Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 18 February 2003
Accepted 2 May 2003

S-SWAT (softer single-wavelength anomalous technique): potential in high-throughput protein crystallography

Andrzej Olczak,[a] Michele Cianci,[b] Quan Hao,[c] Pierre J. Rizkallah,[b] James Raftery[d] and John Richard Helliwell[d]*

[a]Institute of General and Ecological Chemistry, Technical University of Łodz, Zwirki 36, 90-924 Łodz, Poland, [b]CCLRC Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, England, [c]MacCHESS, 273 Wilson Synchrotron Laboratory, Cornell University, Ithaca, NY 14853-8001, USA, and [d]Department of Chemistry, University of Manchester, Manchester M13 9PL, England. Correspondence e-mail: john.helliwell@man.ac.uk

The drive for yet higher output of protein crystal structures sustains an interest in streamlining data-collection protocols. Highly organized automated robotics attend to one aspect. The data-collection elapsed time also depends on the number of wavelengths used, the data redundancy and the desired diffraction resolution limit. Synchrotron-radiation beamlines offer a centralized data-collection approach with considerable flexibility of choice of wavelength(s) with high beam intensity and fine collimation. The arrival of automated IPs (imaging plates) and CCDs (charge-coupled devices) with sample freezing has widened the scope for ease of data collection and achieving a high redundancy, the latter at the cost of elapsed time per sample. This paper reports a single-wavelength approach where a large optimized f'' signal involving the Xe $L1$-absorption edge is harnessed. Even with modest data redundancy (7), a high-quality electron-density map was obtained from that single data set when combined with phase improvement methods. Two phasing protocols are compared for the same data set. Resolution extension (to 1.4 Å), in this test, was also done but *via* a second data set. In future, with a tilted detector geometry, full diffraction resolution will be collectable in one experiment on the one beamline. Moreover, careful minimization of beamline Be window thickness on the new Daresbury SRS Multipole Wiggler Beamline 10 will maximize the 2–3 Å (softer) wavelength intensity performance specification. The work has application to the harnessing of iodine f'' optimized signals and enhanced sulfur f'', as well as xenon f'', in protein crystallography. This approach, softer-SWAT, should 'swat' many protein structures when it comes online.

1. Introduction

Multiple- (*i.e.* three or more) wavelength anomalous dispersion (MAD) (*e.g.* see Phillips & Hodgson, 1980; Karle, 1980; Kahn *et al.*, 1985; Hendrickson, 1991; Hendrickson & Ogata, 1997) techniques have dominated synchrotron-radiation-based phasing of protein crystal structures in the last decade. Two-wavelength 'TW' (Okaya & Pepinsky, 1956; Helliwell, 1979, 1984; Peterson *et al.*, 1996; Hädener *et al.*, 1999; González, 2003) phasing at synchrotron-radiation (SR) sources is now growing in popularity as a more beam-time-efficient method. Selenomethionine incorporation into recombinant proteins (Hendrickson *et al.*, 1990) is the most favoured method of preparing samples suitable for MAD or OAS (one-wavelength anomalous scattering, otherwise known as SAD) phasing. This paper addresses an alternative approach involving use of softer X-rays at a single wavelength (softer-SWAT; SWAT is single-wavelength anomalous technique) in optimizing Xe f''. It is also relevant in enhancing the use of sulfur f''. 2 Å wavelength was termed softer X-rays (Chayen *et al.*, 2000), since the tradition for diffraction involves Cu $K\alpha$ radiation (1.54 Å) or harder X-rays delivered by synchrotron sources (typically 0.9 Å), and soft is usually taken to mean 5 Å or longer wavelength (Behrens *et al.*, 1998; Carpentier *et al.*, 2000). The feasibility of using softer X-rays had already been determined with experimental work using a lysozyme crystal up to 2.6 Å wavelength (Helliwell, 1983). This was based on the development and utilization of 2, 1.89, 1.74

Table 1
Data-collection statistics (numbers in parentheses are for the highest resolution shell) of the softer-wavelength data set used.

For details of the λ = 0.87 Å data set to 1.4 Å resolution, see Cianci *et al.* (2001).

Station at SRS	7.2
Wavelength (Å)	2.0
Detector	MAR345 (300 mm)
Distance (mm)	109.4
Oscillation range per image (°)	1.0
Number of images	200
Resolution (Å) (last shell)	64.5–2.3 (2.38–2.3)
Total number of reflections	118157
Number of unique reflections	16703
Multiplicity	7.1
Overall completeness	99.7 (100.0)
$R_{\rm merge}$ (%)	7.3 (14.5)
Overall $I/\sigma(I)$	23.4 (12.3)
Reflections with $I > 3\sigma(I)$ in the highest resolution shell (%)	80.4
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	
a	41.11
b	79.81
c	109.86

and 1.488 Å wavelengths with a variety of protein crystals and anomalous-dispersion utilization (Helliwell, 1984).

The relatively high atomic number (54) and anomalous dispersion of xenon makes it useful for MIR (multiple isomorphous replacement) and SIROAS (single isomorphous replacement with optimized wavelength anomalous scattering) experiments in protein crystallography. The history of xenon in crystallography started in 1965 as an investigation of its anaesthetic action (Schoenborn *et al.*, 1965). The use of xenon, as a heavy atom for phase determination, was later suggested by Schoenborn & Featherstone (1967). A recent review on the binding of xenon to proteins and the development of associated techniques has been published by Fourme *et al.* (1999). Xenon protein derivatives have been obtained by subjecting a native protein crystal to a xenon-gas pressure in the range from 1×10^5 to 50×10^5 Pa (Schoenborn *et al.*, 1965; Tilton *et al.*, 1984; Vitali & Robbins, 1991; Schiltz *et al.*, 1994; Djinovic Carugo *et al.*, 1998; Vernède & Fontecilla-Camps, 1999). Xenon atoms diffuse rapidly towards potential interaction sites through the solvent channels of a protein crystal (Tilton *et al.*, 1984; Montet *et al.*, 1997; Prangé *et al.*, 1998). The interaction of xenon with proteins is the result of non-covalent weak-energy van der Waals forces (Tilton *et al.*, 1984, 1986), and therefore the process of xenon binding is reversible (Tilton & Kuntz, 1982). The number and occupancy of xenon binding sites varies with the applied pressure (Tilton *et al.*, 1984). The xenon binding induces very limited perturbations to the crystal structure, so that xenon protein complexes are almost always highly isomorphous with the native crystal (Schoenborn & Featherstone, 1967; Prangé *et al.*, 1998). Xenon is able to bind to a protein in closed intramolecular hydrophobic cavities, accessible active sites, intermolecular cavities and channel pores (Prangé *et al.*, 1998). The use of flash-cooled loop-mounted crystals proved to be a solution (Sauer *et al.*, 1997) to the problem of xenon clathrate formation in capillaries (Schiltz *et al.*, 1994). Since the half-life of xenon residence has been estimated to be of the order of 20 s (Soltis *et al.*, 1997), crystals can be flash-cooled shortly after being removed from a xenon atmosphere (Soltis *et al.*, 1997; Sauer *et al.*, 1997).

Recently, the structure of apocrustacyanin A1 (Cianci *et al.*, 2001) has been solved by means of the single isomorphous replacement (Xe) with optimized wavelength (2 Å) anomalous scattering method (SIROAS). The hand was determined using the sulfur anomalous signal. As a result of a large xenon f'' at this wavelength (11.5 e, 4 e more than the value at Cu $K\alpha$ (7.2 e) and 8 e more than at 0.9 Å (~3 e), a popular generic data-collection wavelength), more than a thousand reflections with $\Delta F_{\rm ano} > 3\sigma(\Delta F_{\rm ano})$ to 2.3 Å resolution were collected. In view of the mounting evidence that one-wavelength anomalous scattering may be sufficient to solve protein structures [starting with Wang (1985); and more recently see, for example, Hao *et al.* (2000) and Dauter *et al.* (2002)], it is valuable to consider the apocrustacyanin A1 case to assay the viability of a single-wavelength anomalous technique (SWAT) being applied to phase the structure, even with low occupancies of four xenon sites (~0.4, ~0.4, ~0.3, ~0.1, respectively).

Here we report Xe *L*1 anomalous-based phasing *via* two different computational routes. One route used *SnB* (Miller *et al.*, 1994; Weeks & Miller, 1999), *MLPHARE* (Otwinowski, 1991) and *DM* (Collaborative Computational Project, Number 4, 1994). The second route used *OASIS* (Hao *et al.*, 2000) and *DM* (Collaborative Computational Project, Number 4, 1994). *OASIS* was especially written to determine phases from one-wavelength anomalous-scattering data instead of using additional multiwavelength diffraction data (Hao *et al.*, 2000). As the xenon derivatives can be obtained

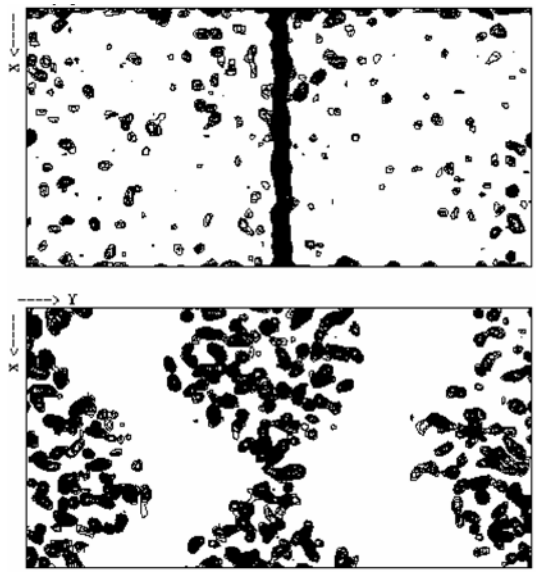

Figure 1
(top) Slab of 10 Å [*MLPHARE*(Xe-ano)] phased $F_{\rm obs}$ electron density at z = 1/2; (bottom) the same slab after density modification with *DM*. Electron density contoured at 2.5±0.5 r.m.s.

from native protein crystals with commercially available equipment in a relatively short time (even less than half an hour), the method described in this paper may provide a good alternative to TW, MAD or MIR phasing, in particular, when high-throughput is desirable. This case study has also led to the optimization of the new SRS MPW MAD 10 beamline under construction at Daresbury, namely to achieve firstly Be windows as thin as possible to maximize softer X-ray intensity and secondly a high-2Θ data-collection coverage. The case study provides a foundation for new biological crystallography studies and synchrotron PX-beamline design. We might be asked why another acronym, SWAT? SWAT can stand for strategy, weapons and tactics. This case study does inform in that way. Also, to swat a problem is to see it solved; that also seems appropriate here in solving the phase problem. In both ways, SWAT is a better acronym than SAD.

2. Methods

Crystals of lobster apocrustacyanin A1 subunit were prepared as reported in Chayen *et al.* (2000) and Cianci *et al.* (2001). X-ray data collection was undertaken at SRS in Daresbury (UK) using station 7.2 (Helliwell *et al.*, 1982) where the Ge (111) single-crystal monochromator was set to reflect X-rays at $\lambda = 2.0$ Å chosen to optimize the xenon $L1$-edge f'' (wavelength 2.27 Å). A wavelength of 2 Å was the longest practicable to allow reasonable diffraction resolution for the detector solid-angle apertures available and a manageable spot size due to oblique incidence. The crystal was mounted in a nylon loop and pressurized in 14 atmospheres (1.419 MPa)[1] of xenon gas using an OxfordXcell (Oxford Cryosystems, 1999) for an hour and then flash cooled with a nitrogen gas stream at approximately 100 K. No cryoprotectant was used. A MAR 345 image-plate detector system was used and 187 images of 1.0° each were collected with exposure times of about 7 min per frame in dose mode. No special precautions were necessary to collect the data at 2 Å wavelength, *i.e* an air path was used and the detector front-entrance window was a standard one; obviously a helium, instead of an air, path and a thinner detector window would be optimizations in the future. The *HKL* package (Otwinowski, 1993) was used for the integration, scaling and merging of the data set. The unit-cell parameters, data collection and processing statistics are reported in Table 1. With a minimum crystal-to-detector distance of 109.4 mm and a 2 Å wavelength, the maximum accessible resolution was 2.3 Å.

3. SWAT (single-wavelength anomalous technique)

3.1. Locating the xenon sites with *SnB* and phasing with *MLPHARE*

Initial anomalous difference $|E_{\rm anom}|$ values were calculated using the subroutine *DREAR* (Howell *et al.*, 2000) of the *SnB* (Miller *et al.*, 1994; Weeks & Miller, 1999) package. 350

[1] Cianci *et al.* (2001) contains a typographical error: 1419 MPa should be 1.419 MPa.

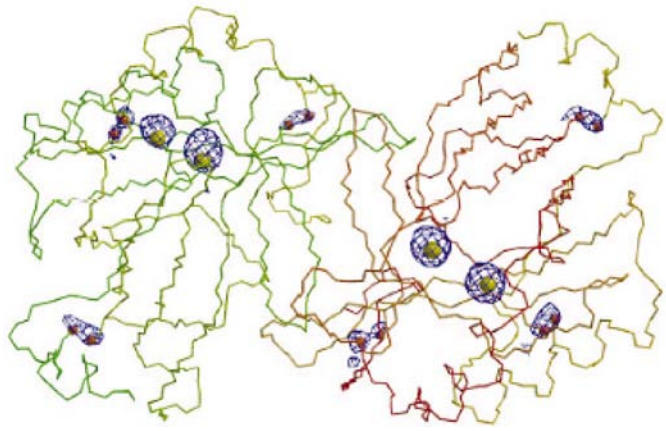

Figure 2
Overview of apocrustacyanin A1 dimer and anomalous difference Fourier electron-density map [*MLPHARE*(Xe-ano)/*DM*] contoured at 2.0σ. Xe atoms are coloured yellow; disulfide bridge sulfur atoms are in orange. Chain *A* is in red (right) and chain *B* in green (left). Figure drawn using *RASTER3D* (Merritt & Bacon, 1997).

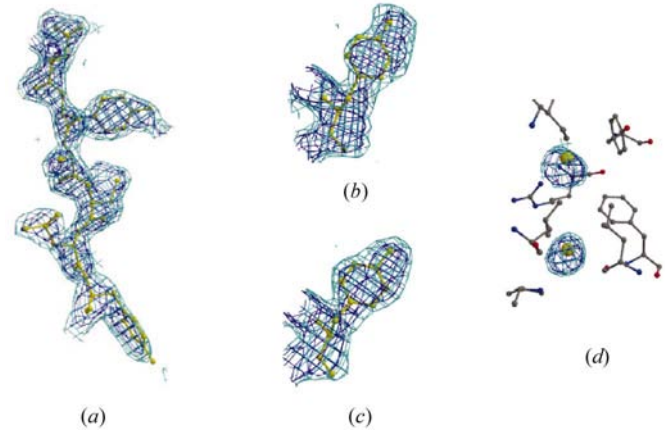

Figure 3
An example of $F_{\rm obs}$ electron-density portions [phases from *MLPHARE*(Xe-ano)/*DM*, 2 r.m.s.] overlayed on fragments of: (*a*) the main chain; (*b*) residue Tyr112 A; (*c*) residue Tyr112B; (*d*) around Xe atoms 1 and 2. The model is the original apocrustacyanin A1 model.

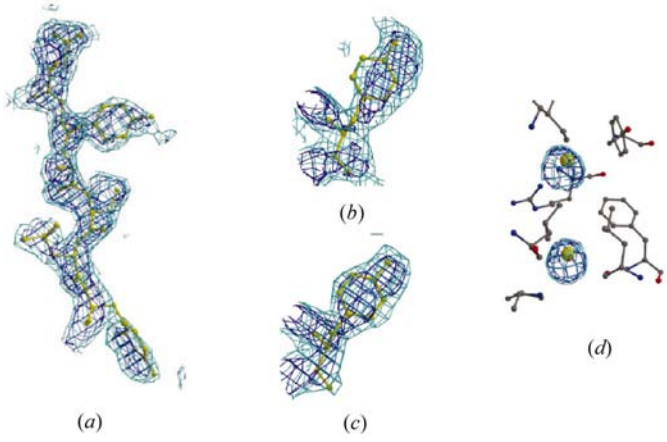

Figure 4
The same portions of $F_{\rm obs}$ electron density as in Fig. 3 in the map phased with *OASIS*/*DM*, 2 r.m.s. The model is the original apocrustacyanin A1 model.

Table 2
The highest ten peaks in the successful *SnB* run.

X	*Y*	*Z*	Height	Rank
0.626529	0.521407	0.504239	23.75	1
0.869992	0.560510	0.990572	23.07	2
0.644868	0.763206	0.878044	19.42	3
0.854668	0.802717	0.623400	12.15	4
0.646469	0.684821	0.996500	9.47	5
0.641464	0.601778	0.510662	7.82	6
0.854857	0.723594	0.622139	7.68	7
0.536744	0.789131	0.847142	7.07	8
0.755788	0.900091	0.684507	6.49	9
0.885455	0.942950	0.864941	5.88	10

reflections ($|E_{\rm anom}| > 3\sigma$) and 5000 triplets to 2.5 Å resolution were used to locate the anomalous-scatterer sites by the *SnB* method. A clear bimodal distribution indicated that a correct solution was found. Table 2 lists the top ten peaks from the *SnB* run (5000 trials). The three strongest peaks represent xenon and were clearly above the fourth peak (the eighth peak was very close to the coordinates of the fourth xenon site located later in an anomalous difference Fourier electron-density map).

The experimental electron-density map, based on FP^{+}_{Xe} and FP^{-}_{Xe}, obtained with *MLPHARE* showed some artificial features (streaks of higher electron density), presumably due to lack of centric reflections in the phasing process and the phase ambiguity in the acentrics (Fig. 1, top). At this point, the figure of merit (FOM) was 0.30. However, after density modification with *DM* (Collaborative Computational Project, Number 4, 1994), the quality of the map greatly improved with overall FOM_{dm} = 0.69 (Fig. 1, bottom). The anomalous difference Fourier map with these *DM* output phases restored the three Xe-atom positions found by *Shake-and-Bake*, and revealed a fourth Xe-atom position with 0.1 occupancy and the six disulfide bridges present in the apocrustacyanin A1 dimer. Fig. 2 shows the anomalous difference electron-density map (2.3 Å resolution, 2 r.m.s. level) overlaid on the refined apocrustacyanin A1 model with the six disulfides and the four xenon-atom sites marked. The Xe-atom and disulfide coordinates found in this way allowed identification of a twofold non-crystallographic symmetry (NCS) axis relating the two expected monomers. Apart from the four major xenon sites and the six disulfides, four minor xenon sites (or ammonium sulfate) were identified in the *DM*-phased anomalous Fourier map. Two of these new sites are essentially sitting on the twofold NCS axis and the other two are NCS related. The main xenon sites were used to calculate new phases and the NCS transformation matrix for electron-density averaging with *DM*. The resultant figure of merit was 0.71. Without NCS averaging, the FOM was 0.70.

3.2. Locating the xenon sites and phasing with *OASIS*

The Xe anomalous scattering sites were located by the conventional direct-methods program *SAPI* (Fan *et al.*, 1990), using the magnitudes of the anomalous differences, for reflections up to 3.0 Å resolution. The solution was found with a default run of the program. The largest 416 $|E_{\rm anom}|$s were used in the tangent formula phase refinement. The resultant electron-density map produced a group of three highest peaks; there was a clear gap between this group and other peaks in terms of peak height. A Karle-recycle refinement (an option in *SAPI*) of these three sites yielded the additional, fourth, minor site. The absolute configuration of these sites was determined by the program *ABS* based on the P_s-function method (Woolfson & Yao, 1994). These four Xe sites were the correct sites and formed the basis for the next phasing step. The computer program *OASIS* (Hao *et al.*, 2000) was then used for the *ab initio* phasing of the OAS data. All Friedel pairs (including centric reflections) were harnessed using *OASIS*. Density modification using the CCP4 program *DM* (Collaborative Computational Project, Number 4, 1994) was then applied to the resulting phase sets. The figure of merit after *DM* was 0.71.

Table 3
FOM statistics for SIROAS/*DM* phases (Cianci *et al.*, 2001), *MLPHARE*(Xe-ano)/*DM* and *OASIS*/*DM versus* resolution.

			Mean_FOM		
Resolution (Å)	$1/(Å)^2$	No. of reflections	SIROAS/*DM*	*MLPHARE* (Xe-ano)/*DM*	*OASIS*/*DM*
9.22	0.012	322	0.75	0.565	0.623
6.52	0.024	536	0.83	0.677	0.678
5.32	0.035	665	0.85	0.804	0.787
4.61	0.047	783	0.88	0.859	0.813
4.12	0.059	878	0.85	0.849	0.804
3.76	0.071	964	0.82	0.814	0.774
3.48	0.082	1021	0.81	0.804	0.775
3.26	0.094	1116	0.73	0.711	0.718
3.07	0.106	1197	0.69	0.694	0.678
2.91	0.118	1228	0.72	0.703	0.685
2.78	0.129	1276	0.72	0.725	0.685
2.66	0.141	1376	0.71	0.725	0.699
2.56	0.153	1427	0.74	0.714	0.698
2.46	0.165	1434	0.74	0.709	0.706
2.38	0.177	1524	0.71	0.681	0.681
2.30	0.188	1560	0.353	0.35	0.662

4. Quality of the electron-density maps

The final electron-density maps clearly revealed the solvent boundary in both cases. The protein Cα trace was clearly visible but there were a number of places where the electron density was discontinuous. A correlation coefficient between the *MLPHARE*(Xe-ano)/*DM* phased map and the SIROAS/*DM* map of Cianci *et al.* (2001) was 0.60. A correlation coefficient between the *OASIS*/*DM*-phased map and the SIROAS/*DM* map of Cianci *et al.* (2001) was 0.67. Table 3 reports the FOM as a function of resolution. Using either set of phases derived for the Xe derivative, which have a resolution limit of 2.3 Å, the visual quality of the maps indicated that the model could be manually traced (Figs. 3 and 4). However, *ArpWarp6* (Morris *et al.*, 2002) failed to auto trace the main chain at this resolution (initial connectivity index was less than 0.40; in general, for *ArpWarp6* to work, the initial connectivity index

Table 4
ArpWarp6 statistics *versus* resolution.

Resolution (Å)	*OASIS*/*DM*			*MLPHARE*(Xe-ano)/*DM*		
	$R_{\rm factor}$	$R_{\rm free}$	Connectivity index	$R_{\rm factor}$	$R_{\rm free}$	Connectivity index
2.2	0.273	0.496	0.45	0.265	0.43	0.69
2.1	0.265	0.456	0.68	0.254	0.411	0.79
2.0	0.259	0.426	0.77	0.243	0.377	0.87
1.8	0.265	0.381	0.86	0.250	0.362	0.90
1.6	0.231	0.307	0.94	0.225	0.294	0.95
1.4	0.227	0.288	0.94	0.224	0.286	0.96

should be greater than 0.5). Also, the number of X-ray reflections should be at least six times higher than the number of atoms in the model, according to the manual, rather than the five achieved in this work (16703 reflections in the 2.3 Å data set, and around 3000 protein atoms and 400 heteroatoms in the final model, PDB code: 1h91).

The two different sets of phases were then extended with *DM* (Collaborative Computational Project, Number 4, 1994) to 1.4 Å to further test their quality. The data set to 1.4 Å was the native collected on SRS station 9.6 and used for refinement of the crystal structure of apocrustacyanin A1 (Cianci *et al.*, 2001).

The phase error statistics against the model (PDB code: 1h91) phases at 1.4 Å were 40° SIROAS/*DM* (Cianci *et al.*, 2001), 46° for *OASIS*/*DM*, and 50° for *MLPHARE*(Xe-ano)/*DM*.

The two new phase sets to 1.4 Å derived with *MLPHARE*(Xe-ano)/*DM* and *OASIS*/*DM* were then tested again with *ArpWarp6* (Morris *et al.*, 2002). This time the runs were successful with very similar results in terms of connectivity index [*MLPHARE*(Xe-ano)/*DM* = 0.96, *OASIS*/*DM* = 0.94], *R* factor [*MLPHARE*(Xe-ano)/*DM* = 0.224, *OASIS*/*DM* = 0.227] and $R_{\rm free}$ [*MLPHARE*(Xe-ano)/*DM* = 0.286, *OASIS*/*DM* = 0.288]. Table 4, reporting *ArpWarp6* run statistics *versus* resolution, clearly shows that the model could be almost completely traced (*i.e.* connectivity index >0.90) combining these sets of phases and data with a resolution of just 1.8 Å.

5. Discussion

The dimer of apocrustacyanin A1 consists of approximately 3000 light atoms and 12 sulfur atoms. In the xenon derivative, we identified four major xenon sites with 0.4, 0.4, 0.3 and 0.1 occupancies.

The structure contains two types of anomalous scatterers. The question arises how to estimate the average anomalous signal $\Delta_{\rm ano} = \langle|\Delta F|\rangle/\langle F\rangle$ in such a situation. Hendrickson & Teeter's (1981) formula can be applied in the case of the single anomalously scattering element only. Weiss *et al.* (2001) gave an approximate estimate of the expected $\Delta_{\rm ano}$ for a protein by using a generalized version of the equation published by Hendrickson & Teeter:

Table 5
Contributions of separate sulfur and xenon anomalous scattering to the overall value at two different wavelengths.

Parameter	$\Delta_{\rm ano}$ (λ = 1.54 Å)	$\Delta_{\rm ano}$ (λ = 2.00 Å)
$f'_{\rm S}$	0.3	0.4
$f''_{\rm S}$	0.55	0.95
$f'_{\rm Xe}$	−0.5	−3.5
$f''_{\rm Xe}$	7.4	11.1
$\Delta^{A}_{\rm ano}$ (%)	2.2	3.3
$\Delta^{B}_{\rm ano}$ (%)	0.7	1.1
$\Delta^{AB}_{\rm ano}$ (%)	0.3	0.4
Total (%)	2.3	3.5

Table 6
$\langle|\Delta F_{\rm ano}|\rangle/\langle F\rangle$, $\langle|\Delta F_{\rm ano}|\rangle/\langle\sigma_{\rm ano}\rangle$ and number of large anomalous differences $N_{\rm ref}$ *versus* resolution.

Resolution shells (Å)		$\langle|\Delta F_{\rm ano}|\rangle/\langle F\rangle$	$\langle|\Delta F_{\rm ano}|\rangle/\langle\sigma_{\rm ano}\rangle$	$N_{\rm ref}$
34.68	5.11	0.063	1.60	261
5.11	4.06	0.045	1.37	174
4.06	3.54	0.045	1.18	124
3.54	3.22	0.048	1.17	111
3.22	2.99	0.057	1.23	107
2.99	2.81	0.066	1.19	112
2.81	2.67	0.074	1.06	69
2.67	2.56	0.075	0.95	50
2.56	2.46	0.077	0.83	38
2.46	2.37	0.079	0.78	26
2.37	2.30	0.085	0.70	26

$$\Delta F/F = \left(2\sum_i N_i \Delta f_i''^2 \Big/ \sum_i N_i f_i^2\right)^{1/2},$$

where both sums run over all atoms i in the structure, N_i is the number of atoms of type i and $\Delta f_i''^2$ and f_i are the anomalous scattering factor and the atomic form factor, respectively, at zero scattering angle of an atom of type i.

In this study, we present (Appendix *A*) a detailed derivation of a formula for the average anomalous signal $\Delta_{\rm ano}$ in the case of two different scattering elements (say *A* and *B*). (It can be easily generalized to any number of anomalously scattering elements.) Interestingly, $\Delta_{\rm ano}$ consists of not only individual contributions $\Delta^{A}_{\rm ano}$ and $\Delta^{B}_{\rm ano}$ coming from scatterers *A* and *B* but also contains a mixed term $\Delta^{AB}_{\rm ano}$:

$$\Delta_{\rm ano} = [(\Delta^{A}_{\rm ano})^2 + (\Delta^{B}_{\rm ano})^2 + (\Delta^{AB}_{\rm ano})^2]^{1/2}. \quad (1)$$

To simplify calculations of the anomalous signal [based on formula (1)], we assumed 0.4 occupation for all four xenon sites and the results are summarized in Table 5. The table shows that the overall increment of the theoretical $\Delta_{\rm ano}$ while moving from λ = 1.54 to λ = 2.0 Å is 50%. If we look at each single contribution, $\Delta^{A}_{\rm ano}$ (related to xenon) increased by 50%, $\Delta^{B}_{\rm ano}$ (sulfur related) by 57% and $\Delta^{AB}_{\rm ano}$ by 33%. Overall, the anomalous signal is strongly dominated by Xe because the total anomalous signal is only slightly larger than the Xe signal itself. This validates the initial choice of using only the four Xe atoms to calculate the phases. The anomalous diffraction signal ($\Delta F_{\rm anom}$) coming from four Xe atoms at, say, 40% occupation each, is 3.3%. To get that average signal, at the

same wavelength, from S atoms only (excluding a reinforcement of signal at lower resolutions from the S atoms being in disulfides), one would need about 80 of them. This arises from a weaker inherent f'' signal and because they are spread over 80 sites.

The experimental value of Δ_{ano}, 4.5% at $\lambda = 2.0$ Å, is larger than the theoretical one (Table 6). This is understandable since experimental data apart from a real anomalous signal contain also statistical fluctuations and possible systematic errors (absorption, scaling errors resulting from beam instability, *etc*). In principle, this experimental value should be taken at $1/d = 0$ resolution limit but, because of the intensity statistics, characteristic to protein structures, the parameter takes its minimum at roughly 5–4 Å resolution (Table 6). It is obvious that this value is a better approximation to the calculated one than the value taken at the lowest resolution shell.

Parameter N_{ref} in Table 6 represents the number of large anomalous differences $[|\Delta F| > 3\sigma(\Delta F)]$ in a given resolution shell. When σ's of structure-factor amplitudes are correctly estimated, the behaviour of this number against resolution could be a useful characteristic allowing estimation of the high-resolution limit of usable anomalous data. This limit is determined at the resolution where N_{ref} starts to increase. In the case of the xenon derivative of the apocrustacyanin A1 data set, almost the full resolution range up to 2.4 Å was usable.

6. Conclusions

The combination of an Xe derivative (total occupation 1.2 Xe atoms per dimer consisting of ~3000 protein atoms) and 2 Å wavelength anomalous data successfully allowed the phasing of the structure of apocrustacyanin A1. This softer-SWAT approach was as successful as the SIROAS approach used in Cianci *et al.* (2001), demonstrated by the successful chain tracing in *ArpWarp6.0* runs. We also showed that the SWAT data were equally viable for structure solution using *OASIS*, *i.e* their utilization was not software-package dependent; with this package, the total CPU time was about 3 min on an Alpha XP10000 workstation. The number of large anomalous differences $[\Delta F_{ano} > 3\sigma(\Delta F_{ano})]$ as a function of resolution served as a simple indicator to estimate the resolution limit for usable anomalous signal data.

As xenon derivatives can be obtained from native protein crystals using commercially available equipment in a relatively short time (tens of minutes), the method described in this paper (softer-SWAT) may provide an attractive alternative to TW, MAD or MIR phasing, in particular when high-throughput is desirable. Also, the phase extension to 1.4 Å worked very well even where the phases of the xenon derivative data set were used directly in the native data set. This is probably due to the high isomorphism of the native and derivative data sets, and the relatively small contribution of the xenon atoms to the scattering properties of the Xe–protein crystal. The good quality of the phases obtained at 2.3 Å could allow automatic tracing of the model to have been achieved even more easily with Xe derivative data extending to just 1.8 Å. The diffraction geometry was the limitation in this experiment. The method of off-setting the detector in the vertical direction or tilting, *i.e.* changing 2Θ, would get round this limitation. This case study, in fact, has indeed driven the design specifications for a new multipole wiggler beamline under construction at the Daresbury SRS. The new SRS MPW MAD 10 beamline, for high-throughput protein crystallography, will be fitted with: (i) a MAR desktop beamline with sample autochanger and will allow a $2\Theta_{max}$ of 110° at maximum tilt and closest detector-to-crystal distance; (ii) by careful minimization of the beamline Be window thickness, the intensity at 2 up to 3 Å wavelength will be maximized. Thus, besides TW and MAD experiments from 0.8 Å wavelength upwards, it will also be feasible on the SRS MPW MAD 10 beamline to collect high-throughput single-wavelength data sets, measured at a softer wavelength as well (softer-SWAT). As a result, optimized xenon $L1$ f'' and enhanced sulfur f'' signals, and signals from a number of other softer λ-edge anomalous scatterers (*e.g.* iodine) will also be obtained. This will extend the use of softer-SWAT, also called 'jolly SAD' applications (Dauter *et al.* (2002), to even more challenging cases, such as those involving fewer sulfur atoms per kDa in the protein crystal asymmetric unit as well as weak Xe substitution.

APPENDIX *A*
Mathematical formalism of the anomalous-scattering contributions

In order to estimate the effectiveness of the xenon contribution to the anomalous scattering when other anomalously scattering atoms are present in a structure (sulfur in this case), it is convenient to analyse the total anomalous signal in terms of the individual (Xe and S) contributions. This has been done by one of us (AO).

The method presented here for two types of anomalous scatterers can be easily extended to any number of them.

A1. Nomenclature

The usual crystallographic conventions have been used. In addition: $\langle X \rangle$ is the average value of X in a resolution shell, F^+, F^-, $F = \frac{1}{2}(F^+ + F^-)$ are the structure-factor amplitudes, $\Delta F = F^+ - F^-$ is the Bijovet difference, $\Delta F^2 = (F^+)^2 - (F^-)^2$ is the squared F^+ and F^- difference, $\Delta_{ano} = \langle|\Delta F|\rangle/\langle F\rangle$ is the anomalous signal, $f_j^c = f_j + f_j' + if_j'' = f_{oj}e^{i\delta_j}$ is the complex atomic scattering factor of the jth atom, $\Delta_{jk} = \delta_j - \delta_k$ by definition, $\mathbf{r}_j$ is the jth-atom position in the unit cell and $\mathbf{H}$ is the reciprocal-space vector.

A2. Anomalous signal in the case of two types of anomalous scatterers

Using structure-factor amplitude distribution $P(F)$ for acentric reflections $P(F) = (2F/\langle F^2\rangle)\exp(-F^2/\langle F^2\rangle)$, one can easily show that

$$\langle F\rangle^2 = (\pi/4)\langle F^2\rangle.$$

Moreover, using the identity $2F\Delta F = \Delta F^2$ and assuming that anomalous difference amplitude $|\Delta F|$ and structure-factor amplitude are not correlated [Crick & Magdoff (1956) justified this approximation in a similar context of isomorphous signal calculations], which means that $\langle F|\Delta F|\rangle = \langle F\rangle\langle|\Delta F|\rangle$, one can write:

$$\langle|\Delta F|\rangle/\langle F\rangle = \langle|\Delta F^2|\rangle/2\langle F\rangle^2$$

and consequently:

$$\Delta_{\rm ano} = (2/\pi)(\langle|\Delta F^2|\rangle/\langle F^2\rangle). \quad (2)$$

Using a classical expression for the structure factor $\mathbf{F}$,

$$\mathbf{F} = \sum_{j=1}^{N} f_j^c \exp(2\pi i\mathbf{H}\cdot\mathbf{r}_j)$$

(N stands for the number of atoms in the unit cell), it can be shown that

$$\Delta F^2 = -4\sum_{k=1}^{N}\sum_{j>k}^{N} f_{oj}f_{ok}\sin\Delta_{jk}\sin(2\pi\mathbf{H}\cdot\Delta\mathbf{r}_{jk}), \quad (3)$$

where $\Delta\mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k$ by definition.

Considering only two types of anomalous scatterers (A and B) present in a given structure, we can put all the atoms in the following order:

$$j = 1,\ldots,N_A, N_A+1,\ldots,N_A+N_B, N_A+N_B+1,\ldots,N,$$

where N_A is the number of anomalous scatterers of type A and N_B is the number of anomalous scatterers of type B.

Now we can split the right-hand side of (3) into three terms:

$$\begin{aligned}\Delta F^2 = &-4\sin\Delta_{AL}f_{oA}f_{oL}\sum_{k=1}^{N_A}\sum_{j=N_A+N_B+1}^{N}\sin(2\pi\mathbf{H}\cdot\Delta\mathbf{r}_{jk})\\ &-4\sin\Delta_{BL}f_{oB}f_{oL}\sum_{k=N_A+1}^{N_A+N_B}\sum_{j=N_A+N_B+1}^{N}\sin(2\pi\mathbf{H}\cdot\Delta\mathbf{r}_{jk})\\ &-4\sin\Delta_{AB}f_{oA}f_{oB}\sum_{k=1}^{N_A}\sum_{j=N_A+1}^{N_A+N_B}\sin(2\pi\mathbf{H}\cdot\Delta\mathbf{r}_{jk}), \quad (4)\end{aligned}$$

where N_L is the number of light, non-absorbing atoms (carbon, nitrogen, oxygen).

Each of the three sums in (4) can be treated as a random variable with effectively Gaussian distributions (because of a large enough number of components in each of them), whose average value is equal to zero as a consequence of the relation $\langle\sin(2\pi\mathbf{H}\cdot\Delta\mathbf{r}_{jk})\rangle = 0$. Since $\langle\sin^2(2\pi\mathbf{H}\cdot\Delta\mathbf{r}_{jk})\rangle = 1/2$, the standard uncertainties of these variables can be expressed in the following way:

$$\begin{aligned}\sigma_A &= 4\sin\Delta_{AL}f_{oA}f_{oL}(\tfrac{1}{2}N_AN_L)^{1/2}\\ \sigma_B &= 4\sin\Delta_{BL}f_{oB}f_{oL}(\tfrac{1}{2}N_BN_L)^{1/2}\\ \sigma_{AB} &= 4\sin\Delta_{AB}f_{oA}f_{oB}(\tfrac{1}{2}N_AN_B)^{1/2}. \quad (5)\end{aligned}$$

Consequently, ΔF^2 is a Gaussian random variable whose average value is zero and variance $\sigma^2 = \sigma_A^2 + \sigma_B^2 + \sigma_{AB}^2$, provided there is no correlation between the anomalous scatterer positions. Since ΔF^2 is a Gaussian random variable, the average value of $\langle|\Delta F^2|\rangle$ is equal to $(2/\pi)^{1/2}\sigma$.

Making the above substitutions to (2), we obtain:

$$\Delta_{\rm ano} = [(\Delta_{\rm ano}^A)^2 + (\Delta_{\rm ano}^B)^2 + (\Delta_{\rm ano}^{AB})^2]^{1/2}, \quad (6)$$

where $\Delta_{\rm ano}^j = (2\times 2^{1/2}/\pi^{3/2})(\sigma_j/\langle F^2\rangle)$ for $j = A, B, AB$.

We can see now that the total anomalous signal coming from two different scatterers is not an additive property and apart from the individual contributions from A and B contains also a mixed term $\Delta_{\rm ano}^{AB}$. In the case of only one type of anomalous scatterer in a structure (let's say A), expression (6) reduces to the following form:

$$\Delta_{\rm ano} = \Delta_{\rm ano}^A = \frac{8}{\pi^{3/2}}\frac{f_A''(f_L+f_L') - f_L''(f_A+f_A')}{N_Af_{OA}^2 + N_Lf_{OL}^2}(N_AN_L)^{1/2}. \quad (7)$$

If $f_L'' \simeq 0$, $f_L' \simeq 0$ and $N_Af_{OA}^2 \ll N_Lf_{OL}^2$, this can be further simplified:

$$\Delta_{\rm ano}^A = \frac{8}{\pi^{3/2}}\frac{f_A''}{f_L}\left(\frac{N_A}{N_L}\right)^{1/2}. \quad (8)$$

This form resembles a formula often used for estimation of anomalous signal (Dauter *et al.*, 2002), namely:

$$\Delta_{\rm ano} = 2^{1/2}\frac{f_A''}{f_L}\left(\frac{N_A}{N_L}\right)^{1/2} \quad (9)$$

(Crick & Magdoff, 1956).

Predictions based on both (8) and (9) are nearly the same because the ratio of $8/\pi^{3/2}$ to $2^{1/2}$ is ~1.016. However, we have to remember that these approximations always lead to slight overestimation. Having calculated individual contributions based on (7), (8) or (9), one can calculate the total anomalous signal using (6).

We thank The Leverhulme Trust and the EU Crystallogenesis Project for salary support (AO) and the EPSRC for studentship support (MC). MC's salary support has continued, as a PDRA, under a BBSRC Award at Daresbury (PI's S. S. Hasnain and J. R. Helliwell with 9 other 'NWSGC' Members). Access to the SRS facilities was awarded by the Joint Biology Program of the Research Councils of UK. This paper is a second article contribution from the North West Structural Genomics Consortium (NWSGC), Manchester node with Daresbury Laboratory. QH is supported under a NIH, USA, grant award RR-01646: MacCHESS, Cornell University.

References

Behrens, W., Otto, H., Sturmann, H. B. & Heyn, M. P. (1998). *Biophys. J.* **75**, 255–263.

Carpentier, P., Berthet-Colonines, C., Capitan, M., Chesne, M. L., Fonclon, E., Leguien, S., Sturhmann, H., Thiadiere, D., Vicart, J., Zelinski, P. & Kahn, R. (2000). *Cell Mol. Biol.* **46**, 915–935.

Chayen, N. E., Cianci, M., Olczak, A., Raftery, J., Rizkallah, P. J., Zagalsky, P. F. & Helliwell, J. R. (2000). *Acta Cryst.* D**56**, 1064–1066.

Cianci, M., Rizkallah, P. J., Olczak, A., Raftery, J., Chayen, N. E., Zagalsky, P. F. & Helliwell, J. R. (2001). *Acta Cryst.* D**57**, 1219–1229.

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

Crick, F. H. C. & Magdoff, B. S. (1956). *Acta Cryst.* **9**, 901–908.

Dauter, Z., Dauter, M. & Dodson, E. (2002). *Acta Cryst.* D**58**, 494–506.
Djinovic Carugo, K., Everitt, P. & Tucker, P. A. (1998). *J. Appl. Cryst.* **31**, 812–814.
Fan, H. F., Hao, Q., Gu, Y. X., Qian, J. Z., Zheng, C. D. & Ke, H. (1990). *Acta Cryst.* A**46**, 935–939.
Fourme, R., Shepard, W., Schiltz, M., Prangé, T., Ramin, M., Kahn, R., de la Fortelle, E. & Bricogne, G. (1999) *J. Synchrotron Rad.* **6**, 834–844.
González, A. (2003). *Acta Cryst.* D**59**, 315–322.
Hädener, A., Matzinger, P. K., Battersby, A. R., McSweeney, S., Thompson, A. W., Hammersley, A. P., Harrop, S. J., Cassetta, A., Deacon, A., Hunter, W. N., Nieh, Y. P., Raftery, J., Hunter, N. & Helliwell, J. R. (1999). *Acta Cryst.* D**55**, 631–643.
Hao, Q., Gu, Y. X., Zheng, C. D. & Fan, H. F. (2000). *J. Appl. Cryst.* **33**, 980–981.
Helliwell, J. R. (1979). Daresbury Study Weekend Proceedings DL/Sci R13, pp. 1–6.
Helliwell, J. R. (1983). Unpublished.
Helliwell, J. R. (1984). *Rep. Prog. Phys.* **47**, 1403–1497.
Helliwell, J. R., Greenhough, T. J., Carr, P. D., Rule, S. H., Moore, P. R., Thompson, A. W. & Worgan, J. S. (1982). *J. Phys. E*, **15**, 1563–1372.
Hendrickson, W. (1991). *Science*, **254**, 51–58.
Hendrickson, W. A., Horton, J. R., LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hendrickson, W. A. & Ogata, C. M. (1997). *Methods Enzymol.* **276**, 494–523.
Hendrickson, W. A. & Teeter, M. (1981). *Nature (London)*, **290**, 107–113.
Howell, P. L., Blessing, R. H., Smith, G. D. & Weeks, C. M. (2000). *Acta Cryst.* D**56**, 604–617.
Kahn, R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J. L., Dideberg, O. & Wery, J. P. (1985). *FEBS Lett.* **179**, 133–137.
Karle, J. (1980). *Int. J. Quantum Chem.* **7**, 356–367.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Miller, R., Gallo, S. M., Khalak, H. G. & Weeks, C. M. (1994). *J. Appl. Cryst.* **27**, 613–621.
Montet, Y., Amara, P., Volbeda, A., Vernède, X., Hetchikian, E. C., Martin, J. F., Frey, M. & Fontecilla-Camps, J. C. (1997). *Nature Struct. Biol.* **4**, 523–526.
Morris, R. J., Perrakis, A. & Lamzin, V. (2002). *Acta Cryst.* D**58**, 968–975.
Okaya, Y. & Pepinsky, R. (1956). *Phys. Rev.* **103**, 1645–1647.
Otwinowski, Z. (1991). Proc. CCP4 Study Weekend, edited by W. Walf, P. Evans & A. G. W. Leslie, p. 80. Daresbury Laboratory, Warrington WA4 4AD, England.
Otwinowski, Z. (1993). Proc. CCP4 Study Weekend, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Daresbury Laboratory, Warrington WA4 4AD, England.
Oxford Cryostystems (1999). *Acta Cryst.* D**55**, 724.
Peterson, M., 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.
Phillips, J. C. & Hodgson, K. O. (1980). *Acta Cryst.* A**36**, 856–864.
Prangé, T., Schiltz, M., Pernat, L., Colloc'h, N., Longhi, S., Bourget, W. & Fourme, R. (1998). *Proteins Struct. Funct. Genet.* **30**, 67–73.
Sauer, O., Schmidt, A. & Kretky, C. (1997). *J. Appl. Cryst.* **30**, 476–486.
Schiltz, M., Prangé, T. & Fourme, R. (1994). *J. Appl. Cryst.* **27**, 950–960.
Schoenborn, B. P. & Featherstone, R. M. (1967). *Adv. Pharmacol.* **5**, 1–17.
Schoenborn, B. P., Watson, H. C. & Kendrew, J. C. (1965). *Nature (London)*, **207**, 28–30.
Soltis, S. M., Stowell, M. H. B., Wiener, M. C. N., Phillips, G. N. & Rees, D. C. (1997). *J. Appl. Cryst.* **30**, 190–194.
Tilton, R. F. & Kuntz, I. D. (1982). *Biochemistry*, **21**, 6850–6857.
Tilton, R. F., Kuntz, I. D. & Petsko, G. A. (1984). *Biochemistry*, **23**, 2849–2857.
Tilton, R. F., Singh, U. C., Weiner, S. J., Connolly, M. L., Kuntz, I. D., Kollman, P. A., Max, N. & Case, D. A. (1986). *J. Mol. Biol.* **192**, 443–456.
Vernède, X. & Fontecilla-Camps, J. C. (1999). *J. Appl. Cryst.* **32**, 505–509.
Vitali, J. & Robbins, A. H. (1991). *J. Appl. Cryst.* **24**, 931–935.
Wang, B. C. (1985). *Methods Enzymol.* **115**, 90–112.
Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.
Weiss, M. S., Sicker, T. & Hilgenfeld, R. (2001). *Structure*, **9**, 771–777.
Woolfson, M. M. & Yao, J. X. (1994). *Acta Cryst.* D**50**, 7–10.