Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

MAD analyses of yeast 5-aminolaevulinate dehydratase: their use in structure determination and in defining the metal-binding sites

P. T. Erskine,[a] E. M. H. Duke,[b] I. J. Tickle,[c] N. M. Senior,[d] M. J. Warren[d] and J. B. Cooper[a]*

[a]Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, England, [b]CCLRC Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, England, [c]Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, England, and [d]School of Biological Sciences, Queen Mary and Westfield College, Mile End Road, London E1 4NS, England

Correspondence e-mail: jbc2@soton.ac.uk

Received 15 April 1999
Accepted 10 January 2000

PDB References: Hg complex, 1qml; Pb complex, 1qnv.

MAD experiments attempting to solve the structure of 5-aminolaevulinic acid dehydratase using Zn and Pb edges are described. The data obtained proved insufficient for a complete structure solution but were invaluable in subsequent identification of metal-binding sites using anomalous difference Fourier analyses once the structure of the enzyme had been solved. These sites include the highly inhibitory substitution of an enzymic cofactor Zn^{2+} ion by Pb^{2+} ions, which represents a major contribution towards understanding the molecular basis of lead poisoning. The MAD data collected at the Pb edge were also used with isomorphous replacement data from the same Pb co-crystal and a Hg co-crystal to provide the first delineation of the enzyme's quaternary structure. In this MADIR analysis, the Hg co-crystal data were treated as native data. Anomalous difference Fouriers were again used, revealing that Hg^{2+} had substituted for the same Zn^{2+} cofactor ion as had Pb^{2+}, a finding of fundamental importance for the understanding of mercury poisoning. In addition, Pt^{2+} ions were found to bind at the same place in the structure. The refined structures of the Pb- and the Hg-complexed enzymes are presented at 2.5 and 3.0 Å resolution, respectively.

1. Introduction

The enzyme 5-aminolaevulinic acid dehydratase (ALAD or porphobilinogen synthase, E.C. 4.2.1.24) catalyses an early step common to the biosynthetic pathways leading to the tetrapyrroles. ALAD joins two molecules of 5-aminolaevulinic acid (ALA) to form the pyrrole porphobilinogen (PBG). Four PBG molecules are joined up by subsequent enzymes to form uroporphyrinogen III. From these key precursors, the biosynthetic pathway then diverges to such biologically important tetrapyrroles as the corrins, chlorophyll and haem. ALAD has been reported to perform separate functions – human ALAD as an inhibitory component of the 26S proteasome (Guo *et al.*, 1994) and rabbit reticulocyte ALAD as a promoter of renaturation by heat shock protein 70 (Gross *et al.*, 1999).

The enzyme is homo-octameric ($M_r \simeq 300$ kDa). Biochemical and EXAFS data indicate that the enzyme possesses at least two zinc cofactor ions in each subunit (Dent *et al.*, 1990; Senior *et al.*, 1996). A distinguishing characteristic of the enzyme is its apparent sensitivity to metal ions (Anderson & Desnick, 1979; Gibbs & Jordan, 1981; Schroeder & Caspers, 1996), most notably lead (Simons, 1995). Inactivation of the enzyme is a central feature of lead poisoning, the neurological symptoms of which are thought to stem from an accumulation of substrate ALA which has been shown to be neuroactive

(Warren *et al.*, 1998, and references therein). Recently, we solved the structure of ALAD from *Saccharomyces cerevisiae* (Erskine, Senior, Awan *et al.*, 1997) and *Escherichia coli* (Erskine *et al.*, 1999). Here, we describe the work stemming from attempts to try to solve the structure of the yeast enzyme using multiwavelength anomalous diffraction (MAD) methods on the zinc cofactor ions and bound lead ions. It is worth noting that the experiments were undertaken at a time when MAD was not the routine technique it may now be considered to be. The phasing information from both the Zn-edge and the Pb-edge MAD experiments proved insufficient to solve the structure, but the data from the Pb-edge experiment were used to produce an electron-density map showing for the first time the overall arrangement of the octameric quaternary structure. Following successful structure analysis by the selenomethionyl MAD approach, the Zn, Hg and Pb data sets subsequently yielded a large amount of information on how metal ions interact with and inhibit ALAD. For example, the main active-site zinc ion was located by use of the Zn^{2+} anomalous data. We also report the refined structures of the Pb and Hg complexes of ALAD refined at 2.5 and 3.0 Å resolution, respectively. Both metal ions replace the enzyme's catalytic zinc ion bound by three conserved cysteine side chains. In addition, there are a number of minor metal-binding sites in the active site of the enzyme and some evidence of disruption to the cysteine residues which normally coordinate the catalytic zinc ion. Hereditary deficiencies in ALAD have neurological symptoms (Doss *et al.*, 1979; Thunell *et al.*, 1987; Hassoun *et al.*, 1989; Mercelis *et al.*, 1990). A clearer picture of the molecular events that contribute to the neurological symptoms of poisoning by these and, most likely, other heavy-atom reagents is provided.

2. Experimental

2.1. Data collection

Full details of the crystallization methods have been described previously (Erskine, Senior, Maignan *et al.*, 1997). The crystals were grown using the hanging-drop vapour-diffusion method and grew in space group *I*422. Co-crystallizations were set up by the introduction of 1 µl of a 10 m*M* heavy-atom solution to a 10 µl hanging drop. The heavy-atom compounds used were lead acetate and mercury chloride. All data collection was carried out at 100 K. The crystals were flash-cooled by immersion in liquid ethane prior to storage under liquid nitrogen. In general, the crystals were cryoprotected by slow addition of glycerol to around 30%(*v*/*v*).

Multiwavelength data were collected at Station 9.5 (PX 9.5) at the Synchrotron Radiation Source (SRS) at Daresbury. The Zn *K* absorption edge of native ALAD and the Pb $L_{\rm III}$ absorption edge of a lead co-crystal were utilized. A calculation was made to estimate whether the Bijvoet signal from the two Zn atoms in each subunit would be strong enough for MAD phasing using the approximate values for f'' calculated for unbound atoms by Cromer & Liberman (1970). The formula $(1/2)^{1/2}(N_A/N_P)^{1/2}(2f''/Z_{\rm eff})$ was used, where N_A is the number of anomalously scattering atoms per molecule, N_P is the number of non-H atoms in the molecule and $Z_{\rm eff}$ is the average atomic number of non-H atoms in proteins (~6.7; Krishna Murthy, 1996). Although this showed that the signal would be borderline (2.3% of the average structure-factor amplitude on the basis of an M_r of 37 800 Da and an f'' of 3.87), the experiment proceeded on the basis that the phasing information would be useful in combination with data from derivatives.

The silicon 111 monochromator on Station 9.5 beamline had $2d_{\rm Si111}$ = 6.271 Å. The detector was a 30 cm MAR Research image plate. The fluorescence detector for analysis of the absorption-edge structure (XANES) was placed beside the collimator about 5 cm away from the crystal. X-ray absorption spectra were collected using a Daresbury-written program to control the monochromator and a Canberra S100 multichannel analyser and software on an IBM laptop to determine the windowing for the signal. An Oxford Cryosystems Cryostream was used to maintain the crystals at 100 *K* and the goniostat was a three-circle Enraf–Nonius CAD-4 goniometer.

Large imperfect crystals from the same crystallization setups as the crystals used in the main experiment were used for the scans of the X-ray absorption edges. No absorption edge for mercury was detected in the XANES scans of the mercury co-crystals for the Hg $L_{\rm III}$ edge (λ = 1.0093 Å). The scan for the Zn *K* edge at 1.283 Å was more successful and is shown in Fig. 1. From this plot, the wavelengths for collection on a native crystal were selected as 1.2808 Å for a value of f' close to zero (λ_1), 1.2817 Å for the peak in f'' (λ_2) and 1.2826 Å for the maximal negative peak in f' (λ_3).

In a MAD experiment, alignment of the crystal can be performed in order to maximize the collection of Friedel pairs

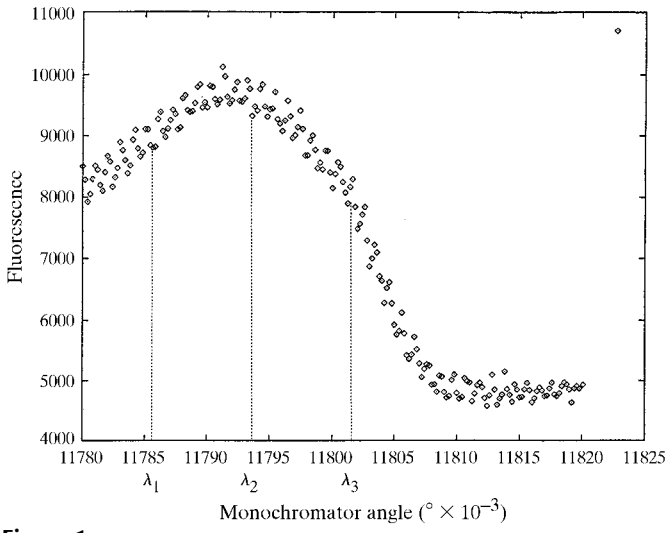

Figure 1
XANES scan for the Zn *K* edge at 1.283 Å made at PX 9.5 at the SRS for the Zn MAD data set. From this scan, the following wavelengths for collection were selected: λ_1, 1.2808 Å (11785.5° × 10^{-3}) for a value of f' close to zero; λ_2, 1.2817 Å (11793.5° × 10^{-3}) for the peak in f''; λ_3 1.2826 Å (11801.6° × 10^{-3}) for the maximal negative peak in f' occurring at the point of inflection in the rise to the peak in f''.

in the same image in order to minimize the effect of scaling errors on the anomalous differences. The native crystal could only be partially aligned with the a^*b^* diagonal as the axis of rotation. A rotation-block collection schedule was used, designed to ensure that the data at three different wavelengths were collected close together in time. Each rotation block was a set of six images recording two images for each wavelength. The oscillation angle for each image was 2.5°. The exposure time for each image was 480 s; 50° of data were collected to 2.9 Å resolution at each of the three wavelengths.

For the lead co-crystal, data to a resolution of 2.5 Å were collected at $\lambda = 0.875$ on Station 9.6 (PX 9.6) at the SRS prior to the multiwavelength collection. This was used as the remote wavelength in an analysis using data from both multi-wavelength anomalous diffraction and isomorphous replacement (MADIR). The same crystal was used for the PX 9.5 MAD collection, in which it was partially aligned. The XANES scan for the lead L_{III} edge is shown in Fig. 2. Wavelengths were selected from the absorption-edge scan in such a way as to maximize the chances of still collecting at the peak in f'' and the maximal negative peak in f' even if the beam shifted slightly during the course of collection. This was performed by choosing an additional interleaving wavelength between the wavelengths of these two maxima. The wavelengths chosen from the plot shown in Fig. 2 were $\lambda_1 = 0.94945$ Å to collect the peak in f'', $\lambda_2 = 0.95002$ Å to collect the interleaving wavelength and $\lambda_3 = 0.95036$ Å to collect the maximal negative peak in f'. The data collection provided 95° of data to 3.0 Å resolution at each wavelength.

Single-wavelength data also mentioned in the following analyses were collected on a native crystal on PX 9.5 ($\lambda = 1.13$ Å, 2.3 Å resolution) and a mercury co-crystal on PX 9.6 ($\lambda = 0.875$, 3.0 Å resolution).

In the anomalous difference Fourier syntheses the phases were those derived from the refined ALAD structure resulting from the selenomethionyl MAD analysis.

2.2. Data processing

All data sets were processed using *MOSFLM* (Leslie, 1992) and were scaled and merged using *ROTAVATA* and *AGROVATA* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). The MAD data collected at three wavelengths around the Zn edge of the native crystal were scaled to the λ_3 data using *SCALEIT* (Collaborative Computational Project, Number 4, 1994). Dispersive difference Pattersons were calculated for λ_3 against λ_1 with varying amounts of data excluded. No non-axial peaks above 4 r.m.s. were found in the Harker sections. Anomalous difference Pattersons were calculated for λ_2 with no data excluded and with those reflections with anomalous differences greater than 12 times the mean $|\Delta_{ano}|$ excluded. In the latter Patterson, there was a peak of ~7 r.m.s. in the Harker section at $u = 0.5$ (Fig. 3). A vector search was made in the Harker sections of this map using the program *VECSUM*. This showed the site to have the most promising occupancy but that it was on a diagonal (and thus liable to have a contribution from noise). Vector searching using this site as the basis for a superposition function failed to reveal any other promising sites. The site was entered on its own into a pseudo-MIR phasing approach in *MLPHARE* using λ_3 as 'native', but the phasing statistics were very poor (overall FOM = 0.1).

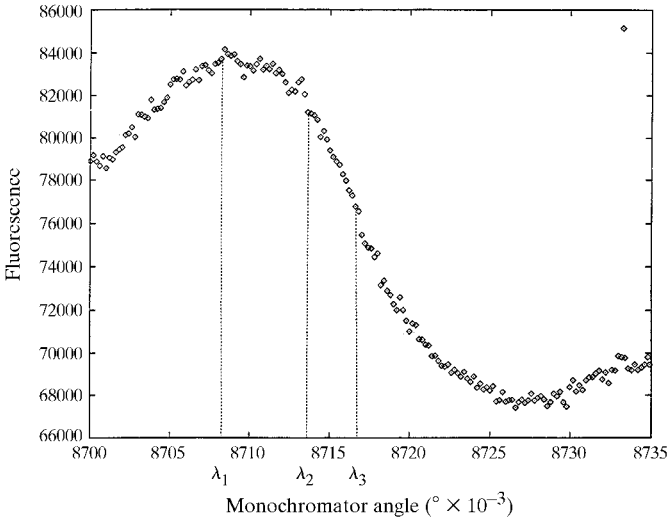

Figure 2
XANES scan for the Pb L_{III} edge at 0.9511 Å made at PX 9.5 at the SRS for the Pb MAD data collection. In order to maximize the chances of still collecting at the peak in f'' and the maximal negative peak in f' even if the beam shifted during the collection, an interleaving wavelength was collected between the expected wavelengths of these two maxima. The wavelengths chosen from this plot were therefore: λ_1, 0.94945 Å (8708.3° × 10^{-3}) to collect the peak in f''; λ_2, 0.95002 Å (8713.6° × 10^{-3}) to collect the interleaving wavelength; λ_3, 0.95036 Å (8716.7° × 10^{-3}) to collect the maximal negative peak in f'.

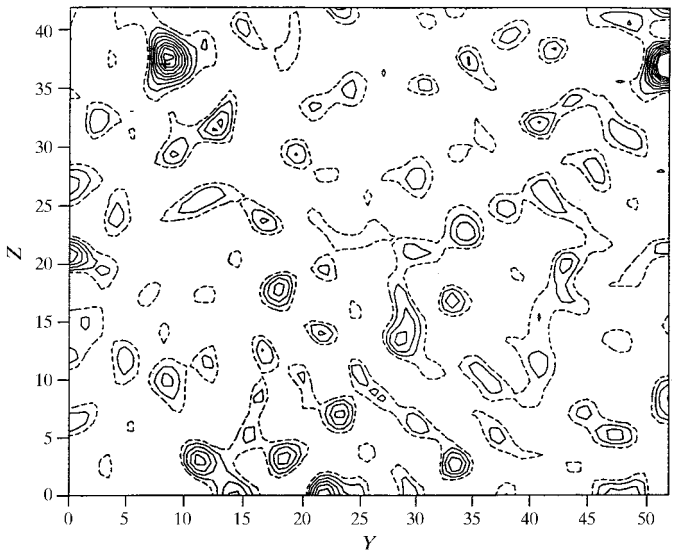

Figure 3
S. cerevisiae ALAD Zn edge crystal (PX 9.5 MAD data) anomalous difference Patterson λ_2 at 2.9 Å resolution $u = \frac{1}{2}$ Harker section. The site marked by a cross (top left) was in the same position as a lead site from the isomorphous replacement work, refined to a promising value but was on a diagonal. In a superposition function in *VECSUM* it failed to bring back other sites and in a pseudo-MIR MAD approach in *MLPHARE* using λ_3 as native the phasing statistics were very poor. The map is contoured at 83% r.m.s.

Table 1
Summary of statistics for *S. cerevisiae* ALAD native and co-crystal data.

All the data sets were collected at 100 K. The space group for all crystals was *I*422.

Data set	Resolution† (Å)	Unit-cell parameters (Å); $a = b$, c	R_{merge}‡ (%)	Completeness (%)	$I > 3\sigma(I)$ (%)
Native	2.3	103.7, 167.8	4.5	95.8	93.9
$HgCl_2$ co-crystal	3.0	102.1, 170.5	9.9	99.0	83.5
$Pb(CH_3CO_2)_2$ co-crystal	2.5	102.4, 170.5	9.8	93.0	81.7

† R_{merge} is defined as $\sum_{i=1}^{n} |I(hkl)_i - \bar{I}(hkl)| / \sum_{hkl} \sum_{i=1}^{n} I(hkl)_i$. ‡ Resolution to which data were collected, completeness and percentage of data $>3\sigma(I)$ are given to this resolution.

Table 2
Refined positions (in Å) and occupancies of lead sites from PX 9.6 SIR data from *S. cerevisiae* ALAD lead acetate co-crystal, using the mercury co-crystal as native, using Patterson refinement with *VECREF* at 3.0 Å resolution (see also Fig. 5).

Atom	*X* (Å)	*Y* (Å)	*Z* (Å)	Relative occupancy
Pb 1	26.64	23.75	23.14	0.34
Pb 2	22.87	21.86	23.14	0.50
R.m.s. residual†	0.624			
Correlation‡	0.896			

† R.m.s. residual is the root-mean-square difference between observed and calculated values for the Patterson map at positions corresponding to the termini of vectors calculated from the input sites. ‡ Correlation is the correlation coefficient between the observed and calculated values in the Patterson map.

The MAD data collected at the Pb edge of the lead co-crystal were locally scaled to the λ_1 data set. Following analysis of the size of the dispersive and anomalous differences using *SCALEIT*, dispersive difference Pattersons were calculated for λ_3 against λ_1 and, in case the beam had shifted making λ_2 closer to the peak in f', for λ_2 against λ_1. No non-axial peaks above 4 r.m.s. were seen in the Harker sections. However, in an anomalous difference Patterson calculated for λ_1, a non-axial peak of about 5 r.m.s. was found (Fig. 4). Vector searching in the Harker sections identified a site which corresponded with this peak, but again a phasing attempt based on this site proved unsatisfactory (*MLPHARE* FOM = 0.1).

The mercury co-crystal and lead co-crystal data which had been collected on PX 9.6 were tried in isomorphous replacement against the native crystal collected at λ = 1.13 Å on PX 9.5 (a summary of data-quality statistics is shown at Table 1). However, analysis in *SCALEIT* following local scaling showed that both co-crystals had an isomorphous difference *R* factor which was notably high against this native data set (42.4 and 44.8%, respectively). The isomorphous difference Patterson for the mercury co-crystal showed no non-axial peaks higher than 4 r.m.s.. A vector search was conducted in the Harker sections with *VECSUM*, but failed to produce any sites whose occupancy refined to a promising value.

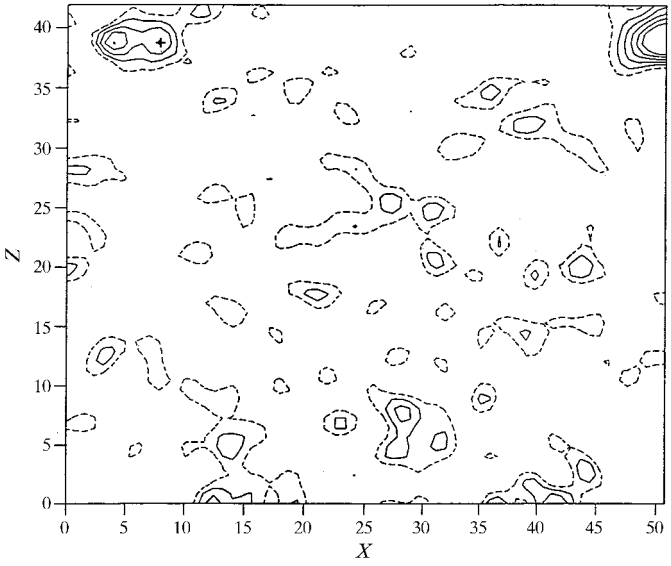

Figure 4
S. cerevisiae ALAD Pb edge (PX 9.5 MAD λ_1 data) anomalous difference Patterson at 3.0 Å resolution, $u = \frac{1}{2}$ Harker section. The site marked by a cross was tried in a superposition function in *VECSUM* but failed to bring back other sites. It produced poor phasing statistics in *MLPHARE*. From the isomorphous replacement processing of the PX 9.6 data from this crystal against the isomorphous Hg co-crystal the site corresponds to one of the Pb sites. The major axial peak at the right corresponds to peaks in that processing found to have arisen from cross vectors between the two lead sites. The map is contoured at 94% r.m.s.

In view of the XANES scans during the multiwavelength experiment, which indicated that there was no mercury present in the Hg co-crystals, it was decided to treat the mercury co-crystal as a native and see whether it might be more isomorphous with the lead co-crystal. After local scaling, the isomorphous difference *R* factor between these crystals was found to be 17.8%. An isomorphous difference Patterson was calculated at 3.0 Å. A vector search of this Patterson using *VECSUM* produced a site from the Harker sections in the same position as those seen in the earlier Zn and Pb MAD analyses which had not proved satisfactory in phasing. However, by treating the Hg complex data as the native data set, the lead site's occupancy refined to a promising value. When this site was entered into a superposition function, a second peak close to the input site could be seen (Fig. 5). The occupancy of the new site refined to a higher value than that of the input site (see Table 2 for the occupancies and refined positions). Both sites were entered into *MLPHARE*, yielding an overall FOM of 0.31, a phasing power for centrics of 1.12 and a Cullis *R* factor for centrics of 0.80 (see Table 3 for a full summary of this refinement). Solvent flattening was carried out with *DM* and an electron-density map of the native was calculated at 3.0 Å using *FFT*. There was a clear solvent boundary, but the map was not interpretable. The anomalous data were found not to improve the map and the anomalous differences were thought not to be sufficiently large. Searches for further sites were unfruitful. A difference Fourier with the coefficients $F_{PH} - F_P$ was calculated using the improved phases output by *DM* following refinement of one of the sites alone with *MLPHARE*. This gave back the other site, tending to confirm, independently of the isomorphous difference Patterson, the site's reality.

Table 3
Refined positions (in Å) and occupancies of lead sites from PX 9.6 SIR data from *S. cerevisiae* ALAD lead acetate co-crystal, using the mercury co-crystal as native, using *MLPHARE* (3.0 Å resolution, see also Fig. 5).

Atom	*X* (Å)	*Y* (Å)	*Z* (Å)	Real relative occupancy	Anomalous relative occupancy
Pb 1	26.74	23.88	23.36	0.219	—
Pb 2	22.86	21.84	23.02	0.317	—

Phasing statistics	
Phasing power† (acentrics)	1.50
Cullis *R* factor‡ (acentrics)	0.92
Phasing power (centrics)	1.12
Cullis *R* factor (centrics)	0.80
FOM§	0.3097

† Phasing power = $\langle|F_H|\rangle/\langle$(lack of closure)$\rangle$, where (lack of closure) = $\big||F_{PH\text{obs}}| - |F_{PH\text{calc}}|\big|$. ‡ Cullis *R* factor = $\langle$(lack of closure)$\rangle/\langle|F_{PH} - F_P|\rangle$. § FOM = $\sum_\varphi P_\varphi \cos(\Delta_\varphi)/\sum P_\varphi$.

It was then decided to try a MADIR approach using these sites obtained from isomorphous replacement, using the data collected on the mercury and lead co-crystals at PX 9.6 with the data from the three Pb wavelengths collected from the same Pb co-crystal on PX 9.5. The five data sets were local scaled together and phase calculation was performed at 3.0 Å resolution using *MLPHARE* on the two sites identified in the isomorphous difference Patterson. The phasing statistics were improved, with an overall FOM of 0.6, a phasing power for centrics of 1.15 and a Cullis *R* factor for centrics of 0.70 (see Table 4 for the occupancies and refined positions of the sites in angstroms and Table 5 for a full summary of the refinement giving details for each wavelength and positions of the sites in

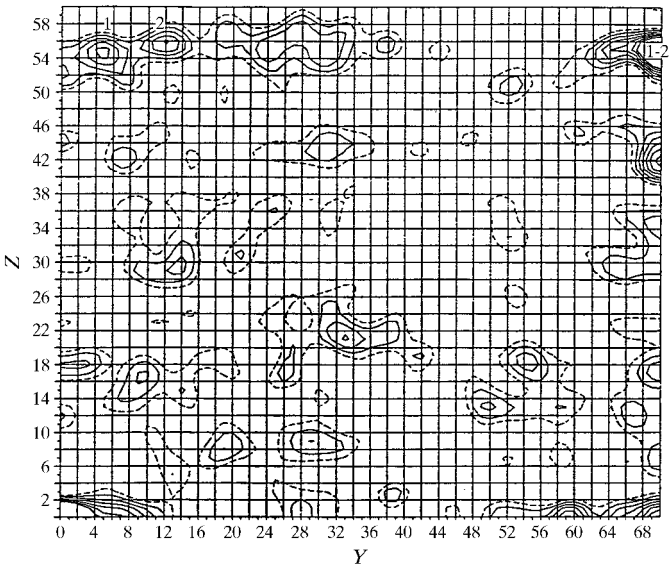

Figure 5
Lead acetate co-crystal PX 9.6 SIR data from *S. cerevisiae* ALAD were scaled against data from a mercury co-crystal. An isomorphous difference Patterson was calculated at 3.0 Å. The $u = \frac{1}{2}$ Harker section is shown. The map is contoured at 71% r.m.s. A vector search produced site 1. A superposition function using this site produced site 2, whose occupancy refined to a higher value. See also Tables 2 and 3.

Table 4
Refined positions (in Å) and occupancies of lead sites from MADIR approach using PX 9.6 SIR data and PX 9.5 MAD data from *S. cerevisiae* ALAD lead acetate co-crystal, treating the mercury co-crystal as native, using *MLPHARE* (3.0 Å resolution).

The figures given here are for the PX 9.6 data set ($\lambda = 0.875$); see Table 5 for the refined lead positions in fractional coordinates and occupancies for all four wavelengths.

Atom	*X* (Å)	*Y*(Å)	*Z*(Å)	Real relative occupancy	Anomalous relative occupancy
Pb 1	26.58	23.74	23.35	0.273	0.066
Pb 2	22.72	21.40	22.68	0.321	0.214

Phasing statistics	
Phasing power† (acentrics)	1.90
Cullis *R* factor‡ (acentrics)	0.63
Phasing power (centrics)	1.15
Cullis *R* factor (centrics)	0.70
FOM§	0.6163

† Phasing power = $\langle|F_H|\rangle/\langle$(lack of closure)$\rangle$, where (lack of closure) = $\big||F_{PH\text{obs}}| - |F_{PH\text{calc}}|\big|$. ‡ Cullis *R* factor = $\langle$(lack of closure)$\rangle/\langle|F_{PH} - F_P|\rangle$. § FOM = $\sum_\varphi P_\varphi \cos(\Delta_\varphi)/\sum P_\varphi$.

fractional coordinates). Solvent flattening was performed with *DM* and, using the improved phases from this, an electron-density map was calculated at 3.0 Å resolution for the asymmetric unit. The map again had a clear solvent boundary. The map was extended using *EXTEND* allowing the visualization of the overall organization of the octamer (Fig. 6; inter-

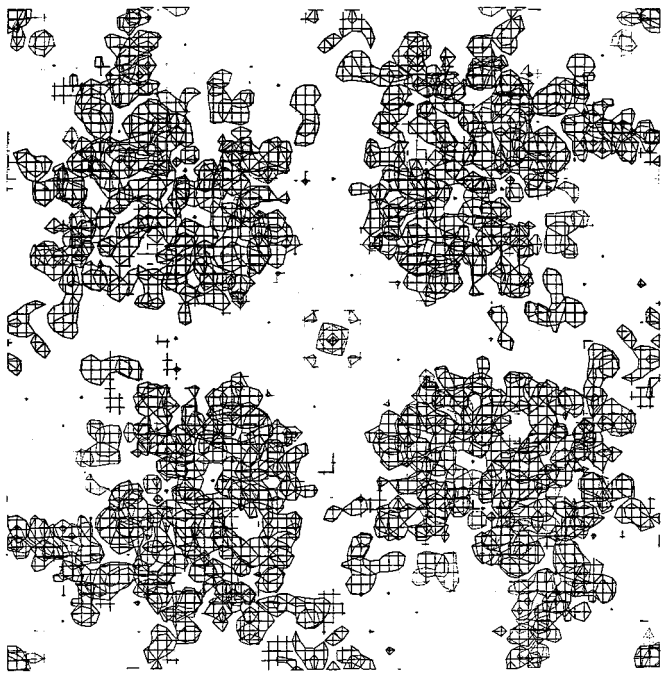

Figure 6
A MADIR approach with data from the *S. cerevisiae* ALAD lead acetate co-crystal allowed the calculation of an electron-density map at 3.0 Å resolution in which there was a clear solvent boundary. The map was extended to show the overall organization of the octamer. Here, four subunits of the ALAD octamer can be seen disposed around the crystallographic fourfold axis.

Table 5
Refined positions from *MLPHARE* in fractional coordinates and occupancies of lead sites from the MADIR approach using PX 9.6 SIR data and PX 9.5 MAD data from the *S. cerevisiae* ALAD lead acetate co-crystal, treating the mercury co-crystal as native (3.0 Å resolution).

		Fractional coordinates				
Wavelength	Atom	*X*	*Y*	*Z*	Real relative occupancy	Anomalous relative occupancy
λ_1, PX 9.5 MAD, max f''	Pb 1	0.262	0.234	0.138	0.250	0.049
	Pb 2	0.224	0.211	0.134	0.287	0.175
λ_2, PX 9.5 MAD, interleaf	Pb 1	0.262	0.234	0.138	0.248	0.049
	Pb 2	0.224	0.211	0.135	0.285	0.166
λ_3, PX 9.5 MAD, max f'	Pb 1	0.262	0.234	0.138	0.248	0.048
	Pb 2	0.224	0.211	0.134	0.284	0.165
Remote PX 9.6, SIR, λ = 0.875	Pb 1	0.262	0.234	0.138	0.273	0.066
	Pb 2	0.224	0.211	0.134	0.321	0.214

Phasing statistics

	λ_1, max f''	λ_2, interleaf	λ_3, max f'	Remote
Phasing power† (acentrics)	1.97	1.98	1.97	1.90
Cullis *R* factor‡ (acentrics)	0.60	0.60	0.60	0.63
Phasing power (centrics)	1.13	1.13	1.14	1.15
Cullis *R* factor (centrics)	0.74	0.74	0.74	0.70
FOM§	0.6163			

† Phasing power = $\langle|F_H|\rangle/\langle$(lack of closure)$\rangle$, where (lack of closure) = $\big||F_{PH\rm obs}| - |F_{PH\rm calc}|\big|$. ‡ Cullis *R* factor = $\langle$(lack of closure)$\rangle/\langle|F_{PH} - F_P|\rangle$. § FOM = $\sum_\varphi P_\varphi \cos(\Delta_\varphi)/\sum P_\varphi$.

Table 6
Summary of statistics for PX 9.5 MAD experiments on *S. cerevisiae* ALAD.

Data set	λ (Å)	Resolution (Å)	$R_{\rm merge}$ (%)	$I > 3\sigma(I)$ (%)	$\lvert\Delta_{\rm ano}\rvert > 3\sigma(\Delta_{\rm ano})$ (%)	Mean multiplicity	Completeness (%)
Zn λ_1 reference	1.2808	2.9	5.9	94.8	3.1	4.7	81.6
λ_2 max f''	1.2817	2.9	5.2	94.5	3.4	4.7	81.5
λ_3 max f'	1.2826	2.9	5.2	94.7	2.5	4.7	81.4
Pb λ_1 max f''	0.94945	3.0	10.6	89.9	5.1	7.4	95.3
λ_2 interleaf	0.95002	3.0	10.5	90.1	4.5	7.4	95.2
λ_3 max f'	0.95036	3.0	10.3	90.2	4.6	7.4	95.2

pretative work was undertaken using the density skeletonization program *BONES* from the *O* suite; Jones *et al.*, 1991), but the map proved not to possess sufficient detail for successful interpretation.

3. Results

The structure of *S. cerevisiae* ALAD is described elsewhere (Erskine, Senior, Awan *et al.*, 1997). Here, a description is given of how the data from the Zn- and Pb-edge MAD experiments, while being unable to provide sufficient phasing power for a successful structure solution, proved invaluable in defining in detail where the metal ions bind in the final structure and what the implications of this binding are for inhibition and catalysis.

A summary of the statistics for the multiwavelength data is shown in Table 6, from which it can be seen that for both the Zn MAD and Pb MAD data collections the percentages of anomalous differences at each wavelength which were $>3\sigma(\Delta_{\rm ano})$ are very small. The statistics for the eventual solution provided by selenomethionyl MAD are shown in Table 7, where it can be seen that the anomalous signal is much stronger, especially at the white line (λ_2). The initial map for the protein obtained by selenomethionine MAD indicated that there was a possible metal-binding site between three cysteine residues (133, 135 and 143; yeast numbering). This was confirmed using the Zn^{2+} optimized anomalous data set from the Zn-edge multiwavelength experiment (λ_1) in an anomalous difference Fourier which was found to have its strongest peak at this location. In addition to the main site, there appeared to be a less well defined site involving His142 and Cys234 (this latter residue had been previously identified as a putative zinc ligand after covalent modification using 5-chlorolaevulinic acid; Jaffe *et al.*, 1992). This second site was on the border of a disordered region between residues 220 and 233.

These Zn^{2+} sites were found to appear in both anomalous difference and dispersive difference Fouriers phased using phases calculated from the model. The data from the Zn edge collected at λ_3 and λ_1 were used for the dispersive difference Fourier. The triple cysteine site produced a peak of 7.85 r.m.s. in the anomalous difference Fourier and 7.61 r.m.s. in the dispersive difference Fourier. The less well defined site produced the second highest peak in each at 4.88 and 4.31 r.m.s., respectively. In the dispersive difference Fourier a third peak at 4.30 r.m.s. occurred, indicating a possible metal site near Cys285.

To locate the lead ions bound to ALAD in the Pb co-crystal, an anomalous difference Fourier was calculated using the highest resolution data set (2.5 Å) collected from the Pb co-crystal on PX 9.6 (λ = 0.875 Å), using phases calculated from the selenomethionyl MAD solution but omitting the Zn atoms from the calculation. This showed that *S. cerevisiae* ALAD binds lead ions principally at the triple cysteine site. The peak height in this Fourier for this site was 9.74 r.m.s. Two additional volumes of positive difference density appear within 4.5 Å of this main site and are probably of lower occupancy. When the data from λ_1 of the MAD data collection to 3.0 Å were used to calculate an anomalous difference Fourier, the peak height at the principal triple cysteine site was increased to 17.89 r.m.s. The heights of the adjacent peaks in the λ_1 anomalous difference Fourier were 6.99 r.m.s. and 6.52 r.m.s., with a third adjacent peak appearing at 6.50 r.m.s.. Surprisingly, calculation of a dispersive difference Fourier using λ_3 against λ_1 produced no significant peaks. However, a dispersive difference Fourier using λ_3 against the remote wavelength

Table 7
Summary of statistics for ESRF BM14 Se-edge MAD experiment on selenomethionyl *S. cerevisiae* ALAD.

Data set	λ (Å)	Resolution (Å)	R_{merge} (%)	$I > 3\sigma(I)$ (%)	$\lvert\Delta_{ano}\rvert > 3\sigma(\Delta_{ano})$ (%)	Mean multiplicity	Completeness (%)
Se λ_1, max f'	0.97934	2.45	5.0	92.1	5.1	9.6	99.4
λ_2, max f''	0.97913	2.45	5.1	92.1	19.0	9.5	99.4
λ_3, reference	0.95377	2.40	4.9	91.5	8.7	9.4	99.5
Se high-resolution opt. anom.	0.97913	2.3	8.8	92.3	7.7	16.8	99.9

Table 8
X-ray refinement parameters for Pb- and Hg-complexed yeast ALAD at 2.5 and 3.0 Å, respectively.

The data were collected at PX 9.6 at the SRS (λ = 0.875 Å).

	Pb	Hg
R_{merge}† (%)	9.8	9.9
Outer shell R_{merge} (%)	24.3	26.8
Completeness (%)	93.0	99.0
$I \geq 3\sigma(I)$ (%)	81.7	83.5
Mean multiplicity	6.1	7.5
R factor (%)	25.2	24.9
R_{free} (%)	37.9	33.0
Correlation coefficient, working set (%)	93.8	94.2
Correlation coefficient, free test set (%)	72.2	74.0
Resolution range (Å)	20–2.5	14.5–3.0
$\sigma(F)$ cutoff	3.0	3.0
Number of reflections used	13008	8224
Number of parameters refined	11823	8576
Number of geometric restraints	6730	6814
R.m.s. deviation of bond lengths (Å)	0.006	0.009
R.m.s. deviation of 'angle distances' (Å)	0.019	0.028
R.m.s. deviation for van der Waals contacts (Å)	0.026	0.045
R.m.s. deviation of main-chain planes (Å)	0.012	0.008
R.m.s. deviation of side-chain planes (Å)	0.011	0.008
R.m.s. deviation of chiral tetrahedra (Å)	0.010	0.015
Mean isotropic B factor main-chain atoms (Å^2)	29.3 (1312 atoms)	Not refined
Mean isotropic B factor side-chain atoms (Å^2)	29.7 (1234 atoms)	Not refined
Mean isotropic B factor for solvents (Å^2)	34.3 (405 atoms)	Not refined

† R_{merge} is defined as $\sum_{i=1}^{n} |I(hkl)_i - \bar{I}(hkl)| / \sum_{hkl} \sum_{i=1}^{n} I(hkl)_i$.

produced a peak at 8.08 r.m.s. at the triple cysteine site with no other significant peaks.

Although the Hg co-crystal failed to give a XANES edge characteristic of Hg and had been treated as the native in phase calculation, it was decided to see if Hg had bound to the protein at all. An anomalous difference Fourier was calculated for the mercury co-crystal data collected on PX 9.6 using the selenomethionyl MAD phases. The Fourier showed a peak at 8.46 r.m.s. at the triple cysteine site, with no other significant peaks. It remains unclear why no Hg signal was detected in the XANES scan – high background noise level from the fluorescence detector is perhaps the most likely explanation, but it may be worth noting that difficulties in picking up the Hg signal are not uncommon (A. Thompson, personal communication). In addition, an anomalous difference Fourier map was calculated using data from a platinum tetrachloride co-crystal. The data had been obtained during another MAD experiment at ESRF (Grenoble). The platinum data which were used for this calculation had been collected at a wavelength of 1.074 Å. The data extend to a resolution of 2.8 Å and have a merging R factor of 14.6. As before, phases from the selenomethionyl MAD solution with the Zn atoms omitted were used. The difference Fourier confirms that platinum also binds at the triple cysteine site. A single peak was found at this position at a height of 6.94 r.m.s.

The lead and mercury complexes were refined and rebuilt for several rounds and the relevant statistics are given in Table 8. The remodelling during refinement of the lead complex was performed initially using the *O* suite (Jones *et al.*, 1991) and subsequently using *TURBO FRODO* (Roussel & Cambillau, 1989). For the mercury complex, *TURBO FRODO* was used throughout. Least-squares restrained refinement for both structures was performed using *RESTRAIN* (Haneef *et al.*, 1985) and a summary of the refinement and validation statistics for the two structures is given in Table 9.

4. Discussion

It appears from three-dimensional structural superposition of the Pb and Hg complexes with the native structure that neither metal ion has caused any substantial conformational changes in the protein (using a maximum distance cutoff of 2.0 Å, the r.m.s. deviation for 326 structurally equivalent C^{α} atoms was 0.4 Å for both structures). It thus appears that the inhibitory effect of both metals is a consequence of displacement of the catalytic Zn^{2+} rather than any induced conformational change.

The first step in the reaction catalysed by ALAD is the binding of the substrate ALA destined to become the propionate side of product porphobilinogen. This eponymously termed 'P-side' substrate molecule binds by means of a Schiff base link to an active-site lysine residue (Lys263 in yeast ALAD). The NH_2 group of this lysine is approximately 5 Å from the triple cysteine site from which zinc is displaced by both lead and mercury. Zinc is also lost on oxidation of the metal-binding thiol groups, which is accompanied by loss of activity (Spencer & Jordan, 1993; Tsukamoto *et al.*, 1979; Jaffe & Hanes, 1986). In contrast, plant dehydratases require magnesium ions and lack the zinc-binding cysteines. For example, in the pea, moss and spinach enzymes, two of the conserved cysteines and one histidine are replaced by aspartates. Correspondingly, the plant enzymes are less sensitive to oxidation. It has been shown in mammalian ALADs that

Table 9
Validation statistics for Pb- and Hg-complexed yeast ALAD at 2.5 and 3.0 Å, respectively.

	Pb	Hg
Estimated coordinate error from *SIGMAA*† (Å)	0.50	0.50
Main-chain torsion angles within 'most favoured regions' of the *PROCHECK*‡ Ramachandran plot	77.9	73.9
Main-chain torsion angles within 'additional allowed regions' of *PROCHECK* Ramachandran plot	20.4	23.6

† Read (1986). ‡ Laskowski *et al.* (1993).

formation of the Schiff-base complex between P-side substrate and the active-site lysine does not require Zn^{2+} or enzyme sulfhydryl groups (Jaffe & Hanes, 1986; Jordan *et al.*, 1976), suggesting that either the zinc ion interacts with the other (acetic acid side or 'A-side') substrate molecule or is important for the structure of the site to which that molecule binds.

It is very unusual to find an active-site zinc ion coordinated by three cysteine ligands, as zinc shows a preponderance of imidazole and carboxyl ligands at the active sites of other enzymes, such as the functionally related metalloaldolases. However, the active site of alcohol dehydrogenase is similar in that its zinc ion is held by two cysteines and a histidine. In this enzyme, the zinc is presumed to coordinate the carbonyl of ethanal (acetaldehyde) to facilitate its reduction. In ALAD, the active-site zinc ion may perform a similar task by coordinating the carbonyl of A-side substrate prior to its condensation with the P-side ALA. The zinc ion may function by polarizing this carbonyl group, thus making it more susceptible to nucleophilic attack by the P-side substrate. During catalysis, there is a requirement for a strongly basic group which deprotonates the A-side C-3 methylene. It is feasible that this could be a zinc-bound hydroxide ion. A solvent molecule datively bound to the zinc ion is the most likely candidate for a hydroxyl species in the vicinity of the substrate. The zinc ion could also coordinate and therefore stabilize the putative oxyanion intermediates in the reaction mechanism. Clearly, the potent inhibition of the enzyme by lead owing to its binding at the main zinc site reflects the fact that lead has inappropriate properties either for interacting with the substrate or for stabilizing the catalytic intermediates.

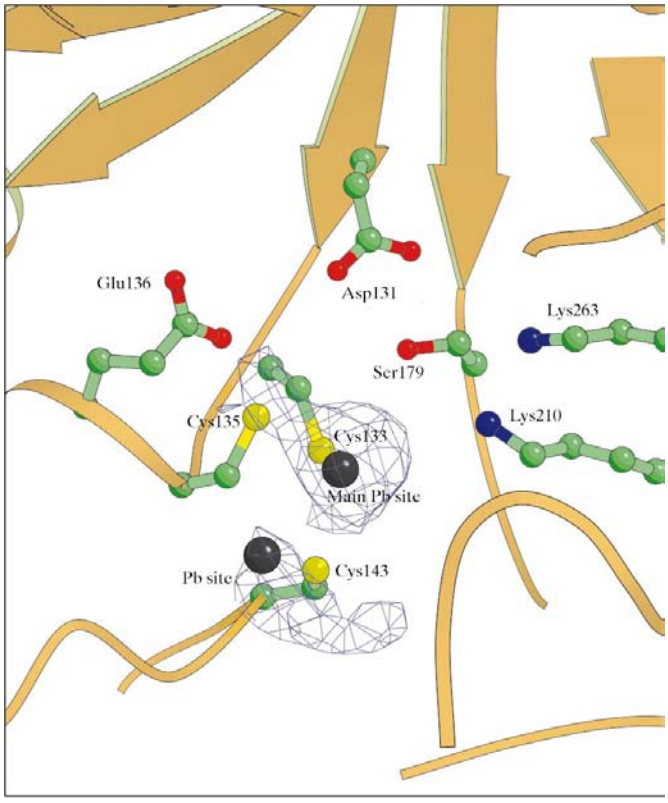

Figure 7
The refined positions of the Pb^{2+} sites in the model built using the 2.5 Å resolution PX 9.6 data are shown with selected active-site residues. At the right are the two lysines involved in catalysis. It is to Lys263 that the substrate molecule which forms the propionic acid side of product porphobilinogen binds, making a Schiff base. The 17.89 r.m.s. peak in the anomalous difference Fourier map at the main Pb site in the triple cysteine cluster is also shown. The Fourier is calculated from the 3.0 Å resolution λ_1 data of the Pb L_{III} edge MAD experiment conducted at PX 9.5 at the SRS. The map is contoured at 5 r.m.s. The subsidiary Pb site coinciding with the 6.99 r.m.s. peak in the Fourier can be seen below and to the left of the main Pb site. The two smaller lobes of density contiguous with each of the peaks at the refined Pb sites are the adjacent peaks in the Fourier of about 6.5 r.m.s. This figure and the following figure were prepared with *BOBSCRIPT* (Esnouf, 1997).

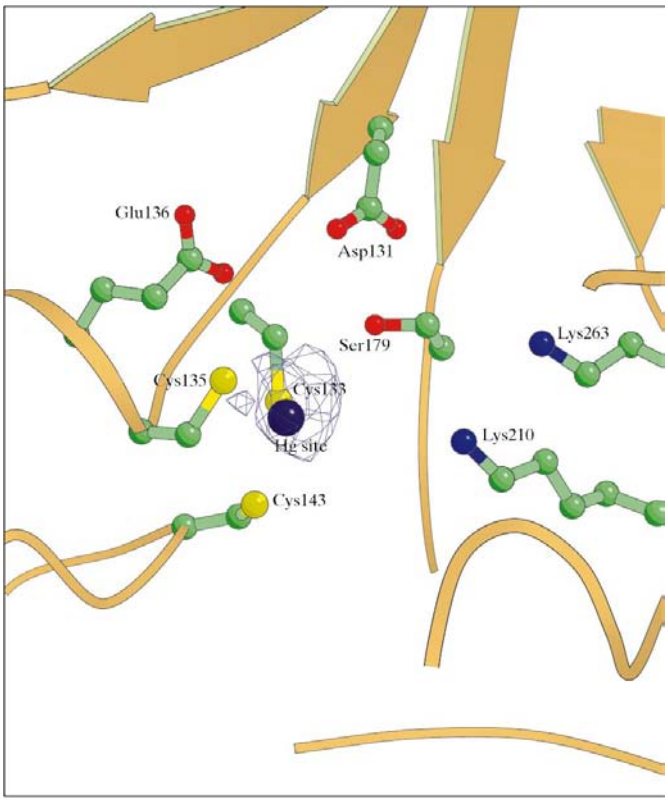

Figure 8
The refined position of the Hg^{2+} site found at the triple cysteine cluster using the 3.0 Å resolution data collected at PX 9.6 at the SRS. The electron density shown is from the anomalous difference Fourier calculated from this data. The peak height was 8.46 r.m.s. The map is contoured at 4 r.m.s.

It seems likely that the phasing exercise based on the Pb-edge MAD data did not produce an interpretable map because the dispersive differences were not sufficiently large. However, by means of the anomalous difference Fourier it has been possible to successfully locate the Pb sites, providing for the first time a graphic explanation of the cause for the effects of lead poisoning at this point on the tetrapyrrole biosynthetic pathway (Fig. 7). Similarly, the zinc-edge MAD data, whilst being insufficient to solve the structure of ALAD, have provided, by means of the difference Fouriers, valuable information on the Zn^{2+} sites, which greatly aided interpretation of the selenomethionyl MAD-derived map. Evaluation of the peaks other than the unambiguous main peak at the triple cysteine site is not straightforward. For example, the absence from the anomalous difference Fourier of the third peak in the vicinity of Cys285 found in the dispersive difference Fourier may indicate that both it and the second site, which is at the same height, may be at the level of noise. The cysteine near the second site (Cys234) occurs in a part of the electron-density map which so far has proved notably refractory to elucidation and final interpretation may require more data on inhibitor complexes. The stretch of primary structure adjacent to the putative metal-binding residues looks likely to constitute a flexible active-site lid and it is possible that these secondary sites may be less well occupied as a consequence of inherent dynamic movement.

The difficulty in locating the Hg and Pb sites in difference Pattersons calculated against the native data and the failure of the Hg and Pb derivatives to provide phases of sufficient quality to solve the structure could be attributable to non-isomorphism. However, the unit-cell parameters of the crystals described here are essentially the same – the Pb and Hg co-crystal cells were only 1.6% longer along *c* than the native collected at $\lambda = 1.13$ Å and only 1.5 (Hg) and 1.3% (Pb) smaller along *a*/*b* (see Table 1 for unit-cell parameters). It is interesting to rationalize the initially somewhat puzzling result from the isomorphous difference Patterson between the lead co-crystal and the Hg co-crystal being treated as native. Patterson vector-search methods identified an initial peak which when put into a superposition function identified an additional peak of higher occupancy. The initial site produced by the vector search was not the triple cysteine site but one of the adjacent peaks for which the full 82 electrons of the lead ion would potentially be contributing to the difference Patterson peak. At the triple cysteine site the occupancy is probably higher (judging from the refined structure), but because the putative native had in fact been substituted by mercury at this position, only two electrons were contributing to the difference Patterson peak or at best 52 electrons from that fraction of unsubstituted molecules retaining the zinc cofactor at this position. It is worth noting that whilst there is no peak in the Hg anomalous difference Fourier at the initial site (not the main site), there is some density in the Hg $2F_o - F_c$ map, possibly indicative of a low degree of substitution by mercury at this position. While the pattern of substitution may not be quite as clear-cut as suggested in the above rationalization, the entire analysis may not be vitiated.

The result of the anomalous difference Fourier for the mercury co-crystal was revelatory, not least for showing what a dubious choice the mercury ion is for a MAD experiment, but more importantly for showing that the main site of substitution in the enzyme is identical to that of the main site of substitution by lead (Fig. 8). This, together with our demonstration that platinum binds at the same site, provides a valuable insight into the causes of the neurological effects of these major heavy-metal poisons (Goyer, 1996).

Many helpful discussions with Dr I. Glover (Department of Physics, University of Keele) and Dr A. Thompson (ESRF, Grenoble) are gratefully acknowledged. The support of the EPSRC, BBSRC and the Wellcome Trust is gratefully acknowledged.

References

Anderson, P. M. & Desnick, R. J. (1979). *J. Biol. Chem.* **254**, 6924–6930.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cromer, D. T. & Liberman, D. (1970). *J. Chem. Phys.* **53**, 1891–1898.
Dent, A. J., Beyersmann, D., Block, C. & Hasnain, S. S. (1990). *Biochemistry*, **29**, 7822–7828.
Doss, M., von Tiepermann, R., Schneider, J. & Schmid, H. (1979). *Klin. Wochenschr.* **57**, 1123–1127.
Erskine, P. T., Norton, E., Cooper, J. B., Lambert, R., Coker, A., Lewis, G., Spencer, P., Sarwar, M., Wood, S. P., Warren, M. J., Schoolingin-Jordan, P. M. (1999). *Biochemistry*, **38**, 4266–4276.
Erskine, P. T., Senior, N., Awan, S., Lambert, R., Lewis, G., Tickle, I. J., Sarwar, M., Spencer, P., Thomas, P., Warren, M. J., Shoolingin-Jordan, P. M., Wood, S. P. & Cooper, J. B. (1997). *Nature Struct. Biol.* **4**, 1025–1031.
Erskine, P. T., Senior, N., Maignan, S., Cooper, J., Lambert, R., Lewis, G., Spencer, P., Awan, S., Warren, M., Tickle, I. J., Thomas, P., Wood, S. P. & Shoolingin-Jordan, P. M. (1997). *Protein Sci.* **6**, 1774–1776.
Esnouf, R. (1997). *J. Mol. Graph. Mod.* **15**, 132.
Gibbs, P. N. B. & Jordan, P. M. (1981). *Biochem. Soc. Trans.* **9**, 232–233.
Goyer, R. A. (1996). *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 5th ed., edited by C. D. Klassen, pp. 703–712. New York: McGraw–Hill.
Gross, M., Hessefort, S. & Olin, A. (1999). *J. Biol. Chem.* **274**, 3125–3134.
Guo, G. G., Gu, M. & Etlinger, J. D. (1994). *J. Biol. Chem.* **269**, 12399–12402.
Haneef, I., Moss, D. S., Stanford, M. J. & Borkakoti, N. (1985). *Acta Cryst.* A**41**, 426–433.
Hassoun, A., Verstraeten, L., Mercelis, R. & Martin, J.-J. (1989). *J. Clin. Chem. Clin. Biochem.* **27**, 781–786.
Jaffe, E. K., Abrams, W. R., Kaempfen, H. X. & Harris, K. A. Jr (1992). *Biochemistry*, **31**, 2113–2123.
Jaffe, E. K. & Hanes, D. (1986). *J. Biol. Chem.* **261**, 9348–9353.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Jordan, P. M., Gore, M. G. & Chaudhry, A. G. (1976). *Biochem. Soc. Trans.* **4**, 762–763.

Krishna Murthy, H. M. (1996). *Crystallographic Methods and Protocols. Methods in Molecular Biology*, Vol. 56, edited by L. Jones, B. Mulloy & M. Sanderson, pp. 127–151. New Jersey: Humana Press.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leslie, A. G. W. (1992). In *Jnt CCP4/ESF–EACMB Newslett. Protein Crystallogr.* Vol. 26.
Mercelis, R., Hassoun, A., Verstraeten, L., De Bock, R. & Martin, J.-J. (1990). *J. Neurol. Sci.* **95**, 39–47.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Roussel, A. & Cambillau, C. (1989). *Silicon Graphics Geometry Partner Directory*, pp. 77–88. Mountain View, California: Silicon Graphics.
Schroeder, T. M. & Caspers, M. L. (1996). *Biochem. Pharmacol.* **52**, 927–931.
Senior, N. M., Brocklehurst, K., Cooper, J. B., Wood, S. P., Erskine, P., Shoolingin-Jordan, P. M., Thomas, P. G. & Warren, M. J. (1996). *Biochem J.* **320**, 401–412.
Simons, T. J. B. (1995). *Eur. J. Biochem.* **234**, 178–183.
Spencer, P. & Jordan, P. M. (1993). *Biochem. J.* **290**, 279–287.
Thunell, S., Holmberg, L. & Lundgren, J. (1987). *J. Clin. Chem. Clin. Biochem.* **25**, 5–14.
Tsukamoto, I., Yoshinaga, T. & Sano, S. (1979). *Biochim. Biophys. Acta*, **570**, 167–178.
Warren, M. J., Cooper, J. B., Wood, S. P. & Shoolingin-Jordan, P. M. (1998). *Trends Biochem. Sci.* **23**, 217–221.