Acta Cryst. (1998). D**54**, 629–635

Structure Determination of a 16.8 kDa Copper Protein at 2.1 Å Resolution Using Anomalous Scattering Data with Direct Methods

Ian Harvey,[a,b] Quan Hao,[a] Elizabeth M. H. Duke,[b] W. John Ingledew[c] and S. Samar Hasnain[a,b]*

[a]*School of Applied Sciences, De Montfort University, The Gateway, Leicester LE1 9BH, England,* [b]*CLRC Daresbury Laboratory, Warrington WA4 4AD, England, and* [c]*School of Biomedical Sciences, St Andrews University, St Andrews, Fife KY16 9TS, Scotland. E-mail: s.hasnain@dl.ac.uk*

(*Received 13 February 1998; accepted 9 April 1998*)

Abstract

The structure of rusticyanin, an acid-stable copper protein, has been determined at 2.1 Å resolution by direct methods combined with the single-wavelength anomalous scattering (SAS) of copper ($f'' = 3.9$ e$^-$) and then conventionally refined ($R_{\rm cryst}$ = 18.7%, $R_{\rm free}$ = 21.9%). This is the largest unknown protein structure ($M_r \simeq 16.8$ kDa) to be determined using the SAS and direct-methods approach and demonstrates that by exploiting the anomalous signal at a single wavelength, direct methods can be used to determine phases at typical (∼2 Å) macromolecular crystallographic resolutions. Extrapolating from the size of the anomalous signal for copper ($f'' \simeq 4$ e$^-$), this result suggests that the approach could be used for proteins with molecular weights of up to 33 kDa per Se ($f''_{\rm max} = 8$ e$^-$ at the 'white line') and 80 kDa for a Pt derivative ($f''_{\rm max} = 19$ e$^-$ at the 'white line', L_3 edge). The method provides a powerful alternative in solving a *de novo* protein structure without either preparing multiple crystals (*i.e.* isomorphous heavy-atom derivative plus native crystals) or collecting multi-wavelength anomalous diffraction (MAD) data.

1. Introduction

Recently there has been tremendous interest in the use of direct methods for phase determination for macromolecules. This surge of interest has primarily resulted from two factors: the ability to obtain atomic resolution (1.2 Å) data in favourable cases and the development of powerful phasing methods including traditional direct methods (*Shake & Bake*), so-called *Half-baked*, and combinations of direct methods with isomorphous replacement and/or anomalous scattering (Hauptman, 1997). The ultimate potential of traditional direct methods is still unclear but one limit appears to be the requirement for atomic resolution data (Hauptman, 1997). Following our successful phasing of the 13 kDa copper protein azurin II using only the single wavelength anomalous scattering (SAS) data collected at a non-optimal wavelength ($f'' = 2.2$ e$^-$) (Zheng *et al.*, 1996), we proposed that with careful wavelength optimization it should be possible to phase a copper protein of up to 20 kDa using only the SAS data. Here we demonstrate the successful phasing and refinement of the copper protein rusticyanin (16.8 kDa) from *Thiobacillus ferrooxidans* using only SAS data collected at a wavelength close to, but not at the $f''_{\rm max}$ ('white line').

Rusticyanin is one of the largest of the type-1 blue copper proteins and is thought to be a principal component in the iron respiratory electron-transport chain of the acidophile *T. ferrooxidans*. The protein possesses extreme acid stability and the highest known redox potential of the single blue-copper proteins (680 mV): the average redox potential of the family is ∼300 mV. Homology with the C terminal of the other single copper-containing blue proteins (Ronk *et al.*, 1991; Yano *et al.*, 1991; Nunzi *et al.*, 1993; Grossmann *et al.*, 1995) and an interesting homology to parts of the blue-copper domain of the multi-copper proteins such as nitrite reductase have been noted (Grossmann *et al.*, 1995) and several models have been proposed to explain these unusual properties (Ronk *et al.*, 1991; Nunzi *et al.*, 1993; Grossmann *et al.*, 1995). These models suggested two possible candidates for the fourth ('N-terminal') copper ligand: Asp73 and His85. Mutagenesis (Casimiro *et al.*, 1995), together with this and other structural studies† have confirmed that the fourth copper ligand is His85.

The development of our approach which relies on SAS data and direct methods has a direct bearing on other efforts for *de novo* structure determination including the multi-wavelength anomalous dispersion (MAD) method (Hendrickson, 1991; Fourme & Hendrickson, 1990). In fact, the first unknown structure to be solved by the MAD technique in 1988 (Guss *et al.*, 1988) was also a type-1 copper protein, the considerably smaller (10 kDa) basic cucumber protein. MAD experiments generally rely on the use of at least three

† During the course of this work a MAD structure determination of oxidized rusticyanin as well as an NMR structure of reduced rusticyanin have appeared (Walter *et al.*, 1996; Botuyan *et al.*, 1996). The structure reported here was determined independently prior to the release of coordinates.

ISSN 0907-4449

wavelengths (four were used in the case of cucumber basic protein) to reach a least-squares determination of the phase angle, which puts increased demand on scarce synchrotron beam time. Moreover, the stringent requirement of selecting wavelengths optimized for f'_{min} and f''_{max} requires high-definition fluorescence XANES data from a single-crystal. Obtaining these data is experimentally demanding and is feasible only on certain specialized MAD beamlines, *e.g.* Brammer *et al.* (1988), Staudenmann *et al.* (1989). Furthermore, care is also required for tracking wavelength with crystal orientation if there is a significant dichroism in the XANES data as the position of f'_{min} and f''_{max} may shift with the crystal orientation. Therefore, the development of a method that is capable of resolving the phase ambiguity arising from the single-wavelength anomalous scattering (SAS) technique is particularly significant, especially since many native proteins already contain anomalous scatterers such as iron, copper, zinc or molybdenum. In this respect, it is worth noting that other strategies for MAD experiments in terms of minimizing the number of wavelengths (especially involving two wavelengths) have been recently explored for a brominated oligonucleotide (Peterson *et al.*, 1996).

There are successful procedures to break the phase ambiguity of SAS data for solving unknown protein structures. Ramachandran & Raman (1956) proposed that for the two possible phases of each reflection one can always make that choice which has a phase closer to that of the heavy-atom contribution. Hendrickson & Teeter (1981) used a similar but improved method in the structure determination of the small hydrophobic protein crambin. Their method combined the bimodal SAS phase distribution with the Sim weight calculated from the known positions of anomalous scatterers. Apart from the above, procedures based on P_s-function-related techniques (Hao & Woolfson, 1989; Fan *et al.*, 1990) and Wilson statistics (Ralph & Woolfson, 1991) have also been proposed to break the SAS phase ambiguity. Tests with data from known proteins showed that these methods are of use in practice. In a different context, direct methods have continuously been trying to break the SAS phase ambiguity (Fan & Gu, 1985; Hauptman, 1982). A common feature of this kind of method is to use three-phase structure invariants. So far the procedure of Fan & Gu (1985) is the only direct-methods procedure which has been successfully tested with experimental SAS data (Fan, Hao, Gu *et al.*, 1990; Zheng *et al.*, 1996). However, none of the above methods have been used to solve unknown protein structures except in the case of the 4 kDa protein crambin using data to 1.5 Å (Hendrickson & Teeter, 1981). In that case the high resolution of the data was essential for successful phasing.

In the SAS study of azurin II (Zheng *et al.*, 1996), the importance of data quality was assessed by attempting the direct-methods structure determination using several data sets varying in factors such as multiplicity, completeness and resolution. A minimum multiplicity of six was found to be essential for a successful solution of azurin II. Therefore, a high multiplicity SAS data set was collected for rusticyanin.

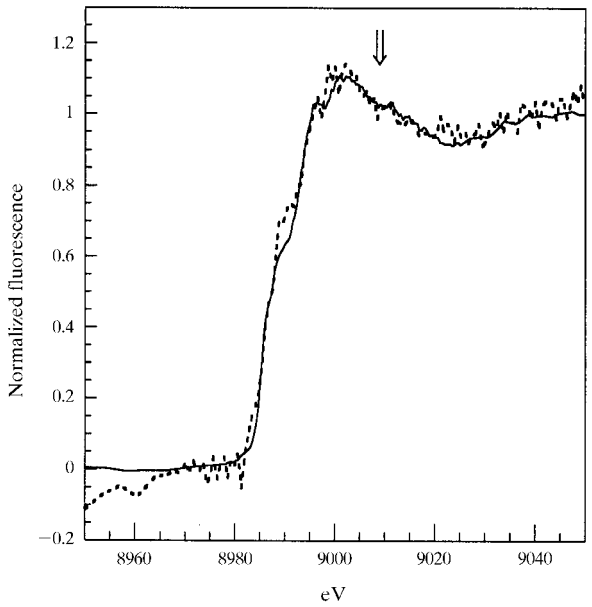

Fig. 1. XANES data for reduced rusticyanin in solution (solid line) and single crystal (dashed line). The two spectra are essentially the same and clearly show that protein is reduced and that the oxidation state of Cu is one. The poorer quality of the single-crystal data in comparison to the solution data results from a number of factors, including the nature and geometry of the detector available for the single-crystal data collection. The wavelength used for diffraction data collection (1.376 Å, 9008 eV) is shown with an arrow.

2. Structure determination

2.1. *Data collection and reduction*

Rusticyanin was crystallized in the reduced form (Cu^+) in space group $P2_1$ using conditions similar to those described previously (Djebli *et al.*, 1992). Anomalous scattering diffraction data extending to 2.1 Å were collected at room temperature on station 9.5 at the SRS, Daresbury Laboratory. The crystals diffract to higher resolution but the maximum resolution attainable was limited by a combination of the incident wavelength, detector radius and crystal-to-detector distance. Following a fluorescence XANES scan of a crystal, Fig. 1, the single wavelength (λ = 1.376 Å, f''_{Cu} = 3.88) was chosen to be close to the absorption edge, but approximately 6 eV above the peak. This was carried out to avoid dichroic effects of the crystal which are generally most pronounced at the white line. The absorption edge and features associated with it further confirm that the crystal was that of reduced protein. The crystal was

mounted with the spindle aligned along its *b* axis so that Friedel pairs were recorded on the same image plate. 3° oscillation images were collected, and the whole data set was obtained from one crystal. The 80 752 measurements (full reflections only) were reduced using *DENZO* (Otwinowski & Minor, 1997), then scaled and merged using the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) giving a merging $R_{\rm merge}$ of 6.6% ($R_{\rm ano}$ 2.9%) for 7898 independent Friedel pairs (or unique reflections for centric ones). The data are 95.5% complete to 2.1 Å (see Table 1).

2.2. *Locating the copper site*

The copper anomalous scatterer was located by using magnitudes of the anomalous differences

$$|\Delta\mathbf{F}(\mathbf{H})| = \big||\mathbf{F}(\mathbf{H})| - |\mathbf{F}(-\mathbf{H})|\big|$$

by the conventional direct methods program *SAPI*91 (Fan *et al.*, 1991) using data to 3.0 Å. The highest peak in the resultant sharpened (E) electron-density map was three times higher than the next peak. The highest peak was chosen as the copper site (fractional coordinates x = 0.1474, y = 0.2500, z = 0.4228). An anomalous Patterson map agreed with this solution (peak height 22.2 x's r.m.s. using all data).

2.3. *Evaluation of phase doublets*

The phase doublets inherent in the SAS method are expressed as,

$$\varphi_H = \varphi'_H \pm |\Delta\varphi_H|, \tag{1}$$

where φ'_H is the phase of

$$\mathbf{F}''_{\rm ano} = \sum_{j=1}^{N} if''_j \exp(i2\pi\mathbf{H}\cdot\mathbf{r}_j), \tag{2}$$

which can be calculated from the known positions of the anomalous scatterers and the known value of f''; $|\Delta\varphi_H|$ is obtained from

$$\cos\Delta\varphi_H = (F_H^+ - F_H^-)/2|F''_{\rm ano}|, \tag{3}$$

[see *e.g.* Blundell & Johnson (1976), and references therein]. The phase problem in the SAS case is in fact a sign problem according to (1). The probability for $\Delta\varphi_H$ is given by Fan & Gu (1985),

$$\begin{aligned} P_+(\Delta\varphi_H) = \tfrac{1}{2} + \tfrac{1}{2}\tanh\{\sin|\Delta\varphi_H|[\sum_{H'} m_{H'}m_{H-H'}\kappa_{H,H'} \\ \times \sin(\Phi'_3 + \Delta\varphi_{H',\rm best} + \Delta\varphi_{H-H',\rm best}) \\ + \chi\sin\delta_H]\}. \end{aligned} \tag{4}$$

Although anomalous differences for centric reflections are zero, *i.e.* $\Delta\varphi_H = 90°$, their phases can be evaluated using (4) the same way as for acentric reflections. This space group has a relatively small fraction of centric reflections (~5%), and the direct experimental information from SAS only involves the acentrics. Nevertheless, the the centric estimates benefit from the acentric estimates, and, therefore, where a larger fraction of centrics occur in other space-group cases, this should not be a limitation of the method demonstrated here with this space group as the example (see Table 2). The procedure to use (4) for *ab initio* phasing of the SAS data of rusticyanin is the same as that described by Fan *et al.* (1990) using the program *OASIS* (Hao *et al.*, 1996). All 7898 Friedel pairs (including centric reflections) to 2.1 Å were used as input. The noise level of the electron-density map was reduced by a minimum function of the direct-method map and a P_s-function map (Hao & Woolfson, 1989).

Table 1. *Data collection and refinement*

Values in parentheses refer to the highest resolution shell (2.15–2.10 Å)

Space group	$P2_1$
Unit-cell dimensions (Å, °)	a = 38.43
	b = 60.68
	c = 32.01
	γ = 107.82
Source	SRS beamline 9.5 (λ = 1.376 Å)
Resolution (Å)	50–2.1
No. of observed reflections	80752 (full only)
No. of unique reflections	7898 (417 centric, 7481 acentric)
Completeness (%)	95.5 (91.9)
Multiplicity	10.2 (7.7)
$R_{\rm sym}$† (%)	6.6 (10.6)
$R_{\rm ano}$ (%)	2.9 (4.6)
$I > 3\sigma$ (%)	95.9 (89.8)
Refinement	
Data range	8.0–2.1 Å
Reflections used ($R_{\rm free}$ set)	7044 (612)
Protein non-H atoms	1161
Solvent molecules	57
R.m.s. Δ bond lengths (Å)‡	0.011
R.m.s. Δ bond angles (°)‡	1.88
$R_{\rm free}$ (%)§	21.9
$R_{\rm cryst}$ (%)¶	18.7

† $R(I) = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| \sum_{hkl}\sum_i I_i(hkl)$. ‡ Root-mean-square deviation (R.m.s. Δ) are with respect to Engh and Huber ideal values. § $R_{\rm free}$ is the same as $R_{\rm cryst}$, but calculated on the 8% of data excluded from refinement ¶ $R_{\rm cryst} = 100\sum\big||F_o| - |F_c|\big|/\sum F_o$ for all reflections.

The phases obtained from the above procedure were not sufficiently accurate to produce a readily interpretable Fourier map. The density-modification technique, *DM* (Cowtan, 1994) was used to improve the phases, using an estimated solvent content of 38%. The electron-density map calculated from the *DM*-derived phases was then examined in the modelling suite *O* (Jones *et al.*, 1991). Use of the *bones* skeletonization

2.4. *Phase improvement, structure determination and refinement*

Table 2. *Phase errors at different stages*

Reflections were sorted in descending order of F_{obs} and then cumulated into seven groups with number of reflections shown in the first column. Phase errors were calculated against the refined structure model phases. Weighted average phase errors were calculated using F_{obs} as weights. I, phase errors of direct-methods results. II, phase errors of minimum function of direct methods and P_s function. III, phase errors after density modification. Centric reflections (417 in total) were included throughout the calculations. The separate statistics for these reflections were: at stage I, the mean F_{obs}-weighted phase errors were 79.4 and 62.9, respectively; stage II, 76.4 and 67.4; stage III, 65.0 and 46.6. Although the number of theses reflections is small (5% of the total) these were kept in the data set to avoid systematic bias.

		Phase errors					
		I		II		III	
Number of reflections	F_{obs} range	Mean	Weighted	Mean	Weighted	Mean	Weighted
1000	226.4–774	49.5	49.5	51.3	51.4	36.4	35.5
2000	157.5–774	51.2	50.8	52.6	52.3	40.1	38.6
3000	122.7–774	51.6	51.1	52.5	52.3	42.1	40.2
4000	97.1–774	53.0	52.0	54.4	53.4	44.5	41.8
5000	76.4–774	55.1	53.2	56.2	54.4	47.5	43.7
6000	57.5–774	56.4	53.9	57.4	55.1	49.7	44.9
7000	36.9–774	59.1	54.9	59.7	56.0	53.2	46.3
7998	5.9–774	61.9	55.5	62.0	56.5	56.4	47.1

routine resulted in an unambiguous tracing of the β-sheet core of the protein. Some surface residues could only be tentatively located at this stage. However, subsequent refinement confirmed the initial backbone trace as essentially correct. The complete backbone was built, based on the *bones* skeleton. The initial model, which consisted of one Cu and residues 1–155 (poly-alanine except the four Cu ligands: His85, Cys138, His143, Met148) had an R_{cryst} of 55.9% (R_{free} = 55.4%). The model refinement was then carried out with *X-PLOR* (Brünger *et al.*, 1987) using the Engh & Huber (1991) parameter libraries. A single cycle of simulated annealing was performed using 8–3 Å data over a temperature range 4000–300 K with a global *B* factor of 15 Å^2. The copper site was only weakly restrained during simulated annealing with energies of 335 kJ mol^{-1} (80 kcal mol^{-1}) for ligands His85, Cys138, His143 and 125.5 kJ mol^{-1} (30 kcal mol^{-1}) for Met148. During the initial cycle of refinement the Cu position was fixed and the R_{cryst} dropped from 55.9 to 35.0% (R_{free} 55.4 to 41.0%). The Cu position was then allowed to refine and the data was extended to 2.3 Å in two steps (R_{cryst} = 39.4%, R_{free} = 44.8%). At this stage side chains were clearly visible and were added over two further cycles of simulated annealing (R_{cryst} = 24.6%, R_{free} = 28.1%). The data was then extended to 2.1 Å with a further cycle of simulated annealing followed by restrained individual *B*-factor refinement (R_{cryst} =

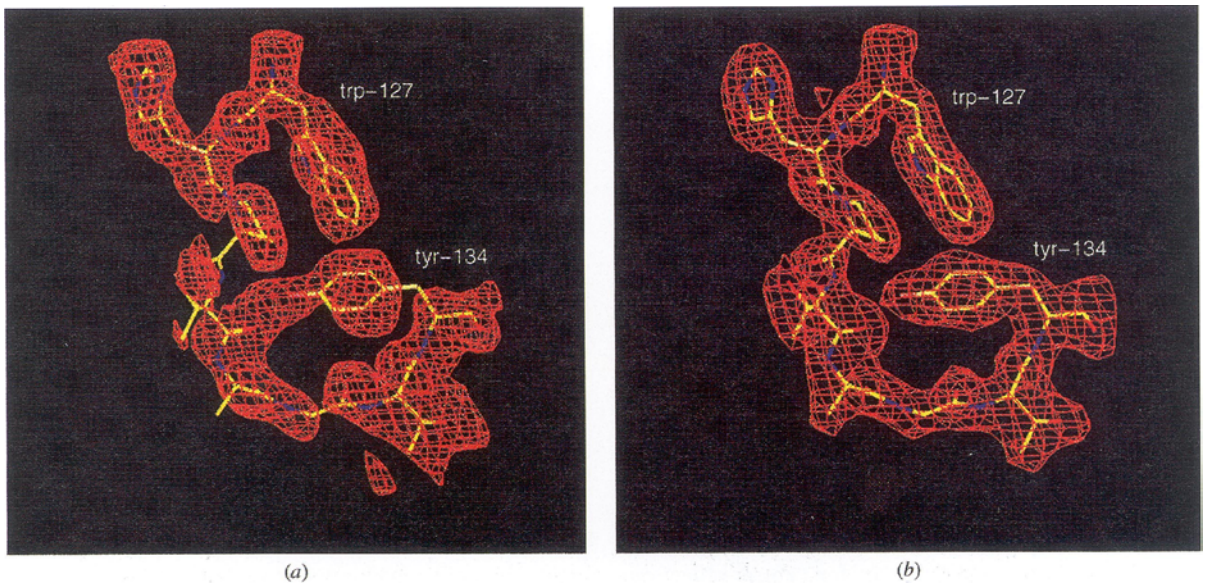

Fig. 2. The initial (*a*) and final (*b*) $2F_o - F_c$ electron density for the surface tyrosine loop (residues 127–134, final coordinates shown), contoured at 1 r.m.s.. Although the initial density for Thr130 is poor, there is clear initial density for most side chains.

22.4%, R_{free} = 26.2%). At this stage it was clear that the sequence being used was incorrect for five residues. Comparison of the sequence suggested by the electron density with the various known sequences showed that we had crystals with a known, but unpublished sequence (Ambler, personal communication). Four further cycles of positional and restrained individual *B*-factor refinement, concurrent with deletion of residues 1–2, the corrections of the sequence just referred to, and the introduction of 57 water molecules gave an R_{cryst} of 18.7% (R_{free} = 21.9%), See Table 1. The entire data set, including the 8% of reflections which were set aside for calculating R_{free}, was then used for a final cycle of positional refinement, followed by restrained individual *B*-factor refinement, which yielded a final R_{cryst} of 18.7%. The final model consists of residues 3–155, one copper and 57 waters. Analysis with *PROCHECK* (Laskowski *et al.*, 1993) shows good stereochemistry, with 89.3% of residues within the Ramachandran most favoured region and all other residues within the allowed region.

3. Results and discussion

The initial and final (from refined structure) electron density for the surface tyrosine loop (residues 127–134) is shown in Fig. 2. Although the initial density for Thr130 is poor, there is clear density for most side chains in this map. However, we erred on the side of caution by omitting side chains from the early cycles of refinement. This was one of the poorer sections of the map: density for the β-sheet core of the protein was significantly better. Fig. 3 shows the initial and final electron density at the copper site superimposed. Again, the quality of the initial map is good. The structure indicates that the extreme acid stability and high redox potential of this protein must, in part, arise from the extensive hydrophobic and hydrogen-bonding interactions. Further details of the structure will be presented when structure refinement with higher resolution data collected at a shorter wavelength (0.9 Å) is complete.

Table 2 shows the phase errors of different phasing stages for all 7898 reflections calculated using the final refined coordinates. These phase errors are typically 8° lower than the azurin II results (Zheng *et al.*, 1996) which reflects the stronger anomalous scattering signal at the Cu *K* edge due to higher resolution of the diffraction from rusticyanin crystals. There is also a significant correlation between the phase errors and F_{obs}, *i.e.* the stronger F_{obs} have lower errors.

The present case demonstrates that the direct method (4) is capable of discriminating the correct phase in a bimodal phase distribution of a protein reflection by exploiting single-wavelength anomalous scattering (SAS) diffraction data which extends to modest resolution (~2 Å). It is clear that despite a very weak anomalous scatterer like Cu ($f'' \simeq 4\ e^-$), phase determination is possible for a ~17 kDa protein using this approach. The XANES of rusticyanin (Fig. 1) shows no significant white line. This made data collection extremely tolerant in the choice of wavelength: the wavelength was selected such that it is 6 eV above the absorption edge. In our earlier study of azurin II (Zheng *et al.*, 1996), the data was collected at 1 Å, more than 3 keV above the absorption edge. This suggests that in practice it may not even be necessary to collect a flourescence XANES scan from a single-crystal prior to diffraction data collection. We note that in one of the earliest applications, two Mn ions were accurately located in pea lectin (a dimer of 50 kDa in the asymmetric unit, with one Mn per monomer) using f'' without obtaining fluorescence XANES data (Einspahr *et al.*, 1985). This tolerance in wavelength selection frees up the need for beam time to be committed to take a fluorescence scan at a regular interval during data collection. We can estimate the size of proteins for which this approach could be used to resolve the phase ambiguity simply by scaling to the magnitude of f'' of the scatterer if the quality of data can be rigorously main-

Fig. 3. Stereoview of the initial (red) and final (grey) $2F_o - F_c$ electron density at the copper site superimposed (both contoured at 1 r.m.s.).

tained. Thus, from (2), we estimate that without a wavelength optimization, the phase ambiguity can be broken for proteins with molecular weight of up to 17 kDa per Se using its K edge ($f'' \simeq 4\ e^-$), and 44 and 52 kDa respectively for single-site Pt and Hg derivatives, using their L_3 edges ($f'' \simeq 9\ e^-$). Although for these scatterers, this approach still requires soaking or co-crystallization of an anomalous scatterer into protein crystals, it retains the advantage of the MAD method of maintaining perfect isomorphism (*i.e.* Δ_{ano} is derived from the 'derivative' crystal alone). Since many heavy scatterers of interest have significant white lines, wavelength optimization will nevertheless significantly increase the f'' component of the data. In such cases, it should be possible to break the phase ambiguity for proteins with molecular weight of up to 33 kDa per Se using its K edge ($f''_{max} = 8\ e^-$) (Hendrickson *et al.*, 1990) and 80 kDa for a Pt derivative ($f''_{max} = 19\ e^-$) (Helliwell, 1992) and 100 kDa in the case of a single Eu site ($f''_{max} \simeq 23\ e^-$) (Hendrickson, 1991), using their L_3 edge white lines (which are usually larger than L_1 white lines). In these cases beam time for a fluorescence scan would need to be allowed for. The measurement of a single-wavelength diffraction data set, rather than two or more wavelengths would still result in a significant saving of beam time. This is especially so where the area detector read-out time is long with respect to the exposure time per image.

In comparison with the MAD method, this approach offers a considerable time saving in data collection and in some cases much simpler experiment (*i.e.* no on-line fluorescence scan is needed).

In summary, the SAS with direct methods provides a powerful alternative in solving a *de novo* protein structure without either preparing isomorphous heavy-atom derivative crystals or collecting multi-wavelength anomalous diffraction (MAD) data as demonstrated here.

We would like to thank the CLRC for provision of the facilities at Daresbury Laboratory. We would like to thank members of the group for their help and interest throughout the project. We are thankful to Professor Richard Ambler for making available unpublished sequence data. We are also grateful to Professors Kenneth Barker and Roger Linford for their support of the project. Professor Haifu Fan is thanked for useful discussions. Coordinates and structure factors have been deposited with the Protein Data Bank.†

† Atomic coordinates and structure factors have been deposited with the Protein Data Bank at Brookhaven (Reference: 1A8Z and R1A8ZSF). Free copies may be obtained through the Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference HE0228).

References

Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*, p. 338. New York: Academic Press.

Botuyan, M. V., Toypalmer, A., Chung, J., Blake, R. C., Beroza, P., Case, D. A. & Dyson, H. J. (1996). *J. Mol. Biol.* **263**, 752–767.

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

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Casimiro, D. R., Toypalmer, A., Blake, R. C. & Dyson, H. J. (1995). *Biochemistry*, **34**, 6640–6648.

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

Cowtan, K. (1994). *Jnt CCP4 ESF-EACBM Newslett. Protein Crystallogr.* **31**, 319–324.

Djebli, A., Proctor, P., Blake, R. C. & Shoham, M. (1992). *J. Mol. Biol.* **227**, 581–582.

Einspahr, H., Suguna, K., Suddath, F. L., Ellis, G., Helliwell, J. R. & Papiz, M. (1985). *Acta Cryst.* B**41**, 336–341.

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

Fan, H. F. & Gu Y. X. (1985). *Acta Cryst.* A**41**, 280–284.

Fan, H. F., Hao, Q., Gu, Y. X., Qian, J. Z., Zheng, C. D. & Ke, H. (1990). *Acta Cryst.* A**46**, 935–939.

Fan, H. F., Hao, Q. & Woolfson, M. M. (1990). *Acta Cryst.* A**46**, 659–664.

Fan, H. F., Yao, J. X., Zheng, C. D., Gu, Y. X. & Qian, J. Z. (1991). *SAPI*91, *a computer program for automatic solution of crystal structures from X-ray diffraction data*, Institute of Physics, Chinese Academy of Sciences, Beijing 100080, China.

Fourme, R. & Hendrickson, W. A. (1990). *Synchrotron Radiation and Biophysics*, edited by S. S. Hasnain, pp. 156–175. Chichester: Harwood.

Grossmann, J. G., Ingledew, W. J., Harvey, I., Strange, R. W. & Hasnain, S. S. (1995). *Biochemistry*, **34**, 8406–8414.

Guss, J. M., Merritt, E. A., Phizackerly, R. P., Hedman, B., Murata, M., Hodgson, K. O. & Freeman, H. C. (1988). *Science*, **241**, 806–811.

Hao, Q., Gu, Y. X., Zheng, C. D. & Fan, H. F. (1996). *OASIS*: a computer program for breaking the phase ambiguity in OAS or *SIR* protein data, School of Applied Sciences, De Montfort University, Leicester LE1 9BH, England, & Institute of Physics, Chinese Academy of Sciences, Beijing 100080, PR China.

Hao, Q. & Woolfson, M. M. (1989). *Acta Cryst.* A**45**, 794–797.

Hauptman, H. (1982). *Acta Cryst.* A**38**, 632–641.

Hauptman, H. (1997). *Curr. Opin. Struct. Biol.* **7**, 672–680.

Helliwell, J. R. (1992). *Macromolecular Crystallography with Synchrotron Radiation*, ch. 9. Cambridge University Press.

Hendrickson, W. A. (1991). *Science*, **254**, 51–58.

Hendrickson, W. A., Horton, J. R. & Lemaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.

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

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Nunzi, F., Woudstra, M., Campese, D., Bonicel, J., Morin, D. & Bruschi, M. (1993). *Biochim Biophys Acta*, **1162**, 28–34.
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.
Ralph, A. C. & Woolfson, M. M. (1991). *Acta Cryst.* A**47**, 533–537.
Ramachandran, J. N. & Raman, S. (1956). *Curr. Sci.* **25**, 348–351.
Ronk, M., Shively, J. E., Shute E. A. & Blake, R. C. (1991). *Biochemistry*, **30**, 9435–9442.
Staudenmann, J. L., Hendrickson, W. A. & Abramowitz, R. (1989). *Rev. Sci. Instrum.* **60**, 1939–1942.
Walter, R. L., Ealick, S. E., Friedman, A. M., Blake, R. C., Proctor, P. & Shoham, M. (1996). *J. Mol. Biol.* **263**, 730–751.
Yano, T., Fukumori, Y. & Yamanaka, T. (1991). *FEBS Lett.* **288**, 159–162.
Zheng, X. F., Fan, H. F., Hao, Q., Dodd, F. E. & Hasnain, S. S. (1996). *Acta Cryst.* D**52**, 937–941.