UK) and the Biotechnology and Biological Sciences Research Council provided grant support and 33% salary support of JR. AH and WNH are grateful to Steve Wood and Gordon Louie, Birkbeck College, London, for helpful discussions, and AH thanks Johan N. Jansonius and Kasper Kirschner, Biozentrum, University of Basle, and Reuben Leberman, EMBL outstation, Grenoble, for use of laboratory facilities. Since it was the first time the MAD method was tried with a SeMet-labelled protein at the SRS Daresbury this meant that we had to establish the important details of the protocol for the first time and, accordingly, the

support by members of the Daresbury staff was excellent. Likewise the use of the ESRF CCD detector allowed considerable quantities of data to be collected on ESRF BM14, and we are grateful to ESRF for supporting this development.

References

Allen, F. H., Bellard, S., Brice, M. D., Cartright, B. A., Doubleday, A., Higgs, H., Hummelink-Peters, B. G., Kennard, O., Motherwell, W. D. S., Rodgers, J. R. & Watson, D. G. (1979). *Acta Cryst.* B**35**, 2331–2339.

Baker, E. N., Rumball, S. V. & Anderson, B. F. (1987). *Trends Biochem. Sci.* **12**, 350–353.

Battersby, A. R., Fookes, C. J. R., Gustafson-Potter, K. E., McDonald, E. & Matcham, G. W. J. (1982). *J. Chem. Soc. Perkin Trans. I*, pp. 2427–2444.

Battersby, A. R., Fookes, C. J. R., Matcham, G. W. J. & McDonald, E. (1980). *Nature (London)*, **285**, 17–21.

Battersby, A. R. & Leeper, F. J. (1990). *Chem. Rev.* **90**, 1261–1274.

Becker, W. & Sheldrick, W. S. (1978). *Acta Cryst.* B**34**, 1298–1304.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Biou, V., Leonard, G. A., Stojanoff, V., Laboure, S., Mattenet, M., Helliwell, J. R., Felisaz, F., Claustre, L., Lapeyre, F., Brown, K. & Thompson, A. W. (1997). *ESRF Newslett.* **28**, 21–25.

Bonnett, R., Davies, J. E., Hursthouse, M. B. & Sheldrick, G. M. (1978). *Proc. R. Soc. London Ser. B*, **202**, 249–268.

Bonnett, R., Hursthouse, M. B. & Neidle, S. (1972). *J. Chem. Soc. Perkin Trans. II*, pp. 1335–1340.

Brammer, R. C., Helliwell, J. R., Lamb, W., Liljas, A., Moore, P. R., Thompson, A. W. & Rathbone, K. (1988). *Nucl. Instrum. Methods A*, **271**, 678–687.

Brünger, A. T. (1992). *X-PLOR*, Version 3.1, A System for X-ray Crystallography and NMR. New Haven: Yale University Press.

Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.

Chadwick, D. J. & Ackrill, K. (1994). Editors. *The Biosynthesis of the Tetrapyrrole Pigments*. Chichester: John Wiley.

Chayen, N. E., Boggon, T. J., Cassetta, A., Deacon, A., Gleichmann, T., Habash, J., Harrop, S. J., Helliwell, J. R., Nieh, Y. P., Peterson, M. R., Raftery, J., Snell, E. H., Hädener, A., Niemann, A. C., Siddons, D. P., Stojanoff, V., Thompson, A. W., Ursby, T. & Wulff, M. (1996). *Quart. Rev. Biophys.* **29**, 227–278.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowtan, K. D. & Main, P. (1993). *Acta Cryst.* D**49**, 148–157.

Deacon, A., Habash, J., Harrop, S. J., Helliwell, J. R., Hunter, W. N., Leonard, G. A., Peterson, M., Hädener, A., Kalb (Gilboa), A. J., Allinson, N. M., Castelli, C., Moon, K., McSweeney, S., Gonzalez, A., Thompson, A. W., Ealick, S., Szebenyi, D. M. & Walter, R. (1995). *Rev. Sci. Instrum.* **66**, 1287–1292.

Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.

Engel, N., Gossauer, A., Gruber, K. & Kratky, C. (1993). *Helv. Chim. Acta*, **76**, 2236–2238.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Eschenmoser, A. (1988). *Angew. Chem. Int. Ed. Engl.* **27**, 5–39.

Hädener, A., Alefounder, P. R., Hart, G. J., Abell, C. & Battersby, A. R. (1990). *Biochem. J.* **271**, 487–491.

Hädener, A., Matzinger, P. K., Malashkevich, V. N., Louie, G. V., Wood, S. P., Oliver, P., Alefounder, P. R., Pitt, A. R., Abell, C. & Battersby, A. R. (1993). *Eur. J. Biochem.* **211**, 615–624.

Harrop, S. J., Helliwell, J. R., Wan, T. C. M., Kalb, A. J., Tong, L. & Yariv, J. (1996). *Acta Cryst.* D**52**, 143–155.

Hart, G. J., Abell, C. & Battersby, A. R. (1986). *Biochem. J.* **240**, 273–276.

Hart, G. J., Miller, A. D., Beifuss, U., Leeper, F. J. & Battersby, A. R. (1990). *J. Chem. Soc. Perkin Trans. I*, pp. 1979–1993.

Hart, G. J., Miller, A. D., Leeper, F. J. & Battersby, A. R. (1987). *J. Chem. Soc. Chem. Commun.*, pp. 1762–1765.

Helliwell, J. R. (1984). *Rep. Prog. Phys.* **47**, 1403–1497.

Helliwell, J. R., Hädener, A., Deacon, A., Cassetta, A., Thompson, A. W., Hammersley, A. P. & Moy, J.-P. (1996). *ESRF Highlights 1995/1996*, edited by D. Cornuéjols, pp. 78–79. Grenoble, France: ESRF.

Helliwell, J. R., Nieh, Y. P., Cassetta, A., Raftery, J., Hädener, A., Niemann, A. C., Battersby, A. R., Carr, P. D., Wulff, M., Ursby, T., Moy, J. P. & Thompson, A. W. (1997). *Time-Resolved Diffraction*, edited by J. R. Helliwell & P. M. Rentzepis, pp. 187–194. New York: Oxford University Press.

Helliwell, J. R., Nieh, Y.-P., Raftery, J., Cassetta, A., Habash, J., Carr, P. D., Ursby, T., Wulff, M., Thompson, A. W., Niemann, A. C. & Hädener, A. (1998). *J. Chem. Soc. Faraday Trans.* **94**, 2615–2622.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Jordan, P. M. (1994). *The Biosynthesis of the Tetrapyrrole Pigments*, edited by D. J. Chadwick & K. Ackrill, pp. 70–89. Chichester: John Wiley.

Jordan, P. M., Thomas, S. D. & Warren, M. J. (1988). *Biochem. J.* **254**, 427–435.

Jordan, P. M. & Warren, M. J. (1987). *FEBS Lett.* **225**, 87–92.

Jordan, P. M., Warren, M. J., Mgbeje, B. I. A., Wood, S. P., Cooper, J. B., Louie, G., Brownlie, P., Lambert, R. & Blundell, T. L. (1992). *J. Mol. Biol.* **224**, 269–271.

Jordan, P. M. & Woodcock, S. C. (1991). *Biochem. J.* **280**, 445–449.

Kappas, A., Sassa, S., Galbraith, R. A. & Nordmann, Y. (1995). *The Metabolic and Molecular Bases of Inherited Disease*, edited by C. R. Scriver, A. L. Beaudet, W. S. Sly & D. Valle, pp. 2103–2159. New York: McGraw-Hill.

Kinder, S. H., McSweeney, S. M. & Duke, E. M. H. (1996). *J. Synchrotron Rad.* **3**, 296–300.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Lander, M., Pitt, A. R., Alefounder, P. R., Bardy, D., Abell, C. & Battersby, A. R. (1991). *Biochem. J.* **275**, 447–452.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Louie, G. V., Brownlie, P. D., Lambert, R., Cooper, J. B., Blundell, T. L., Wood, S. P., Malashkevich, V. N., Hädener, A., Warren, M. J. & Shoolingin-Jordan, P. M. (1996). *Proteins Struct. Funct. Genet.* **25**, 48–78.

Louie, G. V., Brownlee, P. D., Lambert, R., Cooper, J. B., Blundell, T. L., Wood, S. P., Warren, M. J., Woodcock, S. C. & Jordan, P. M. (1992). *Nature (London)*, **359**, 33–39.

Miller, A. D., Hart, G. J., Packman, L. C. & Battersby, A. R. (1988). *Biochem. J.* **254**, 915–918.

Mitchell, L. W. & Jaffe, E. K. (1993). *Arch. Biochem. Biophys.* **300**, 169–177.

Moy, J.-P. (1994). *Nucl. Instrum. Methods Phys. Res. A*, **348**, 641–644.

Moy, J.-P., Hammersley, A. P., Svensson, S. O., Thompson, A., Brown, K., Claustre, L., Gonzalez, A. & McSweeney, S. (1996). *J. Synchrotron Rad.* **3**, 1–5.

Niemann, A. C., Matzinger, P. K. & Hädener, A. (1994). *Helv. Chim. Acta*, **77**, 1791–1809.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Peterson, M. R., Harrop, S. J., McSweeney, S. M., Leonard, G. A., Thompson, A. W., Hunter, W. N. & Helliwell, J. R. (1996). *J. Synchrotron Rad.* **3**, 24–34.

Rajeswaran, M. & Parthasarathy, R. (1984). *Acta Cryst.* C**40**, 647–650.

Reagen, W. K. & Radonovich, L. J. (1989). *J. Am. Chem. Soc.* **111**, 3881–3886.

Sarra, R., Garratt, R., Gorinsky, B., Jhoti, H. & Lindley, P. F. (1990). *Acta Cryst.* B**46**, 763–771.

Sasaki, S. (1989). KEK Report No. 88–14. National Laboratory for High-Energy Physics, Tsukuba, Japan.

Scott, A. I., Roessner, C. A., Stolowich, N. J., Karuso, P., Williams, H. J., Grant, S. K., Gonzalez, M. D. & Hoshino, T. (1988). *Biochemistry*, **27**, 7984–7990.
Sessler, J. L., Mody, T. D., Ford, D. A. & Lynch, V. (1992). *Angew. Chem.* **104**, 461–464.
Sheldrick, W. S. & Becker, W. (1979). *Z. Naturforsch. B*, **34**, 1542–1546.
Spurlino, J., Lu, G.-Y. & Quiocho, F. A. (1991). *J. Biol. Chem.* **266**, 5202–5219.
Stark, W. M., Baker, M. G., Raithby, P. R., Leeper, F. J. & Battersby, A. R. (1985). *J. Chem. Soc. Chem. Commun.*, pp. 1294–1296.
Thompson, A. W. (1996). *ESRF Highlights 1995/1996* edited by D. Cornuéjols, pp. 54–56. Grenoble, France: ESRF.
Thompson, A. W. (1997). *ESRF Beamline Handbook*, edited by R. Mason, pp. 105–107. Grenoble, France: ESRF.
Thompson, A. W., Habash, J., Harrop, S., Helliwell, J. R., Nave, C., Atkinson, P., Hasnain, S. S., Glover, I. D., Moore, P. R., Harris, N., Kinder, S. & Buffey, S. (1992). *Rev. Sci. Instrum.* **63**, 1062–1064.
Weiss, M. S. & Hilgenfeld, R. (1997). *J. Appl. Cryst.* **30**, 203–205.