Structure determination of an FMN reductase from *Pseudomonas aeruginosa* PA01 using sulfur anomalous signal

Rakhi Agarwal,[a] Jeffrey B. Bonanno,[b] Stephen K. Burley[b] and Subramanyam Swaminathan[a]*

[a]Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA, and [b]SGX Pharmaceuticals Inc., San Diego, CA 92121, USA

Correspondence e-mail: swami@bnl.gov

The availability of high-intensity synchrotron facilities, technological advances in data-collection techniques and improved data-reduction and crystallographic software have ushered in a new era in high-throughput macromolecular crystallography. Here, the *de novo* automated crystal structure determination at 1.28 Å resolution of an NAD(P)H-dependent FMN reductase flavoprotein from *Pseudomonas aeruginosa* PA01-derived protein Q9I4D4 using the anomalous signal from an unusually small number of S atoms is reported. Although this protein lacks the flavodoxin key fingerprint motif [(T/S)XTGXT], it has been confirmed to bind flavin mononucleotide and the binding site was identified *via* X-ray crystallography. This protein contains a novel flavin mononucleotide-binding site GSLRSGSYN, which has not been previously reported. Detailed statistics pertaining to sulfur phasing and other factors contributing to structure determination are discussed. Structural comparisons of the apoenzyme and the protein complexed with flavin mononucleotide show conformational changes on cofactor binding. NADPH-dependent activity has been confirmed with biochemical assays.

1. Introduction

A major concern in high-throughput macromolecular structure determination is the high cost of sample preparation. Efficient techniques such as multiple-wavelength anomalous dispersion (MAD) methods utilize selenomethionine proteins. Preparation of selenomethionine (SeMet) proteins is frequently straightforward, typical anomalous signals are acceptably high and readily measurable and the Se absorption edge can be accessed at most synchrotron-radiation sources (0.98 Å, 12.65 keV). If, however, native protein crystals could be used directly for structure determination, it will cut the cost of SeMet protein production/crystallization and obviate the need to record diffraction data at multiple X-ray wavelengths. As most proteins contain either methionines or cysteines or both, the sulfur anomalous signal could be used to phase the X-ray data. Although the X-ray absorption edge of sulfur lies at a prohibitively long wavelength (5.02 Å), its anomalous scattering contribution (f'') at accessible X-ray wavelength (Cu $K\alpha$, 0.56 electrons; λ = 1.74 Å, 0.70 electrons) can be measured and may be useful for phasing diffraction data.

Since the structure of crambin (six Cys/45 residues) was determined (Hendrickson & Teeter, 1981), technological advancements in data-collection software and hardware and the availability of high-energy synchrotron sources have

research papers

rekindled interest in using the sulfur anomalous signal for macromolecular structure determination (Dauter *et al.*, 1999). In some cases, the sulfur anomalous signal has aided in tracing the polypeptide chain (Greenwald *et al.*, 1999). Redetermination of previously known structures using sulfur anomalous signals and *de novo* structure determinations of relatively small proteins or those crystallizing in high-symmetry space groups have been reported (Ramagopal *et al.*, 2003; Olsen *et al.*, 2004). Both high-intensity synchrotron radiation and Cu $K\alpha$ radiation from in-house facilities have been used successfully (Debreczeni, Bunkoczi, Girmann *et al.*, 2003; Debreczeni, Bunkoczi, Ma *et al.*, 2003; Debreczeni, Girmann *et al.*, 2003; Rose *et al.*, 2001). However, in all these cases at least one disulfide bond is present. Here, we report for the first time sulfur anomalous scattering structure determination for a protein entirely lacking Cys residues.

The NADH-dependent reductases are widely distributed flavin-containing proteins (Ernster & Navazio, 1957) that catalyze nicotinamide nucleotide-dependent reductions of quinones, quinoneimines, azo dyes and nitro groups. The most frequently used cofactors in such enzymatic redox reactions are pyridines (NAD, NADPH) and flavins (FMN/FAD) (Liger *et al.*, 2004). Flavin coenzymes can exist in three different redox states, characterized by different colors (oxidized form, yellow; reduced form, colorless; semiquinone form, blue). Flavin molecules can support both one- and two-electron transfer reactions. The flavodoxins are commonly small monomeric proteins that bind noncovalently to a single FMN cofactor that forms the active center (Simondsen & Tollin, 1980). The enzyme effects electron transfer by providing a micro-environment that stabilizes the FMN molecule by sandwiching it between two hydrophobic side chains, promoting the highly negative redox potential required for activity. Flavodoxins are widely involved in the electron-transfer reactions required for the metabolism of pyruvate, nitrogen and pyridine nucleotides (Liger *et al.*, 2004). Induction of quinone reductases protects against the toxic and neoplastic effects of quinines (Benson *et al.*, 1980).

The *Pseudomonas aeruginosa* PA01-derived Q9I4D4 open reading frame codes for a 20 kDa protein with unknown function and is annotated as a hypothetical protein (T1501; http://www.nysgrc.org). This protein was chosen as a target of the New York Structural Genomics Research Consortium (NYSGXRC) because no structural information was available at the time of target selection. Here, we report structure determination of T1501 using sulfur anomalous signal. To the best of our knowledge, this is the first time sulfur SAD has been successfully used to determine the structure of a 193-amino-acid protein containing no cysteine, with signals from only four S atoms (including a sulfate ion). The effects of parameters such as data redundancy, signal-to-noise ratio, X-ray beam stability and diffraction limit on the success of structure determination are discussed. We have also compared the phase information obtained from *SOLVE versus* that from *SHELX*. In addition to the native structure, we report the structure of T1501 bound to FMN, confirming that T1501 is almost certainly an NADPH-dependent FMN reductase.

2. Materials and methods

2.1. Cloning, expression and purification

The target gene for T1501 was amplified *via* PCR from *P. aeruginosa* genomic DNA with the appropriate forward (AGCGACGACATCAAGGTATTG) and reverse (CGCGC-ACCCATAGCTGCAG) primers and Taq DNA polymerase (Qiagen) using standard methods. Following gel purification, the PCR product was inserted into a pET vector modified for topoisomerase-directed cloning (Invitrogen) designed to express the protein of interest followed by a C-terminal hexahistidine tag and was transformed into TOP10 cells. The clone was confirmed for correct sequence. The expression and solubility of the protein were checked by standard procedures.

A medium-scale cell culture was grown by adding 500 ml LB medium, 25 ml 10% glucose solution, 500 µl 30 mg ml^{-1} kanamycin solution and a small amount of transformed cell glycerol stock scraping to a 2 l baffled flask and shaking overnight at 250 rev min^{-1} and 303 K. 10 ml of this mixture was added to each of six flasks containing the same culture medium (except for the glycerol stock) for large-scale expression and held at 250 rev min^{-1} and 310 K until the OD at 595 nm reached ~0.8. The cultures were then induced with 200 µl 1 *M* IPTG.

After shaking overnight at 250 rev min^{-1} and 294 K, the contents of the flasks were poured into 1 l spin bottles and spun at 6500 rev min^{-1} for 10 min. After removal of the supernatant, the pellets were collected in 50 ml conical tubes (total mass 26.0 g) and frozen at 193 K.

The pellet was resuspended in lysis buffer (35 ml per 10 g) containing 50 µl protease-inhibitor cocktail (Sigma) and 5 µl benzonase (Novagen) and subjected to repeated sonication with intervals of cooling. The lysate was then clarified by centrifugation at 38 900g for 30 min. The protein was then immobilized on Ni–NTA resin (Qiagen), placed on a drip column, washed with 25 ml buffer *A* (50 m*M* Tris–HCl pH 7.8, 500 m*M* NaCl, 10 m*M* imidazole, 10 m*M* methionine and 10% glycerol) and eluted into an Amicon concentrator (Millipore)
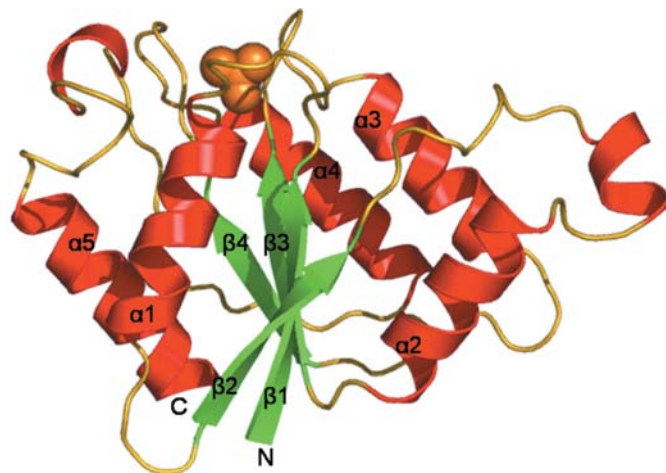
Figure 1
Ribbon representation of T1501 apoenzyme with sulfate bound at the active site. β-Sheets have been labeled $\beta1$–$\beta5$ and α-helices $\alpha1$–$\alpha5$.

with 15 ml buffer *A* containing 500 m*M* imidazole. The solution was concentrated to 6 ml, loaded onto a S200 gel-filtration column and was run off with buffer containing 10 m*M* HEPES pH 7.5, 150 m*M* NaCl, 10 m*M* methionine, 10% glycerol and 5 m*M* DTT. The protein yield was 71.9 mg.

## 2.2. Crystallization and diffraction data collection

Crystallization screening was carried out by the sitting-drop vapor-diffusion method using the TECAN Gemini crystallization robot system with Crystal Screen High Throughput (CSHT) and Index Screen High Throughput (IHT) from Hampton Research in 96-well plates. 1 µl protein solution at 10 mg ml^{-1} was mixed with 1 µl precipitant and equilibrated against 100 µl of the same precipitant at room temperature. Crystals appeared in many conditions within 1 d. Best crystals (yellow in color) were obtained in 0.2 *M* calcium acetate hydrate, 0.1 *M* sodium cacodylate pH 6.5, 18%(*w/v*) polyethylene glycol 8000 after 7–10 d. Following cryoprotection using a mother liquor containing 15% glycerol and flash-freezing in liquid nitrogen, diffraction data were collected under standard cryogenic conditions at the X12C beamline of the National Synchrotron Light Source (Brookhaven National Laboratory, Upton, NY, USA). Crystals diffracted to better than 1.3 Å and belong to space group *I*222, with unit-cell parameters $a = 63.06$, $b = 67.20$, $c = 81.54$ Å (Matthews coefficient $V_M = 2.45$ Å^3 Da^{-1}, assuming one molecule per asymmetric unit). A native data set extending to 1.28 Å resolution was obtained at $\lambda = 1.1$ Å. Since the diffraction quality was excellent, it was decided to collect sulfur SAD data at $\lambda = 1.7$ Å. In order to obtain higher redundancy, two data sets (Set1 and Set2) were collected at this wavelength for a total 720° rotation in φ. Data were reduced and scaled with *HKL*2000 (Otwinowski & Minor, 1997). The data-collection and refinement statistics are given in Table 1.

## 2.3. Structure determination of apoenzyme by sulfur-SAD

SOLVE was used to determine the 'heavy-atom' sulfur position (Terwilliger & Berendzen, 1999). The two data sets, Set1 and Set2, were examined both separately and together (Set1-2). A resolution cutoff of 2.5 Å was used for evaluation of all three data sets. Although only three methionine residues were present in the polypeptide chain, *SOLVE* yielded four sulfur positions. (The fourth sulfur site proved to be a sulfate ion that was not present in the solution used for crystallization.) Phase refinement, extension and automatic model building with *RESOLVE* were not successful. Accordingly, phase refinement was carried out with the sulfur positions using *SHARP* (de La Fortelle & Bricogne, 1997). Phases were further improved by density modification using *SOLOMON* (Collaborative Computational Project, Number 4, 1994) and then used in
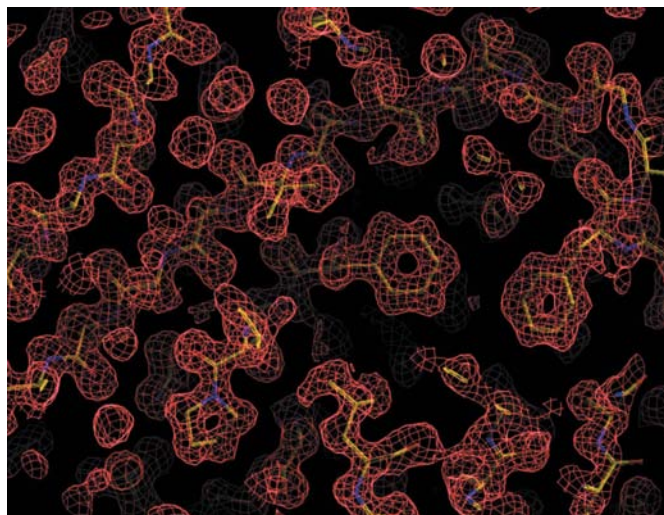
Figure 2
Composite omit electron-density map. A representative region of the apoenzyme calculated with 1.28 Å resolution data and contoured at the 1σ level.
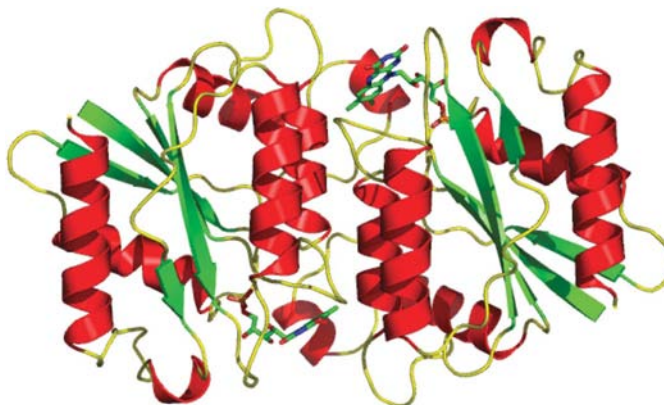
Figure 3
Ribbon representation of the biological dimer of T1501 with bound FMN (shown as a stick model).
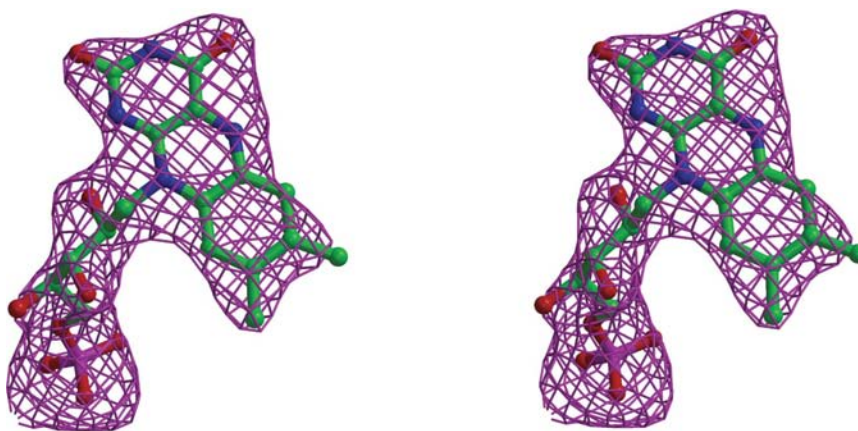
Figure 4
$|F_{obs}| - |F_{calc}|$ difference electron-density map (contoured at 2.5σ) with the final refined atomic model of FMN.

Table 1
Crystal data and refinement statistics.

Values for the highest resolution shell are given in parentheses.

	Apoenzyme		FMN complex
Space group	$I222$		$P2_12_12_1$
Unit-cell parameters (Å)	$a = 66.07$, $b = 70.39$, $c = 85.38$		$a = 61.47$, $b = 65.70$, $c = 117.80$
Method	—	SAD	MR
Wavelength (Å)	1.1	1.7	1.1
Anomalous scatterer	—	Sulfur	—
Resolution (Å)	50–1.28 (1.33–1.28)	50–1.76 (1.83–1.76)	50–2.69 (2.79–2.69)
Completeness (%)	85.3 (40.6)	84.0 (25.0)	99.9 (91.5)
No. of molecules per ASU	1	1	2
No. of reflections	44841	16798	13038
$I/\sigma(I)$	21.2 (2.2)	36.8 (6.8)	12.2 (3.1)
Redundancy	6.9	12.1	4.3
R_{merge}	0.036 (0.23)	0.036 (0.12)	0.070 (0.31)
Refinement statistics			
Resolution range (Å)	50–1.28		50–2.69
No. of reflections	44069		12497
R factor	0.21		0.21
R_{free}	0.22		0.28
No. of protein atoms	1330		2660
No. of heterogen atoms	5		62
No. of water molecules	164		160
R.m.s.d. bonds (Å)	0.044		0.066
R.m.s.d. angles (°)	1.35		1.33
Ramachandran plot			
Most favored region (%)	89.8		90.5
Additionally allowed (%)	8.8		8.2
PDB code	1rtt		1x77
Phasing statistics for S-SAD data Set2			
Phasing power	0.721		
FOM, acentric	0.235		
FOM after solvent flattening	0.947		

Table 2
Statistics of various data sets.

Values for the highest resolution shell are given in parentheses.

	Set1	Set2	Set 1-2
$I/\sigma(I)$	36.8	32.7	34.8
Redundancy	12.1 (2.8)	12.1 (2.8)	24.1 (5.2)
No. of reflections	16798	16787	16848
Resolution (Å)	1.76	1.76	1.76
Completeness (%)	83.9 (22.4)	83.9 (22.4)	84.1 (23.4)

search model using *AMoRe* (Navaza & Saludjian, 1997). The model was further refined with *CNS*. A difference electron-density map ($|F_o| − |F_c|$) showed residual density corresponding to FMN at the active site. The residual density was also seen in the corresponding composite omit map. Accordingly, FMN was built into the electron-density feature and the atomic model refined to convergence. The final atomic model is a dimer related by an NCS twofold (Fig. 3; residues 7–180, one FMN cofactor per protomer and 160 water molecules). Both FMN cofactors have well defined electron density, as shown in Fig. 4, and all φ–ψ values fall within the allowed regions of the Ramachandran diagram (Table 1).

2.5. Enzyme assays

The protein was assayed for the NADPH-dependent FMN reductase activity in a 100 µl reaction volume containing 100 µM FMN and varying concentrations (5, 7.5 and 10 µM) of enzyme in 20 mM HEPES pH 7.5. The reaction was initiated by addition of 150 µM NADPH. Oxidation of NADPH was measured by the decrease in absorption at 340 nm. A control experiment was performed without the enzyme.

3. Results and discussion

3.1. Sulfur phasing

The crystal structure of this conserved hypothetical protein from *P. aeruginosa* PA01 was determined by single-wavelength anomalous dispersion sulfur phasing using native crystals (S-SAD). Since the anomalous signal is small in general and especially so for sulfur ($f'' = 0.68$ at $\lambda = 1.7$ Å), it has to be measured accurately by increasing the signal-to-noise ratio. This can be achieved by one of several methods: using high-intensity synchrotron radiation, increasing the redundancy, using narrow oscillation ranges *etc.* (Rose *et al.*, 2001). However, increasing redundancy and radiation damage may oppose each other (especially with high-intensity radiation) and a careful compromise has to be made during data collection and processing (Garman & Owen, 2006; González, 2003). In addition, high-resolution data (>2 Å) are required for obtaining measurable signals above the noise level to obtain interpretable electron-density maps.

The most important factor to be considered before launching a phasing experiment is to calculate the Bijvoet ratio ($\Delta F_{ano}/F$) owing to sulfur. Bijvoet ratios of 0.68% may be useful in phasing diffraction data with the aid of iterative solvent-flattening methods to resolve phase ambiguity (Wang,

ARP/wARP for automatic model building (Perrakis *et al.*, 1999). 91% of the polypeptide chain was built automatically and the remainder was fitted manually using *O* (Jones *et al.*, 1991). Refinement was carried out with data extending to 1.28 Å resolution using *CNS* (Brünger *et al.*, 1998).

The final model contains 174 amino acids (residues 7–180; Fig. 1), one sulfate ion and 164 water molecules. The electron density is well defined throughout the length of the polypeptide chain (Fig. 2) and all φ–ψ values fall within the allowed regions of the Ramachandran diagram (Table 1; Laskowski *et al.*, 1993).

2.4. Structure determination of T1501 complexed with FMN

In order to confirm the presence and the type of cofactor, the protein crystals were soaked in mother liquor containing a 1 mM final concentration of FMN for 1 min. Diffraction data were collected at an X-ray wavelength of $\lambda = 1.1$ Å and the crystal diffracted to better than 2.7 Å resolution (Table 1). The crystal form used for this soaking experiment proved to be orthorhombic ($P2_12_12_1$; two molecules per asymmetric unit). Since the protein crystallizes in different forms with similar morphology under the same conditions, it was not possible to determine which crystal form was selected for the FMN soaking experiment prior to data collection. The structure was determined *via* molecular replacement with apoenzyme as the

Table 3
SOLVE versus SHELX results for sulfur anomalous phasing.

	Set1	Set2	Set1-2
Trials with *SOLVE*			
Resolution cutoff (Å)	2.5	2.5	2.5
FOM/Z score	0.25/11.0	0.30/18.3	0.23/12.1
Sulfur positions obtained	4	4	4
$\Delta\varphi$ with final model (°)	91.6	67.2	92.0
Trials with *SHELX* (*hkl2map*)			
Resolution cutoff (Å)	2.5	2.5	2.5
Sulfur positions obtained	4	4	4
Pseudo-free CC (%)	51.3	63	68.5
Weight	0.526	0.602	0.674
Map correlation	0.759	0.814	0.878
Connectivity	0.886	0.902	0.914
Contrast	0.391	0.501	0.531
$\Delta\varphi$ with final model (°)	61.4	51.5	46.2

Table 4
The calculations were repeated for various resolution cutoffs on Set2 data.

Positions of S atoms were determined using *SOLVE*; phase refinement took place using *SHARP* and model building with *ARP/wARP*.
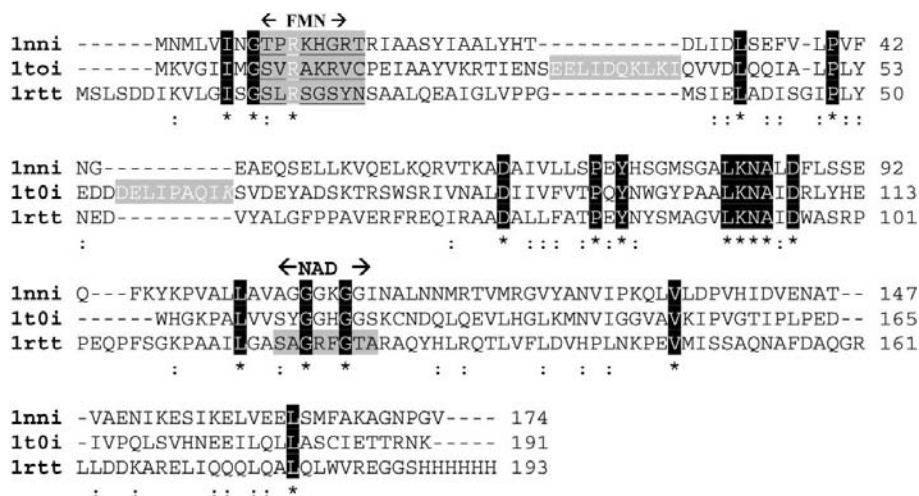
Resolution limit (Å)	No. of S atoms from *SOLVE*	No. of reflections	FOM before/ after *DM*	No. of residues built	Observation/ parameter ratio	Connec- tivity
1.76	4	16648	0.235/0.947	165	2.48	0.95
2.00	4	13484	0.308/0.929	167	2.04	0.95
2.20	4	10266	0.328/0.919	167	1.60	0.94
2.40	4	7971	0.344/0.917	56	1.16	0.71
2.50	4	7081	0.342/0.897	35	1.04	0.71
3.00	4	4163	0.335/0.903	33	0.61	0.62

1985). We chose to collect data at λ = 1.7 Å for two reasons. This energy range was easily attainable with the X12C beamline and represents the optimum for collecting sulfur diffraction data without significant adverse effects arising from absorption (Kwiatkowski *et al.*, 2000; Weiss, Sicker, Djinovic Carugo *et al.*, 2001; Weiss, Sicker & Hilgenfeld, 2001). With three methionines present in the 193-residue target protein (including the His tag), the calculated Bijvoet ratio is 0.68% (which would be 0.78% with the sulfate ion identified after structure refinement). Although the sulfur anomalous signal from methionine is typically weaker from that of super-sulfur (sulfurs of the disulfide of cystine), we proceeded with the experiment because the diffraction quality of the crystal was extremely good.

As described above, we recorded two diffraction data sets from the same crystal and used them both individually (Set1 and Set2) and merged in combination (Set1-2). All of the expected sulfur positions were obtained with each version of the diffraction data using *SOLVE*. However, the statistics from Set2 were superior to both Set1 and Set1-2, with those from Set1 being the worst (Table 2). *ARP/wARP* after *SHARP* and *SOLOMON* yielded atomic models for 160, 165 and 164 residues with Set1, Set2 and Set1-2, respectively. The best experimental phases came from Set2. $\Delta\varphi$ and electron-density map correlations with respect to the final refined model are given in Table 3.

It is remarkable that Set2 gave better results than Set1, which was recorded first. We had anticipated that radiation damage would affect the accuracy of Set2. While data redundancy improved data quality, merging of Set1 with Set2 to give Set1-2 degraded and compromised the anomalous phasing signal present in Set2. In retrospect, we believe that beam instability contributed to the poor quality of Set1, which was obtained immediately after the beam came up following a beam dump. Indeed, we detected fluctuations in scale factor in the range of 6–8% for oscillation frames recorded during the first 30 min of data collection.

Table 4 presents the results of *SOLVE* calculations repeated with Set2 using various resolution cutoffs. *SOLVE* correctly located all four S atoms irrespective of the resolution cutoff, with similar overall figures of merit (0.28–0.30).

However, automatic model building with *ARP/wARP* was not successful for resolution cutoffs worse than 2.20 Å, which may be explained by the observation-to-parameter ratio. We carried out similar calculations with all three data sets with *SHELX* using the same resolution cutoffs (Schneider & Sheldrick, 2002; Sheldrick, 2002). All *SHELXD* calculations yielded the requisite four sulfur positions, giving phasing statistics that were best for Set1-2, followed by Set2 and Set1 (Table 3). The increased data redundancy appears beneficial for Set1-2.

3.2. Automation of model building

Automated model building was attempted using *ARP/wARP* with phases obtained from *SHARP* and *SHELXE*. In both cases we used default parameters. *ARP/wARP* was successful with *SHARP* phases after density modification. The phasing statistics were very good for all three data sets with *SHELXE*. However, the pseudo-free CC was on the border-line. An ideal value for traceability is suggested to be 70%. None of the *SHELXE*-derived phase sets permitted auto-mated model building of the polypeptide chain with *ARP/wARP*, although the resulting map was excellent. The electron density was well defined but fragmented. This could be the reason that the automatic tracing of the polypeptide failed. Interestingly, the average phase difference between *SHELXE* and *SHARP* (after *DM*) was 49° for data to 1.76 Å resolution. The quality of the map suggests that the polypeptide chain could be traced in the conventional manner. *RESOLVE* failed to yield a substantive model of the polypeptide chain with Set1, Set2 or Set1-2. At best, 20–30% of the molecule was produced. With Set2, first we used 2.5 Å resolution data in *SOLVE* and then extended the phases to 1.76 Å in *RESOLVE*. The model building was not successful. Representative sections of electron-density maps computed with phases from *SHARP*, *SOLVE*, *RESOLVE* and *SHELXE* have been deposited as supplementary material[1]. The quality of the maps suggests that manual model building would have succeeded in case of *SOLVE* and *SHELXE*.


```
                   ← FMN →
1nni  ------MNMLVINCTP KHGRTRIAASYIAALYHT-----------DLIDISEFV-LPVF  42
1toi  ------MKVGIIMGSV AKRVCPEIAAYVKRTIENSEELIDQKLKIQVVDIQQIA-LPLY  53
1rtt  MSLSDDIKVLGISGSL SGSYNSAALQEAIGLVPPG---------MSIEIADISGIPLY  50
             :  * *: *                              ::* ::  :*::

1nni  NG----------EAEQSELLKVQELKQRVTKADAIVLLSPEYHSGMSGALKNALDFLSSE  92
1t0i  EDDDELIPAQIASVDEYADSKTRSWSRIVNALDIIVFVTPQYNWGYPAALKNAIDRLYHE 113
1rtt  NED---------VYALGFPPAVERFREQIRAADALLFATPEYNYSMAGVLKNAIDWASRP 101
         :                   * ::: :*:*:  *****:*

              ←NAD →
1nni  Q---FKYKPVALLAVAGGGKGGINALNNMRTVMRGVYANVIPKQLVLDPVHIDVENAT-- 147
1t0i  ------WHGKPAIVVSYGGHGGSKCNDQLQEVLHGLKMNVIGGVAVKIPVGTIPLPED-- 165
1rtt  PEQPFSGKPAAILGASAGRFGVTARAQYHLRQTLVFLDVHPLNKPEVMISSAQNAFDAQGR 161
             * : * *                : :   :  : :*

1nni  -VAENIKESIKELVEEISMFAKAGNPGV---- 174
1t0i  -IVPQLSVHNEEILQLLASCIETTRNK----- 191
1rtt  LLDDKARELIQQQLQAIQLWVREGGSHHHHHH 193
          :  :        :: ::  *
```

Figure 5
Sequence alignment of *P. aeruginosa* Q9I4D4 (PDB code 1rtt) with its structural neighbors *B. subtilis* Yhda (1nni) and yeast YLR011wp (1t0i). Strictly conserved residues are designated with white text on black background and an asterisk; the FMN-binding site is shown with underlined black letters on a grey background and labeled 'FMN'. Partially conserved residues are indicated with ':'. The two insertions in structure 1t0i are shown as white letters on a gray background. The probable NAD(P)H-binding site in 1x77 is shown on a gray background and indicated as 'NAD'.

3.3. Structure description and comparisons

T1501 has a globular α/β structure with the central five-stranded parallel β-sheet (strands numbered sequentially $\beta2$, $\beta1$, $\beta3$, $\beta4$ and $\beta5$) flanked by helices $\alpha1$ and $\alpha5$ on one side, and $\alpha2$, $\alpha3$ and $\alpha4$ on the other (Fig. 1). This target protein belongs to the flavoprotein superfamily, as it resembles flavodoxin with a three-layer $\alpha/\beta/\alpha$ structure with a five-stranded parallel β-sheet. A *DALI* search yielded several hits among the oxidoreductases. The highest match obtained was a *Bacillus subtilis* hypothetical protein Yhda (azobenzene
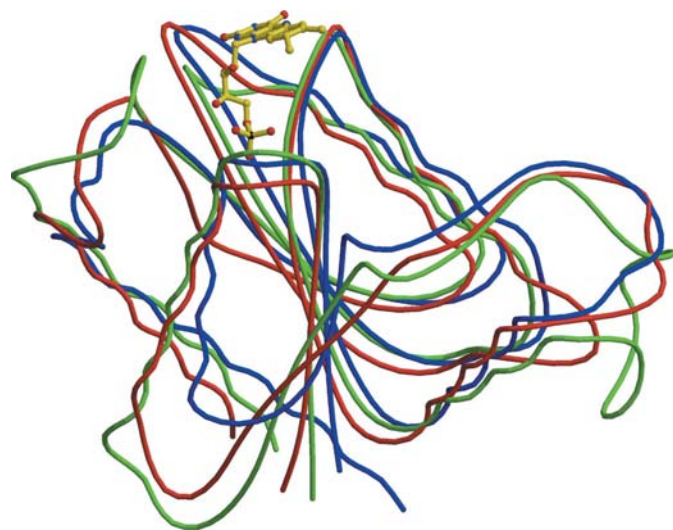
reductase, azored; PDB code 1nni) with a *Z* score of 18, followed by yeast YLR011wp (1t0i), rubredoxin:oxygen oxidoreductase (1e5d), quinone reductase (1d4a) and flavodoxin bound to FMN (1rcf). Sequence alignment of 1nni, 1t0i and 1rtt (T1501, apoenzyme) was performed using *ClustalW* (Thompson *et al.*, 1994) as shown in Fig. 5. Structural alignment of 1nni, 1t0i and 1x77 (T1501, FMN complex) monomers is shown in Fig. 6. The root-mean-square deviation (r.m.s.d.) between 1x77 and 1nni is 1.21 Å for 131 C^{α}-atom pairs and that between 1x77 and 1t0i is 1.6 Å for 117 C^{α}-atom pairs. All three protomers share the flavodoxin fold, with the only significant differences arising from the presence of two extended loops in 1t0i owing to insertions of ten and nine residues (Fig. 5). This difference in loop lengths may reflect differences in substrate specificity/binding. 1t0i is a ferric ion reductase (Liger *et al.*, 2004).

3.4. The FMN-binding site

Sequence analyses of the target protein suggested similarity to NADPH-dependent reductases and NADPH:quinone oxidoreductases. Within the active site of the apoenzyme we found a bound sulfate ion which occupied a phosphate-binding site, as shown for ribityl phosphate in Fig. 7. The electron-density map for the yellow target protein (suggesting the presence of an oxidized flavin ring) failed to show any evidence of the expected FMN cofactor, which may reflect loss during crystal growth at room temperature in the absence of added ligand. Structural analysis of the apoenzyme docu-
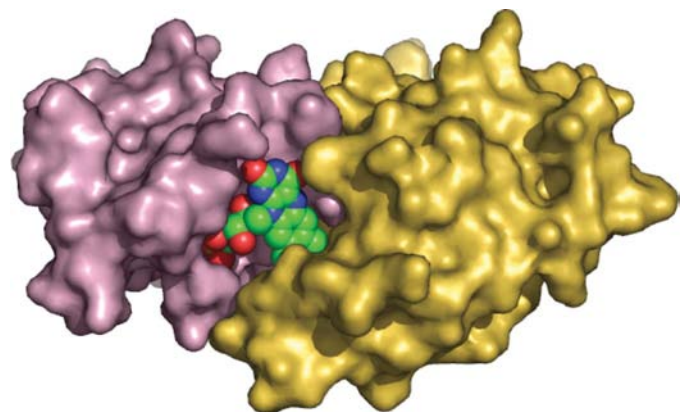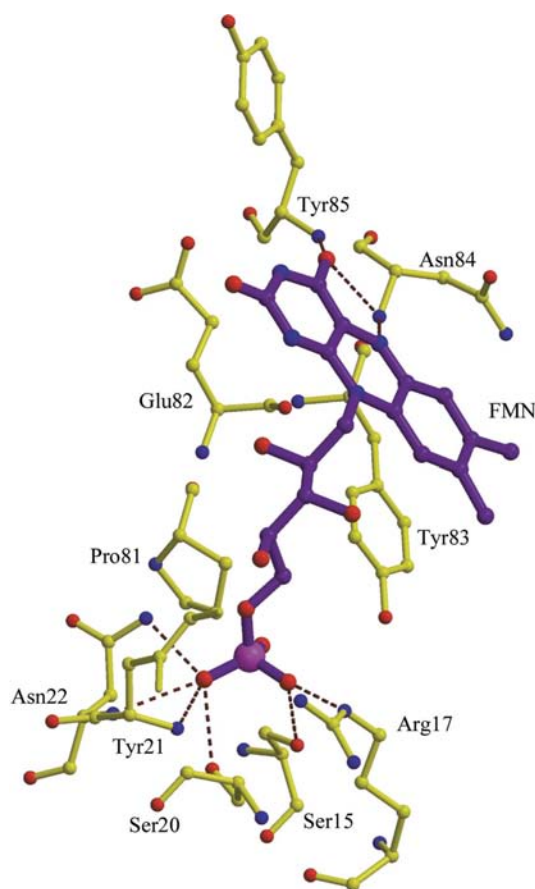

Figure 6
Structural alignment of monomers of 1x77, 1nni and 1t0i shown in red, blue and green, respectively. The FMN molecule in all three is shown only once in yellow. *MOLSCRIPT* and *RASTER3D* were used to produce this figure (Kraulis, 1991).

Figure 7
A surface drawing of the target T1501 dimer with FMN, shown as a hard-sphere representation. The isoalloxazine ring of FMN interacting with both monomers is exposed to solvent, where it can interact with NAD(P)H and substrate. Figs. 1, 3 and 7 were prepared using *PyMOL* (DeLano, 2002).

mented the presence of a large pocket which could accommodate either an FMN or an FAD. An apoenzyme crystal-soaking experiment revealed the presence of FMN in the pocket, with sulfate ion replaced by the ribityl phosphate ion. The cofactor binds to the C-terminal end of the central β-sheet within a positively charged pocket (data not shown). A molecular surface drawing of the dimer with bound FMNs is shown in Fig. 7. The phosphoribityl group of the FMN is deeply buried, although the isoalloxazine ring is partially exposed to the solvent. FMN makes interactions with both protomers.

It is remarkable that this target protein does not bear the commonly found flavodoxin key fingerprint motif (T/S*XT*-G*X*T), yet still binds the FMN phosphate (*via* the GSLR-SGSYN loop; amino acids 14–22). A *BLAST* search shows that Ser15, Arg17 and Ser20 are conserved in this family of proteins, suggesting they will make similar interactions with FMN. The sulfate group in the apoenzyme and the FMN phosphate participate in hydrogen bonds with the side chains of Ser15, Arg17, Ser20 and Asn22 (Fig. 8). The isoalloxaxine ring O4 interacts with the backbone amino groups of Asn84 and Tyr85 and O2 interacts with a water molecule, which in turn interacts with the backbone amino groups of Arg120 and Phe121. N5 of the isoalloxazine ring interacts with Asn84 N.

The ribityl O2* of FMN interacts with a water molecule, which is stabilized by Glu82 N and Ala116 O. The major stabilization of the FMN molecule occurs *via* the ribityl phosphate and the isoalloxazine ring. N5 and N2 of FMN are the groups which are the probable sites for accepting electrons from NADPH.

3.5. Structural changes in the enzyme owing to cofactor binding

The r.m.s.d. between 1rtt and a 1x77 monomer is ∼0.64 Å; the r.m.s.d. between the two monomers of 1x77 is ∼0.60 Å. In our apoenzyme structure, the phosphate-binding site is well formed and occupied by a sulfate anion, as seen for apo flavodoxin from *Anabaena* (Genzor *et al.*, 1996). However, in the apoenzyme the isoalloxine ring and ribityl chain-binding sites were constricted and are partially occupied by the side chains of Tyr85, Arg125 and Tyr21. The side chains of Tyr21, Arg125 and Tyr128 reorient upon FMN binding. Movement of C^{α} atoms was observed for resides 119–122 (the C^{α}-position shifts for of 119, 121 and 122 are 2.67, 1.69 and 3.12 Å, respectively).

3.6. T1501 forms a dimer in both structures

Crystals of both the apoenzyme (*I*222 form) and holo-enzyme in complex with FMN ($P2_{1}2_{1}2_{1}$ form) revealed pairs of molecules forming the same kind of dimer, related by twofold symmetry. While the twofold symmetry is crystallographic symmetry parallel to the *c* axis in the *I*222 form, it is *via* an NCS twofold in the $P2_{1}2_{1}2_{1}$ form. Each dimer has two equivalent catalytic sites, which are formed by parts of both monomers, present at the dimer interface (Fig. 3). The solvent-accessible surface area buried on dimer formation is ∼2560 Å^2 (>14%), suggesting that the crystallographically observed dimer may be biologically relevant. This observation is

Figure 8
FMN interaction at the active site of T1501. FMN is shown in magenta and protein residues in yellow; dashed lines represent hydrogen bonds (this figure was prepared with *MOLSCRIPT* and *RASTER3D*; Kraulis, 1991).
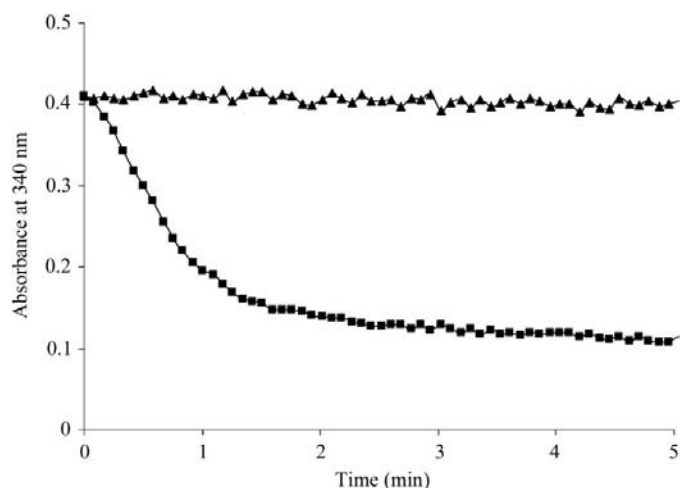

Figure 9
NAD(P)H-FMN reductase activity of T1501. 150 μ*M* NADPH was added in a mixture of 10 μm enzyme containing 100 μ*M* FMN at room temperature and oxidization of NAD(P)H was followed and recorded at 340 nm for 5 min. With 5 and 7.5 m*M* enzyme concentrations, the absorbance decreased to 0.3 and 0.25, respectively (data not shown). Triangles and squares represent the control and the actual experiment, respectively.

analogous to that for the recently reported structure of NADPH-dependent FMN reductase from yeast YLR011wp (PDB code 1t0i). Such dimers are thought to be uncommon in most flavodoxin-related proteins (Liger *et al.*, 2004). However, some of the flavodoxins, such as *Saccharomyces cerevisiae* Ycp4 (Toda *et al.*, 1992) and *Desulfovibrio gigas* rubredoxin-oxygen reductase (Frazao *et al.*, 2000) exist as dimers.

3.7. Catalysis: T1501 is an NADPH-dependent FMN reductase

Our target crystal structure of the holoenzyme confirms that the protein is an FMN-containing enzyme belonging to the flavodoxin family. Biochemical assays have documented the T1501-dependent oxidation of NADPH and NADPH dependence (Fig. 9).

FMN reductases catalyze a two-step reaction.

(1) (*a*) E-FMN + NADPH $\longleftrightarrow$ E-FMN-NADPH

    (*b*) E-FMN-NADPH + H$^+$ $\longleftrightarrow$ E-FMNH$_2$ + NAD(P)$^+$

(2) (*a*) E-FMNH$_2$ + substrate $\longleftrightarrow$ E-FMNH$_2$-substrate

    (*b*) E-FMNH$_2$-substrate $\longleftrightarrow$ E-FMN + product.

The catalytic site of the protein can be divided into three functional regions: one for FMN binding, a second for interaction with the adenine ribose portion of NADPH and the third for binding either the nicotinamide ring of NADPH or the substrate. Overlap of hydride donor and acceptor sites is a common feature of oxidoreductases with a ping-pong kinetic mechanism. The structure of rat liver quinone reductase-bound NADP$^+$ provides some useful insights into NADPH binding and electron transfer (PDB code 1qrd; Li *et al.*, 1995). In this case, the nicotinamide of NADP$^+$ and ring C of the isoalloxazine make a π–π stacking interaction, with an average interplanar distance of about 3.4 Å.

Large C$^\alpha$ displacements of about 2.7 Å in the loop responsible for binding NAD(P)H have been observed in some oxidoreductases (Faig *et al.*, 2000). This finding probably explains our inability to prepare cocrystals of our target protein with NADPH *via* simple crystal soaking, as the crystals crack owing to the large movement of this loop region. As the protein is NADPH-dependent, it has to bind to the enzyme first to donate electrons to the N2/N5 (of isoalloxazine) through parallel stacking of nicotinamide on the isoalloxazine ring, probably similar to the case in 1qrd. However, the binding of the remaining adenosine part of the NAD may be different from that of 1qrd since the NADPH adenosine part has been found to bind in various conformations to specific proteins. The sequence TTGGSGS has been suggested to be the part of the NADPH-binding site (Liu *et al.*, 1989; Ma *et al.*, 1992; Prochaska, 1988). A similar though not identical sequence motif has been identified in this family and the putative binding site is shown in Fig. 5 as the NAD site (probable NADPH-binding site). This site is in close proximity to the isoalloxazine ring of FMN in the structure of 1x77 and probably plays a role in interacting with the nicotinamide ring. The variation of residues present at the interface of the probable nicotinamide ring with the isoalloxazine ring-binding domain in various oxidoreductases possibly serves as the basis of substrate selectivity.

Thus far, we have not attempted to identify biological substrate(s) for T1501. However, its close structural similarity to 1nni, a probable azobezene reductase, suggests that T1501 acts on compound(s) resembling azobenzene.

4. Conclusion

In summary, we have successfully employed a less common strategy for structure determination using long-wavelength radiation for X-ray diffraction measurements of anomalous signals from S atoms, with the attendant needs for a high signal-to-noise ratio and high data redundancy. The homo-dimeric target protein T1501 proved to be a member of the flavodoxin family that uses an unusual GSLRSGSYN sequence for FMN binding instead of the canonical flavodoxin key fingerprint motif (T/S)XTGXT. Initial functional characterization has documented that the protein is an NADPH-dependent reductase which utilizes FMN as a cofactor. Scientists interested in obtaining reagents with which to perform additional functional and mechanistic studies can request them from the senior author at swami@bnl.gov.

References

Benson, A. M., Hunkeler, M. J. & Talalay, P. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 5216–5220.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszwewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.
Debreczeni, J. É., Bunkoczi, G., Girmann, B. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 393–395.
Debreczeni, J. É., Bunkoczi, G., Ma, Q., Blaser, H. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 688–696.
Debreczeni, J. É., Girmann, B., Zeeck, A., Kratzner, R. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 2125–2132.
DeLano, W. L. (2002). *The PyMOL User's Manual.* DeLano Scientific, San Carlos, CA, USA.
Ernster, L. & Navazio, F. (1957). *Biochim. Biophys. Acta*, **26**, 408–415.
Faig, M., Bianchet, M. A., Talalay, P., Chen, S., Winski, S., Ross, D. & Amzel, L. M. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3177–3182.
Frazao, C., Silva, G., Gomes, C. M., Matias, P., Coelho, R., Sieker, L., Macedo, S., Liu, M. Y., Oliveira, S., Teixeira, M., Xavier, A. V., Rodrigues-Pousada, C., Carrondo, M. A. & Le Gall, J. (2000). *Nature Struct. Biol.* **7**, 1041–1045.
Garman, E. F. & Owen, R. L. (2006). *Acta Cryst.* D**62**, 32–47.
Genzor, C. G., Perales-Alcon, A., Sancho, J. & Romero, A. (1996). *Nature Struct. Biol.* **3**, 329–332.
González, A. (2003). *Acta Cryst.* D**59**, 1935–1942.
Greenwald, J., Fischer, W. H., Vale, W. W. & Choe, S. (1999). *Nature Struct. Biol.* **6**, 18–22.

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

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

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

Kwiatkowski, W., Noel, J. P. & Choe, S. (2000). *J. Appl. Cryst.* **33**, 876–881.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–493.

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

Li, R., Bianchet, M. A., Talalay, P. & Amzel, L. M. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 8846–8850.

Liger, D., Graille, M., Zhou, C. Z., Leulliot, N., Quevillon-Cheruel, S., Blondeau, K., Janin, J. & van Tilbeurgh, H. (2004). *J. Biol. Chem.* **279**, 34890–34897.

Liu, X. F., Yuan, H., Haniu, M., Iyanagi, T., Shively, J. E. & Chen, S. A. (1989). *Mol. Pharmacol.* **35**, 818–822.

Ma, Q., Cui, K., Xiao, F., Lu, A. Y. & Yang, C. S. (1992). *J. Biol. Chem.* **279**, 34890–34897.

Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.

Olsen, J. G., Flensburg, C., Olsen, O., Bricogne, G. & Henriksen, A. (2004). *Acta Cryst.* D**60**, 250–255.

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

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Prochaska, H. J. (1988). *Arch. Biochem. Biophys.* **267**, 529–538.

Ramagopal, U. A., Dauter, M. & Dauter, Z. (2003). *Acta Cryst.* D**59**, 1020–1027.

Rose, J. P., Wu, C.-K., Liu, Z.-J., Newton, G. & Wang, B.-C. (2001). *Rigaku J.* **18**, 4–12.

Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.

Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.

Simondsen, R. P. & Tollin, G. (1980). *Mol. Cell. Biochem.* **33**, 13–24.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Toda, T., Shimanuki, M., Saka, Y., Yamano, H., Adachi, Y., Shirakawa, M., Kyogoku, Y. & Yanagida, M. (1992). *Mol. Cell. Biol.* **12**, 5474–5484.

Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–106.

Weiss, M. S., Sicker, T., Djinovic Carugo, K. & Hilgenfeld, R. (2001). *Acta Cryst.* D**57**, 689–695.

Weiss, M. S., Sicker, T. & Hilgenfeld, R. (2001). *Structure*, **9**, 771–777.