Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Making the most of two crystals: structural analysis of a conserved hypothetical protein using native gel screening and SAD phasing

J. Shaun Lott,[a]* Mark J. Banfield,[a,b] Jill Sigrell[a,c] and Edward N. Baker[a]

[a]Laboratory of Structural Biology, School of Biological Sciences, University of Auckland, Private Bag 92-019, Auckland, New Zealand, [b]Cell and Molecular Biosciences, Faculty of Medical Sciences, University of Newcastle, Newcastle-upon-Tyne NE2 4HH, England, and [c]Amersham Biosciences, Research & Development, Bjorkgatan 30, SE-751 84, Uppsala, Sweden

Correspondence e-mail: s.lott@auckland.ac.nz

Received 2 July 2003
Accepted 17 September 2003

The protein PAE2307 is a member of a protein family of unknown function which is conserved among a number of bacterial and archaeal species. The protein was overexpressed in *Escherichia coli*, purified and crystallized in two crystal forms. The prevalent form was twinned, but the other diffracted to 1.45 Å resolution. The non-twinned crystals proved difficult to reproduce, so screening of potential heavy-atom derivatives by native polyacrylamide gel electrophoresis was used to establish suitable derivatization conditions. This process enabled the production of a $K_2Pt(NO_2)_4$ derivative that was used to collect a single-wavelength anomalous diffraction (SAD) data set from the only available crystal. Phase information of high quality was obtained, enabling the calculation of an interpretable electron-density map.

1. Introduction

Many structural genomics initiatives are now under way worldwide and an important aspect of all such projects is the need for rapid determination of phases for new protein crystal structures. Most commonly, multiwavelength anomalous diffraction (MAD) phasing using selenomethionine (Hendrickson, 1991) is the method of choice, although halide soaking (Dauter & Dauter, 2001) and rapid heavy-atom soaks (Nagem *et al.*, 2001; Sun & Radaev, 2002; Sun *et al.*, 2002) have both been suggested as viable high-throughput alternatives. The so-called 'low-hanging fruit' approach to structural genomics, in which the apparently most tractable targets are pursued first, causes a dilemma of time investment when personnel are a limiting factor, as is the case in the majority of laboratories; highly expressed proteins may often be insoluble, soluble proteins may not readily crystallize, crystals may prove hard to reproduce and the expected phasing route may fail. Often, arbitrary decisions have to be made with respect to the best investment of effort. The dilemma becomes more acute further down the experimental path, as more effort has already been expended on a particular target, only for it to become apparent that the low-hanging fruit may in fact be growing on a higher branch than had been supposed.

As part of a pilot structural genomics exercise, we chose a set of proteins of unknown function and structure from the hyperthermophilic archaeon *Pyrobaculum aerophilum*, the complete genome sequence of which is now available (Fitz-Gibbon *et al.*, 2002). One of these proteins, PAE2307, is classified in the COGs database (Tatusov *et al.*, 2001) as an 'ancient conserved domain', being part of a family of uncharacterized proteins that is conserved across both the archaea and the bacteria. The proteins in this family are all approximately 160 amino acids in length and show strong

sequence conservation between distantly related organisms: 50% sequence identity over the length of the protein between PAE2307 and its presumed orthologue in *Pyrococcus horikoshii* and 55% sequence identity between PAE2307 and its presumed orthologue in *Mycobacterium tuberculosis*. Full details of the bioinformatic analysis of PAE2307 will be published elsewhere.

Purified recombinant PAE2307 rapidly and reproducibly yielded diffraction-quality crystals. However, the prevalent crystal form was almost perfectly merohedrally twinned and hence extremely difficult to phase *via* MAD or MIR. Crystallization experiments yielded very few, hard-to-reproduce non-twinned crystals of a second form, which diffracted to high resolution. In our hands, only two such crystals could be grown. We show that the careful use of native polyacrylamide gel electrophoresis (PAGE) to screen heavy-atom derivatization conditions in solution can maximize the likelihood of successfully obtaining a suitable heavy-atom derivative when only a limited number of suitable crystals are available. The derivative obtained using this screening procedure enabled the collection of a single-wavelength anomalous diffraction (SAD) data set using a laboratory X-ray source, which enabled the calculation of an interpretable electron-density map. This strategy has advantages for laboratory-based high-throughput initiatives; no synchrotron access is necessary for structure solution and phasing can be achieved with only a small number of crystals, removing the need for optimization of crystallization conditions after initial crystals have been obtained.

2. Materials and methods

2.1. Protein expression and purification

The open reading frame PAE2307 was amplified from *P. aerophilum* genomic DNA using the polymerase chain reaction with primers designed to introduce *Nde*I and *Bam*HI sites at the 5′ and 3′ ends, respectively, and cloned into the expression vector pET28, yielding a construct containing a 20-amino-acid N-terminal thrombin-cleavable His tag (MGSSHHHHHHSSGLVPRGSH) upstream of the predicted translational start site. The sequence of the coding region was confirmed by DNA sequencing. The construct was transformed into *Escherichia coli* BL21(DE3) cells and grown in LB medium in a shaking culture at 310 K in the presence of a suitable antibiotic selection. When OD_{600} = 0.6, expression was induced by the addition of IPTG to a final concentration of 1 m*M*. The cells were harvested after a further 3 h of growth, resuspended in lysis buffer (20 m*M* Na HEPES pH 8.0, 150 m*M* NaCl, 10 m*M* imidazole; 30 ml per litre of culture) and lysed using a French pressure cell. Insoluble material was removed from the lysate by centrifugation at 13 000*g*. SDS–PAGE analysis indicated that PAE2307 was highly expressed and was found entirely in the soluble fraction of the cell lysate.

The soluble lysate was loaded onto a Hi-Trap NTA column (Amersham Biosciences) which had previously been loaded with Ni^{2+} ions. The column was washed with 10 column volumes of lysis buffer before bound proteins were eluted with a linear gradient of 10–500 m*M* imidazole in lysis buffer, followed by a step wash of 1 *M* imidazole in lysis buffer. The eluted PAE2307 fraction was then concentrated and further purified by size-exclusion chromatography on a Superdex S75 column (Amersham Biosciences).

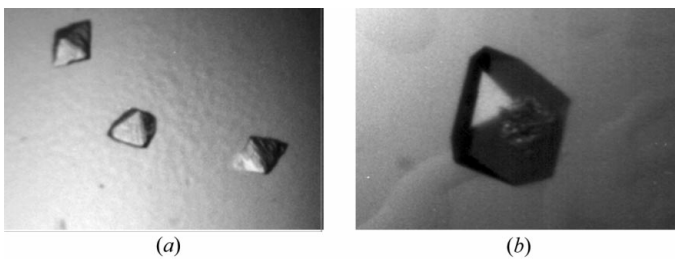
2.2. Crystallization

Crystallization trials were carried out using hanging-drop vapour diffusion at 291 K, with PAE2307 concentrated to 9 mg ml^{-1} in 20 m*M* Na HEPES pH 8.0, 150 m*M* NaCl. Initial experiments using sparse-matrix screens (2 µl drops; 1 µl protein plus 1 µl precipitant) immediately produced bipyramidal crystals, which appeared rapidly and grew over the course of around a week in two different conditions. Crystals of form I grew with 1.5 *M* $LiSO_4$, 0.1 *M* Na HEPES pH 7.5 as precipitant (Fig. 1*a*). Crystal form II grew as a single large crystal (0.25 × 0.25 × 0.2 mm) from an unbuffered solution of 50 m*M* KH_2PO_4, 20% PEG 8000 (Fig. 1*b*). Subsequent refinement of crystallization conditions produced many crystals of form I, which grew readily with 1.4–1.5 *M* $LiSO_4$, 0.1 *M* Na HEPES pH 7.0–7.5. However, further extensive screening around the conditions for crystal form II (varying pH, buffer type, protein concentration and precipitant concentration) yielded very few crystals, only one of which was suitable for subsequent X-ray analysis.

2.3. Data collection and processing

Prior to data collection, crystals of both forms were removed from the hanging drop with a nylon loop and were sequentially washed in mother liquor containing increasing concentrations of glycerol before being flash-cooled under a stream of cold nitrogen in mother liquor containing 20%(*v*/*v*) glycerol as cryoprotectant. X-ray diffraction data were collected at 110 K using a Rigaku rotating-anode generator (λ = 1.5418 Å) and a MAR 345 image-plate detector. Rotation images were collected with an exposure time of 1800 s and an oscillation angle of 1°. X-ray data were indexed, integrated and scaled using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

Crystals of form I routinely diffracted to a maximum resolution of approximately 3 Å on a laboratory source and to



(*a*) (*b*)

Figure 1
Two crystal forms of PAE2307. (*a*) Bipyramids grown from 1.4 *M* $LiSO_4$, 0.1 *M* Na HEPES pH 7.5. (*b*) Single crystal grown from unbuffered 50 m*M* KH_2PO_4, 20% PEG 8000.

Table 1
Data-collection statistics for PAE2307.

Values in parentheses are for the outermost shell.

	Form II		Form I
Data set	SSRL	SAD	SSRL
Space group	$I4_122$	$I4_122$	$I4_1$
Unit-cell parameters			
$a = b$ (Å)	120.04	120.72	120.24
c (Å)	156.47	154.88	155.95
Wavelength (Å)	0.9792	1.5418	0.9184
Resolution range (Å)	50–1.45 (1.50–1.45)	30–2.10 (2.18–2.10)	50–2.10 (2.21–2.10)
No. measured reflections	785 088	408 849	666 696
No. unique reflections	100 414	64 148	64 210
Redundancy	7.8	6.4	10.4
Data completeness (%)	100 (99.9)	100 (100)	99.8 (100)
Average $I/\sigma(I)$	43.0 (5.7)	29.8 (5.7)	14.9 (3.3)
$R_{\rm merge}$† (%)	4.2 (34.7)	6.2 (32.4)	8.4 (23.0)
$R_{\rm anom}$‡ (%)		2.3 (2.9)§	
Phasing statistics			
No. of sites		3	
$R_{\rm cullis}$ (anom)¶		0.95	
Phasing power¶		0.54	
FOM (*SOLVE*)		0.34	
FOM (*RESOLVE*)		0.65	

† $R_{\rm merge} = \sum|I - \langle I\rangle|/\sum I$. ‡ $R_{\rm anom} = \sum|F^+ - F^-|/\sum[(F^+ + F^-/2]$. § Resolution range 30–3.1 Å, outermost shell 3.26–3.10 Å. ¶ As defined in *CNS*.

almost 2 Å at the Stanford Synchrotron Radiation Laboratory (beamline 9-2; ADSC Quantum 4 detector). The reflections were indexed in an *I*-centred tetragonal space group and appeared to have the symmetry of space group $I4_1$, with unit-cell parameters $a = b = 120.24$, $c = 155.95$ Å and a mosaic spread of ~0.45°. The single large form II crystal grown from initial screening diffracted on a laboratory source to a Bragg resolution of better than 2 Å and displayed the symmetry of space group $I4_122$, with unit-cell parameters $a = b = 119.96$, $c = 156.37$ Å. A high-resolution data set was subsequently collected to 1.45 Å resolution from the same crystal at the Stanford Synchrotron Radiation Laboratory (beamline 9-2; ADSC Quantum 4 detector; $\lambda = 0.9792$ Å), which gave refined unit-cell parameters of $a = b = 120.04$, $c = 156.47$ Å and a mosaic spread of 0.35° (Table 1).

2.4. Heavy-atom screening

Conditions for heavy-atom derivatization were screened using native PAGE with a PHAST system (Amersham Biosciences), as has been suggested previously (Boggon & Shapiro, 2000). This method uses shifts in mobility on native PAGE to identify potential heavy-atom derivatives and to identify which heavy-atom reagents cause protein precipitation and may therefore lead to non-isomorphism or destroy diffraction in the crystalline state. It is an easy and rapid way to screen many potential heavy-atom derivatives using protein solutions rather than protein crystals. The theoretical pI of PAE2307 is 6.9, but analytical isoelectric focusing indicated the actual pI to be ≥9.0 (data not shown). Native PAGE was therefore carried out in the 'reverse' direction, *i.e.* migration towards the cathode, using a β-alanine/acetate buffering system at pH 4.1 (Olsson & Tooke, 1988). Initial screening for a heavy-atom derivative was carried out by incubating a solution of purified PAE2307 (9 mg ml^{-1}) with the following compounds, all at a nominal final concentration of 5 m*M*, for 6 h at 298 K: K_2PtCl_4, $KAuCl_4$, K_2HgI_4, uranyl acetate, $HgCl_2$, mercury acetate, trimethyllead acetate, CH_3HgPO_4 and K_2PtCl_6. Native PAGE showed that K_2PtCl_4 and $KAuCl_4$ caused the protein to aggregate or denature (Fig. 2*a*), whereas K_2PtCl_6 showed evidence of forming a stable derivative without causing the protein to precipitate. As both of the tested platinum compounds formed adducts, a further screen was carried out using a variety of platinum compounds of varying reactivity, namely *cis*-$Pt(NH_3)Cl_2$, platinum ethylenediamine dichloride, $K_2Pt(NO_2)_4$, $K_2Pt(CN)_4$ and di-μ-iodobis-(ethylenediamine) diplatinum nitrate (PIP) (Fig. 2*b*). Of the compounds tested, $K_2Pt(NO_2)_4$ gave the clearest evidence of producing a derivative, albeit with some protein precipitation, after a 6 h incubation at 298 K. To refine reaction conditions with this compound, a time-course was carried out at both 298 and 278 K (Fig. 2*c*). Although some derivatization could be detected in as little as 10 min at 298 K (as judged by faint bands of lower mobility being formed), 1–3 h were required for the reaction to approach completion. At 278 K, the reaction proceeded very slowly, with a 3 h incubation giving the same results as a 10 min incubation at 298 K.

3. SAD data collection and phase determination

The second diffraction-quality form II crystal (0.25 × 0.25 × 0.2 mm) was soaked in mother liquor containing 2 m*M* $K_2Pt(NO_2)_4$ for 1 h at 298 K before being mounted in a glass capillary. Initial data were collected at 298 K, again using a Rigaku rotating-anode generator ($\lambda = 1.5418$ Å) and a MAR 345 image-plate detector. Six rotation images were collected with an exposure time of 300 s and an oscillation angle of 1°. Indexing of these initial frames using *DENZO* indicated the space group to be $I4_122$ as expected, but with unit-cell parameters that were sufficiently altered to make the crystal non-isomorphous with the type II native data. However, platinum exhibits significant anomalous scattering at this wavelength ($f' = -4.5932$ and $f'' = 6.9264$) and merging of initial data with *SCALEPACK* indicated the presence of an anomalous signal. The anomalous signal in the data set was detected using a two-stage scaling routine in *SCALEPACK* (Otwinowski & Minor, 1997), where the data are first scaled with the ANOMALOUS flag turned on, thus treating Bijovet pairs separately. The data are then scaled a second time, with all parameters remaining constant and the ANOMALOUS flag turned off, thus combining Bijvoet pairs. The deviation of χ^2 (essentially the goodness-of-fit of the error model) from unity in the second pass is used to quickly assess the presence and extent of anomalous differences in the data.

Having established that the crystal had been derivatized, it was removed from the capillary and soaked for a further 4 h in 5 m*M* $K_2Pt(NO_2)_4$ before being sequentially washed in mother liquor containing increasing concentrations of glycerol, all in the presence of 5 m*M* $K_2Pt(NO_2)_4$. The deri-

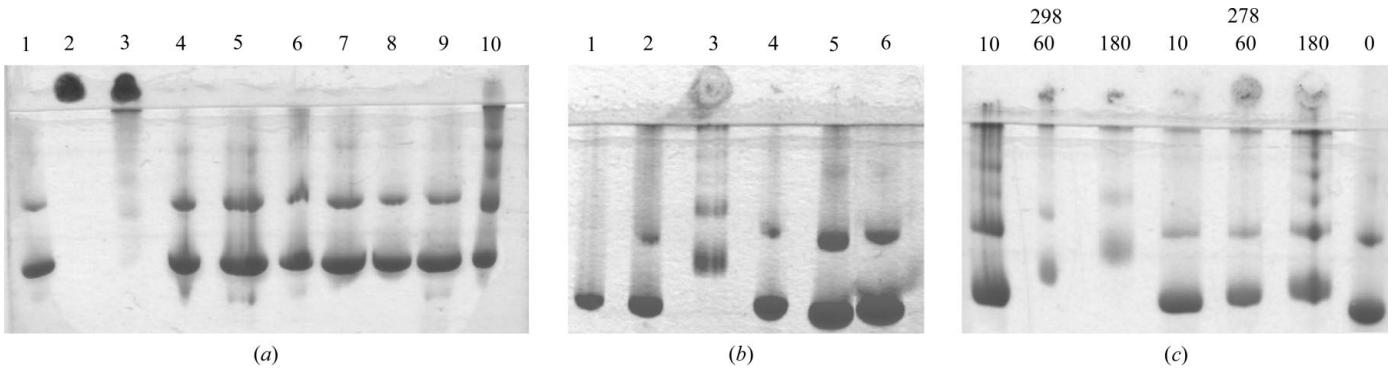


Figure 2
Native PAGE screening for a heavy-atom derivative. (*a*) Initial screening. Lane 1, PAE2307 only; lane 2, K_2PtCl_4; lane 3, $KAuCl_4$; lane 4, K_2HgI_4; lane 5, uranyl acetate; lane 6, $HgCl_2$; lane 7, mercury acetate; lane 8, trimethyllead acetate; lane 9, CH_3HgPO_4; lane 10, K_2PtCl_6. Lanes 2 and 3 show almost total protein precipitation and hence these compounds are unlikely to form crystallographically useful adducts. Lanes 4–9 show no clear evidence of derivatization. Lane 10 shows the appearance of bands of decreased mobility, indicating the production of a potentially useful heavy-atom derivative. (*b*) Screening with platinum compounds. Lane 1, *cis*-$Pt(NH_3)Cl_2$; lane 2, platinum ethylenediamine dichloride; lane 3, $K_2Pt(NO_2)_4$; lane 4, $K_2Pt(CN)_4$; lane 5, PIP; lane 6, PAE2307 only. The formation of a useful adduct by $K_2Pt(NO_2)_4$ is indicated by the change in mobility of the protein band in lane 3. (*c*) Time-course of derivatization with $K_2Pt(NO_2)_4$. Derivatization is more effective at 298 K than at 278 K and the reaction continues for at least 180 min. The upper figures are temperatures in K; the lower figures are times in min.

vatized crystal was flash-cooled as for the native type II crystal, in a final concentration of 20% glycerol, and a sixfold redundant data set was collected (Table 1). Inspection of χ^2 values in *SCALEPACK* as described above indicated that a detectable anomalous signal extended to at least 3.1 Å resolution, using $\chi^2 > 2$ as a cutoff. The positions of the Pt atoms were identified with the program *SOLVE* (Terwilliger & Berendzen, 1999) using all data to 3.1 Å. Although the anomalous signal in the data to 3.1 Å is small ($R_{\rm anom} = 2.3\%$), three clear heavy-atom sites were identified, with a *Z* score of 10.5 and an overall mean figure of merit (FOM) of 0.34. Maximum-likelihood density modification using *RESOLVE* (Terwilliger & Berendzen, 1999) increased the overall mean FOM to 0.65 and resulted in a clearly interpretable electron-density map. Initial manual model building based on helices and strands identified using the program *FFFEAR* (Collaborative Computational Project 4, 1994) enabled the identification of threefold non-crystallographic symmetry, which was used to further improve the electron density by map averaging using the program *DM* (Collaborative Computational Project 4, 1994). An example segment of the modified electron density is shown in Fig. 3.

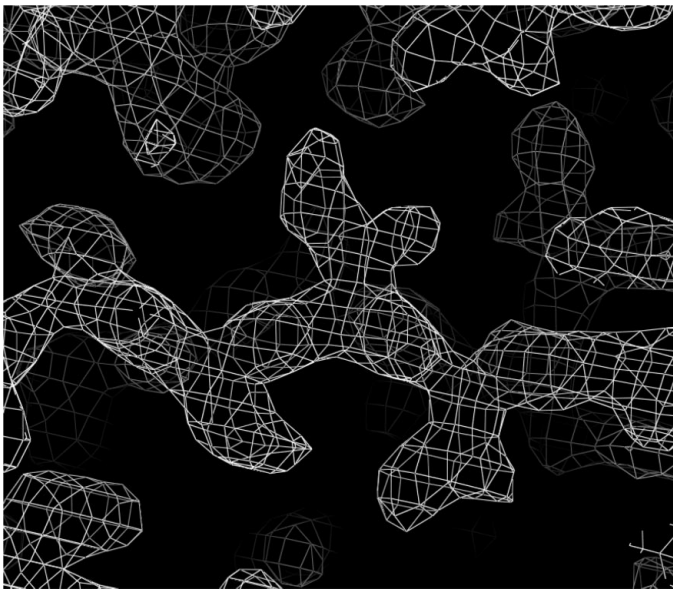


Figure 3
An example of experimentally phased electron density, contoured at 1.5σ, after maximum-likelihood density modification using *RESOLVE* and threefold non-crystallographic symmetry averaging using *DM*. Figure drawn using *PyMol* (DeLano, 2002).

4. Results and discussion

Purified recombinant PAE2307 crystallized in two forms. Crystals of form I grew readily and diffracted to a maximum resolution of 2.1 Å. Data-merging statistics indicated the most likely space group to be $I4_1$ (Table 1). However, inspection of the cumulative intensity distribution and acentric moment analysis using the program *TRUNCATE* (Collaborative Computational Project 4, 1994) indicated that the crystals were merohedrally twinned, with acentric moment values (Yeates, 1997) approaching that expected from a perfect twin (Table 2). Using the partial twin test (Yeates, 1997) as implemented in the program *DETWIN* (Collaborative Computational Project 4, 1994), the twinning fraction (α) was estimated to be 0.44 using all data to 2.1 Å. If only strong data were considered [data where $I/\sigma(I) \geq 3$ in resolution shells where overall $I/\sigma(I) > 5$] then this estimate increases to 0.46. Although cases have recently been reported of successful structure determination by MAD (Yang *et al.*, 2000; Rudolph *et al.*, 2003) or MIR (*e.g.* Terwisscha van Scheltinga *et al.*, 2001) using highly twinned data, in general the anomalous or isomorphous signal is sufficiently degraded by the errors introduced by detwinning so as to make structure solution impossible. This degradation increases as the twin fraction approaches 0.5, at which point the true diffraction intensities cannot be deconvoluted from the measured twinned intensities.

Table 2
Expected and observed moments of I and E for PAE2307 form I crystals.

	Expected values	PAE2307	
Moment	Untwinned	Perfect twin	Form I
$\langle E \rangle$	0.866	0.94	0.94
$\langle E^3 \rangle / \langle E \rangle^3$	1.339	1.175	1.185
$\langle I^2 \rangle / \langle I \rangle^2$	2.0	1.5	1.55
$\langle I^3 \rangle / \langle I \rangle^3$	6.0	3.0	3.31
$\langle I^4 \rangle / \langle I \rangle^4$	24.0	7.5	9.45

However, diffraction-quality crystals of PAE2307 also grew under a second condition. Only a few crystals could be grown in form II and only two were suitable for crystallographic analysis. The first of these form II crystals grew directly out of the initial sparse-matrix crystallization screen. This crystal was found to diffract to 2.1 Å on the laboratory source and displayed the symmetry of space group $I4_122$. Subsequent synchrotron data collection from this frozen crystal gave a complete native data set to 1.45 Å (Table 1). Acentric moment analysis using *TRUNCATE* (Collaborative Computational Project 4, 1994) indicated that this crystal was not twinned, thus leaving a single remaining non-twinned crystal suitable for X-ray analysis. In order to find a suitable heavy-atom derivative, native PAGE was used to screen a range of potential derivatizing agents in solution before treating the single crystal. The use of native PAGE allows compounds that form adducts that change the electrophoretic mobility of the protein to be identified. PAGE also allows those compounds and conditions that denature the protein (and hence presumably will disrupt or destroy order in the crystal form) to be identified and excluded from consideration. From a range of compounds used in the initial screen (Fig. 2*a*), reactive platinum species seemed the most promising and so were screened in more detail (Fig. 2*b*). The most promising of the platinum compounds was $K_2Pt(NO_2)_4$, which was subsequently used to establish a suitable time-course for derivatization (Fig. 2*c*).

Once the behaviour of the derivatizing compound had been established in solution, the single remaining diffraction-quality crystal was treated. Initially, a conservative protocol was adopted [a 1 h soak in 2 m*M* $K_2Pt(NO_2)_4$] based on the observations made in solution. A few images of diffraction data were collected at room temperature and this initial data indicated that the crystal had been derivatized but was not isomorphous with the native. The crystal was therefore soaked for longer in an attempt to ensure maximal occupancy of the platinum sites, again guided by the behaviour of the protein in solution. The advantage of this approach is that by taking initial measurements at room temperature the crystal remained available for further soaking, either in the same or in an alternative derivatizing reagent. However, once the presence of an anomalous signal was confirmed by analysis of the diffraction data, the crystal was frozen to ensure that a redundant data set could be collected for an accurate measurement of the anomalous differences.

Using this approach, we were able to use a single crystal to collect a sufficiently redundant data set to enable heavy-atom positions to be determined and good-quality phases to be calculated using *SOLVE*. Although the initial FOM was low, subsequent maximum-likelihood density modification using *RESOLVE* and non-crystallographic averaging using *DM* produced a readily interpretable electron-density map (Fig. 3). The extent of the anomalous signal and the FOM obtained are in the range of values reported in a recent review of SAD phasing techniques (Dauter *et al.*, 2002) and our experience reiterates the value of using the SAD method as a rapid technique for determining phase information that is easily accessible using a laboratory X-ray source. Careful testing of potential heavy-atom derivatives in solution, combined with a cautious data-collection strategy (utilizing both ambient and cryogenic conditions), can minimize the number of crystals required for structure determination, an outcome that has clear advantages for laboratory-based high-throughput projects where resources are limited.

This work was supported by a grant from the Marsden Fund of New Zealand to ENB. We thank Clyde Smith and Peter Haebel for synchrotron data collection. JB was supported by a Royal Society Travelling Fellowship

References

Boggon, T. J. & Shapiro, L. (2000). *Structure*, **8**, R143–R149.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z. & Dauter, M. (2001). *Structure*, **9**, R21–R26.
Dauter, Z., Dauter, M. & Dodson, E. (2002). *Acta Cryst.* D**58**, 494–506.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://www.pymol.org.
Fitz-Gibbon, S. T., Ladner, H., Kim, U. J., Stetter, K. O., Simon, M. I. & Miller, J. H. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 984–989.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.
Olsson, I. & Tooke, N. E. (1988). PhastSystem Application File No. 300: PAGE of Basic Proteins.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rudolph, M. G., Kelker, M. S., Schneider, T. R., Yeates, T. O., Oseroff, V., Heidary, D. K., Jennings, P. A. & Wilson, I. A. (2003). *Acta Cryst.* D**59**, 290–298.
Sun, P. D. & Radaev, S. (2002). *Acta Cryst.* D**58**, 1099–1103.
Sun, P. D., Radaev, S. & Kattah, M. (2002). *Acta Cryst.* D**58**, 1092–1098.
Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., Kiryutin, B., Galperin, M. Y., Fedorova, N. D. & Koonin, E. V. (2001). *Nucleic Acids Res.* **29**, 22–28.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Terwisscha van Scheltinga, A. C., Valegard, K., Ramaswamy, S., Hajdu, J. & Andersson, I. (2001). *Acta Cryst.* D**57**, 1776–1785.
Yang, F., Dauter, Z. & Wlodawer, A. (2000). *Acta Cryst.* D**56**, 959–964.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.