Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structures of two superoxide dismutases from *Bacillus anthracis* reveal a novel active centre

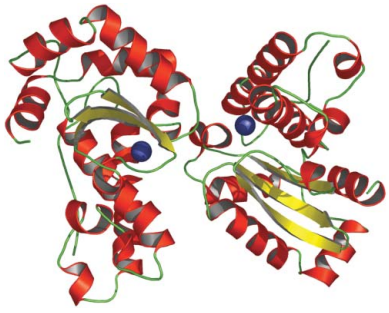
Ian W. Boucher, Anne K. Kalliomaa, Vladimir M. Levdikov, Elena V. Blagova, Mark J. Fogg, James A. Brannigan,* Keith S. Wilson and Anthony J. Wilkinson

Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5YW, England

Correspondence e-mail: jab@ysbl.york.ac.uk

Received 19 April 2005
Accepted 1 June 2005
Online 15 June 2005

PDB References: SodA1, 1xuq, r1xuqsf; SodA2, 1xre, r1xresf.



The BA4499 and BA5696 genes of *Bacillus anthracis* encode proteins homologous to manganese superoxide dismutase, suggesting that this organism has an expanded repertoire of antioxidant proteins. Differences in metal specificity and quaternary structure between the dismutases of prokaryotes and higher eukaryotes may be exploited in the development of therapeutic antibacterial compounds. Here, the crystal structure of two Mn superoxide dismutases from *B. anthracis* solved to high resolution are reported. Comparison of their structures reveals that a highly conserved residue near the active centre is substituted in one of the proteins and that this is a characteristic feature of superoxide dismutases from the *B. cereus*/*B. anthracis*/*B. thuringiensis* group of organisms.

1. Introduction

Bacillus anthracis is a Gram-positive rod-shaped spore-forming bacterium. It is the causative agent of anthrax, which is an acute, infectious and normally lethal disease. Protein structures from this organism are being determined as part of a structural genomics programme (SPINE). Amongst initial candidates for crystallographic analysis are enzymes involved in antioxidant defence. The BA4499 (*SodA1*) and BA5696 (*SodA2*) genes (http://www.tigr.org) of *B. anthracis* encode proteins with 58 and 50% amino-acid sequence identity, respectively, to the Mn superoxide dismutase from *Escherichia coli*. The superoxide radical anion is potentially harmful to cells as it oxidizes and degrades biological molecules such as lipids and proteins (Beyer *et al.*, 1991). Superoxide dismutases counter this threat by catalysing the destruction of the radical species in a dismutation reaction which produces hydrogen peroxide and molecular oxygen (Holm *et al.*, 1996),

$$M^{3+} + O_2^{\cdot -} \rightarrow M^{2+} + O_2$$
$$M^{2+} + O_2^{\cdot -} + 2H^+ \rightarrow M^{3+} + H_2O_2.$$

The hydrogen peroxide species is subsequently broken down to H_2O and O_2 by a peroxidase such as catalase. Four classes of superoxide dismutase have been defined according to their metal cofactor, which can be either iron, manganese, copper/zinc or nickel (Miller, 2004). Fe and Mn dismutases share a highly conserved sequence and a common fold (Stallings *et al.*, 1984). The human genome encodes two superoxide dismutases: the first is a dimeric Cu/Zn-dependent enzyme found in the cytosol, while the second is a tetrameric Mn-dependent enzyme found in the mitochondrion. Differences in metal-dependent activity and oligomerization may provide routes into antibacterial therapies.

2. Materials and methods

2.1. Cloning, expression and purification

The coding sequences of BA4499 and BA5696 were amplified by the polymerase chain reaction (PCR) from *B. anthracis* genomic DNA using KOD Hot Start DNA polymerase (Novagen) and

Table 1
X-ray data-collection and refinement statistics.

Values in parentheses refer to the outer shell (1.86–1.80 Å; for refinement, 1.847–1.80 Å).

	SodA1 (BA4499)	SodA2 (BA5696)
Data collection		
X-ray source	SRS beamline PX9.6	SRS beamline PX9.6
Wavelength (Å)	0.87	0.87
Data-collection temperature (K)	100	100
Resolution range (Å)	50–1.80	50–1.80
Space group	$P12_11$	$P12_11$
Unit-cell parameters (Å,°)	a = 60.88, b = 60.44, c = 61.98, α = 90.0, β = 106.30, γ = 90.0	a = 54.84, b = 69.87, c = 56.57, α = 90.0, β = 106.85, γ = 90.0
Matthews coefficient (Å^3 Da^{-1})	2.4	2.2
Solvent content (%)	48.7	42.6
No. of unique reflections	39919 (3607)	37730 (3554)
Completeness (%)	98.8 (89.7)	98.8 (94.0)
Redundancy	3.9 (3.2)	3.6 (3.0)
$I/\sigma(I)$	16.3 (1.5)	26.3 (1.94)
R_{merge}† (%)	8.4 (74.8)	7.1 (59.7)
Refinement and model statistics		
R factor‡	0.172 (0.253)	0.162 (0.206)
R_{free}§	0.228 (0.299)	0.219 (0.318)
Reflections (working/free)	37842/2001 (2577/162)	35719/1879 (2526/144)
Molecules per AU	2	2
No. of protein non-H atoms	3195	3284
No. of water molecules	475	377
R.m.s. deviation from target¶		
Bond lengths (Å)	0.017	0.015
Bond angles (°)	1.449	1.444
Average B factor (Å^2)	24.88	27.49
Ramachandran plot††	91.0/7.6/1.5/0.0 (Gly111)	91.5/6.8/1.1/0.6 (Asn111)

† $R_{merge} = \sum_{hkl}\sum_i |I_i - \langle I\rangle| / \sum_{hkl}\sum_i \langle I\rangle$, where I_i is the intensity of the ith measurement of a reflection with indexes hkl and $\langle I\rangle$ is the statistically weighted average reflection intensity. ‡ R factor $= \sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively. § R free is the R factor calculated with 5% of the reflections chosen at random and omitted from refinement. ¶ Root-mean-square deviation of bond lengths and bond angles from ideal geometry. †† Percentage of residues in most-favoured/additionally allowed/generously allowed/disallowed regions of the Ramachandran plot according to *PROCHECK* (Laskowski *et al.*, 1993).

complementary gene-specific primers to which were appended sequences to facilitate ligation-independent cloning (LIC; Aslanidis & De Jong, 1990). For LIC, the PCR amplification products were treated with T4 DNA polymerase in the presence of dATP to generate 5′ single-stranded overhangs at both ends of the fragment through the enzyme's combined 3′–5′ exonuclease and DNA polymerase activities. Complementary 5′ single-stranded overhangs were generated in the vector pET-YSBLIC by cleavage with the restriction endonuclease *Bse*RI and treatment with T4 DNA polymerase in the presence of dTTP. This plasmid is a pET28a derivative that has been adapted for LIC and encodes an N-terminal hexahistidine tag to the cloned gene. The vector and PCR products were mixed and used to transform *Escherichia coli* NovaBlue cells (Novagen). Recombinants were selected through kanamycin resistance and colony PCR, using T7 promoter and gene-specific oligonucleotide primers to confirm the presence of the gene insert, the sequence of which was then verified. The pET-YSBLIC*sodA* clones were transformed into *E. coli* BL21(DE3) for protein overexpression.

Two 500 ml cultures of cells were grown with shaking at 310 K in Luria–Bertani broth containing 30 µg ml^{-1} kanamycin to an OD_{600} of 0.6. Expression of *B. anthracis* SodA1 and SodA2, each fused to an amino-terminal sequence (MGSSHHHHHH) encoded by pET-YSBLIC, was then induced by the addition of isopropyl-β-D-thio-galactopyranoside to a final concentration of 1 m*M*. Cells were harvested by centrifugation and resuspended in 20 m*M* Na_2HPO_4 buffer pH 7.5 containing 500 m*M* NaCl and 10 m*M* imidazole (buffer *A*). The cells were lysed by sonication on ice and the cell-free supernatant was collected by centrifugation at 36 890*g* in a Sorvall SS-34 rotor (DuPont). Two 5 ml Hi-Trap chelating columns (Amersham Biosciences), previously charged with nickel and equilibrated with buffer *A*, were attached to an Äkta Explorer three-dimensional purification system (Amersham Biosciences). The soluble fractions were applied to the columns and washed with 25 ml buffer *A* followed by 50 ml buffer *A* containing 70 m*M* imidazole. The proteins were eluted with buffer *A* containing 500 m*M* imidazole and automatically passed to a Superdex G-200 gel-filtration column (Amersham Pharmacia) pre-equilibrated in 50 m*M* Tris–HCl pH 7.5 and 150 m*M* NaCl for size-exclusion chromatography. Fractions containing SodA1 and SodA2 were identified by SDS–PAGE, pooled and concentrated to 46 mg ml^{-1} (SodA1) and 26 mg ml^{-1} (SodA2) by centrifugal ultrafiltration (Millipore).

2.2. Protein characterization

Purified SodA1 and SodA2 proteins were loaded individually onto a 10% native polyacrylamide gel. The gel was run at 100 V for 3 h and soaked in riboflavin and NADPH to generate superoxide radicals. Nitro Blue Tetrazolium (NBT) dye was used to stain the gel. NBT reacts with superoxide to form a blue precipitate when fixed in a gel. Any proteins on the gel that exhibit Sod activity will therefore give

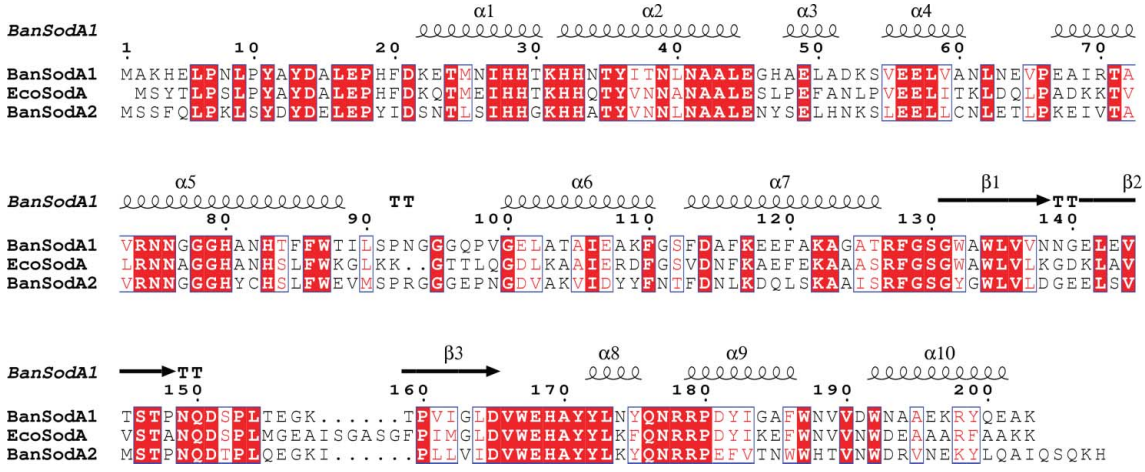


Figure 1
Protein sequence alignment of *B. anthracis* SodA1 and SodA2 and *E. coli* SodA, with the secondary-structure elements of SodA1 superposed. The sequence-alignment figure was generated by *ESPript* (Gouet *et al.*, 1999). Strictly conserved residues are highlighted with red boxes and conservative substitutions are boxed.

rise to a white band against a blue background (Flohe & Otting, 1984). Recombinant SodA1 and SodA2 were both visible as white bands on the gel.

2.3. Crystallization

Automated crystallization screening was carried out using sitting-drop vapour diffusion in 96-well plates (Greiner) using a Mosquito nanolitre pipetting robot (TTP labtech) and screens from Hampton Research (Crystal Screens I, II and Index) and Clear Strategy Screens (CSSI and CSSII; Brzozowski & Walton, 2001) at pH 7.5 and 6.5. Each crystallization drop contained 150 nl protein solution and 150 nl reservoir solution. SodA1 crystallized from a reservoir solution containing 2 *M* sodium acetate trihydrate, 0.1 *M* Tris–HCl pH 8.5 and 30%(*w*/*v*) PEG 4000. The initial conditions for crystallization of SodA2 were optimized manually to a final condition of 34%(*w*/*v*) PEG 3350, 0.2 *M* $MgCl_2$ and 0.1 *M* Tris–HCl pH 8.0.

2.4. Data collection and processing

Diffraction data were collected on beamline PX 9.6 at the Daresbury Synchrotron Radiation Source (SRS). In each case, a single crystal was mounted in a rayon loop, transferred to a solution of the crystallization mother liquor and then rapidly cooled in liquid nitrogen. Diffraction data were recorded on an ADSC Quantum 4 CCD detector. Data were processed and reduced using the *HKL*2000 software suite (*DENZO* and *SCALEPACK*). Data-collection statistics are presented in Table 1.

2.5. Structure solution and refinement

The structures of SodA1 and SodA2 were determined by molecular replacement using the program *MOLREP* (Vagin & Teplyakov, 1997) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Data in the resolution range 50–1.8 Å were used in both rotation and translation calculations. Refinement calculations were performed using *REFMAC*5 (Murshudov *et al.*, 1997) interspersed with sessions of manual modelling using *QUANTA* (Accelrys Inc., San Diego, CA, USA). Refinement and model statistics are presented in Table 1. The quality of the final model was scrutinized using *PROCHECK* (Laskowski *et al.*, 1993) and *SFCHECK* (Vaguine *et al.*, 1999). Atomic coordinates for the refined *B. anthracis* SodA1 and SodA2 models together with experimental structure factors have been deposited as PDB codes 1xuq and 1xre, respectively.

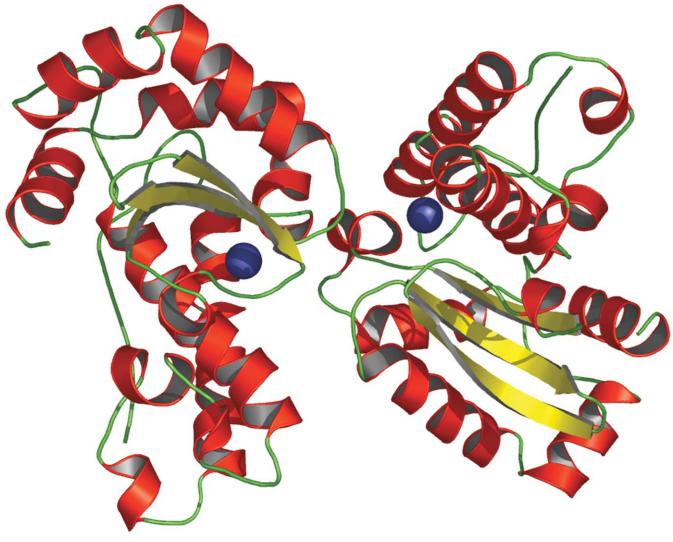


Figure 2
Cartoon representation of the SodA1 structure. Helices are coloured red and strands are in yellow. Mn ions are shown as blue spheres. This figure was drawn with *PyMol* (DeLano, 2002).

Table 2
Bond lengths of residues bonding directly to Mn.

The bond lengths are calculated as an average from chains *A* and *B*.

	Bond length (Å)	
Mn-coordinating residue	SodA1	SodA2
His28	2.04	2.09
His83	2.06	2.05
His169	2.08	2.07
Asp165	1.93	1.92
Water	2.13	2.24

3. Results and discussion

The structures of SodA1 and SodA2 were determined by molecular replacement using the coordinate set for the *B. halodenitrificans* orthologue (70 and 55% amino-acid sequence identity, respectively; PDB code 1jr9; Liao *et al.*, 2002) as a search model. Data-collection, refinement and model-building statistics are summarized in Table 1. The refined models consist of two protein chains *A* and *B* comprising residues 3–202 for SodA1 and 3–203 for SodA2, one manganese ion in each protein chain and a total of 475 and 377 water molecules, respectively. The Matthews coefficient (V_M) for the crystals is 2.4 $Å^3\ Da^{-1}$ (SodA1) and 2.2 $Å^3\ Da^{-1}$ (SodA2) and the estimated solvent content is 48.7 and 42.6%, respectively. The Ramachandran plot shows that 91% (A1) and 91.5% (A2) of residues are in the most favoured regions, with 9% (A1) and 7.9% (A2) in additionally allowed regions. In SodA2, residue Asn111 is in a disallowed region in both chains. Asn111 is situated on an exposed loop and makes interactions with crystallographic waters in the crystal structure. In SodA1 and *E. coli* SodA the equivalent residue is a glycine (Fig. 1).

All the observed X-ray data in the resolution range 50.0–1.8 Å were used in the structure solution and refinement. The maximum-likelihood program *REFMAC*5 (Murshudov *et al.*, 1997) used for refinement leaves the refined models with *R* factors of 0.172 (A1) and 0.162 (A2) and R_{free} factors of 0.228 (A1) and 0.219 (A2), respectively. Superposition of all C^{α} atoms of subunit *A* onto subunit *B* within the asymmetric units of SodA1 and SodA2 gives r.m.s.d. values of 0.27 Å in both cases. Subunits from SodA1 and SodA2 superpose onto each other with r.m.s.d. values in the range 0.75–0.90 Å. SodA1

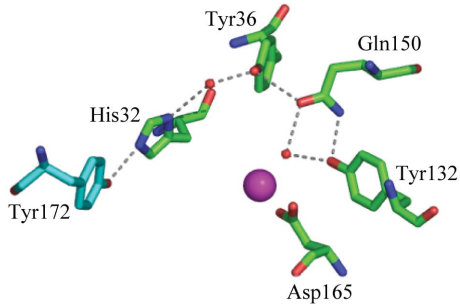


Figure 3
Hydrogen-bonding network in SodA2. Gln150 hydrogen bonds through Tyr36 and His32 to a tyrosine from the other chain (cyan). In the case of SodA1, Gln150 makes a further hydrogen bond with Trp132 (see text), whereas in SodA2, as shown here, this interaction is with Tyr132. Mn is represented by a purple sphere and waters are in red. This figure was drawn in *PyMol* (DeLano, 2002).

and *E. coli* Mn SodA can be superposed to give an r.m.s.d. of 0.73 Å, whereas SodA2 gives an r.m.s.d. of 0.93 Å.

The *B. anthracis* SodA1 and SodA2 monomers (Fig. 2) consist of two domains: an α-helical domain linked to an α/β domain. The monomers form a dimer with the two active-site Mn ions 18 Å apart. In each case the first 16 residues participate in crystal packing and appear to lack any secondary structure. The next 74 residues form the α-helical domain, the dominant feature of which is two long anti-parallel helices connected by an extended loop. The remainder of the residues form the α/β domain. The residues that form the dimer interface are found in the C-terminal domain. As in other dimeric Sod proteins, a relatively small accessible surface area becomes buried in the dimer interface: 9.13% for SodA1 and 8.58% SodA2. Two of the Mn-coordinating residues come from the α-helical domain (His28 and His83); the remaining two (His169 and Asp165) arise from the α/β domain (Table 2).

The Mn ions in both subunits of SodA2 and in the *B* subunit of SodA1 are coordinated in a trigonal bipyramidal fashion, the fifth ligand being a water or hydroxide molecule. In the *A* subunit of SodA1, a sixth ligand (water or hydroxide) appears to have been trapped in the superoxide-binding site. This coordination site is considered to be a fast-exchange position where the ligand rapidly binds and dissociates from the metal. Trapping of a sixth ligand has been reported in diffraction experiments carried out at 100 K with crystals grown at pH 8.5 (Borgstahl *et al.*, 2000). The oxygen of this hydroxide or water is 2.24 Å away from the metal. The geometry of the six-coordinate active site is distorted octahedral. The sixth ligand is within hydrogen-bonding distance of two of the Mn-coordinating histidines (His83 and His169).

Differences in the second-sphere hydrogen-bonding network arise between SodA1 and SodA2. The classical hydrogen-bonding network in Mn dismutases involves a hydrogen bond from the metal-coordinating water to a Gln residue, which subsequently hydrogen bonds to a Tyr hydroxyl group. A conserved water molecule then mediates a hydrogen bond between this tyrosine and a histidine, which in turn forms a hydrogen bond with a tyrosine from the other subunit in the dimer (Ramilo *et al.*, 1999). This network is involved in the supposed 'ping-pong' method of catalysis which relies on communication between the two metal sites. The network is present in both SodA1 and SodA2, yet a subtle difference emerges (Fig. 3). Gln150 hydrogen bonds to the metal-coordinating water and to Tyr36 as part of the expected network. However, its other hydrogen-bonding partner differs between SodA1 and SodA2. In SodA1 it forms a hydrogen bond with Trp132, whereas in SodA2 it forms a hydrogen bond with a tyrosine at the same position. The tyrosine in SodA2 is able to make an additional hydrogen bond to the metal-coordinating water/hydroxide. The vast majority of both Fe and Mn Sods have a tryptophan at this position, so the ocurrence of a tyrosine in SodA2 is unusual. Such a difference in the second-sphere hydrogen-bonding network may have subtle effects on the metal's redox potential and affect the rate of catalysis (Edwards *et al.*, 2001). Sequence searches suggest that in the many hundreds of Sod sequences known this Trp-to-Tyr substitution is rare. It is seen in the atypical cyanobacterium *Gloeobacter violaceus* and some pathogenic strains of *Vibrio*, *e.g. V. cholerae*. It is also characteristic of the *B. cereus* group of organisms, which includes *B. thuringiensis* and *B. anthracis*. There appears to have been an expansion in the number of proteins involved in antioxidation in these organisms. As well as the two MnSod proteins described here, *B. anthracis* encodes a 300-residue protein (BA1489) that has 45% sequence identity to SodA1 over its C-terminal 205 residues. This enhanced repertoire allows specialization of function, which may be reflected in the structure of SodA2. Although the function of this protein is at present unknown, it may be involved in post-translational modifications of outer coat proteins in the spore, *e.g.* cross-linking of the Tyr-rich CotG (Henriques *et al.*, 1998). Alternatively, it may play a protective role against oxidative killing during germination of endospores internalized within phagocytes, thus contributing to both viability and virulence (Krishnakumar *et al.*, 2004; Baillie *et al.*, 2005).

We would like to thank the SRS Daresbury for synchrotron beam time. The work described here was funded by the European Commission as SPINE contract No. QLG2-CT-2002-00988 under the RTD programme 'Quality of Life and Management of Living Resources'. IWB and VML are supported by the BBSRC and JAB by the Wellcome Trust.

References

Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
Baillie, L., Hibbs, S., Tsai, P., Cao, G.-L. & Rosen, G. M. (2005). *FEMS Microbiol. Lett.* **245**, 33–38.
Beyer, W., Imlay, J. & Fridovich, I. (1991). *Prog. Nucleic Acid Res. Mol. Biol.* **40**, 221–253.
Borgstahl, G. E. O., Pokross, M., Chehab, R., Sekher, A. & Snell, E. H. (2000). *J. Mol. Biol.* **296**, 951–959.
Brzozowski, A. M. & Walton, J. (2001). *J. Appl. Cryst.* **34**, 97–101.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *PyMol.* DeLano Scientific, San Carlos, CA, USA. http://pymol.sourceforge.net/.
Edwards, R. A., Whittaker, M. M., Whittaker, J. W., Baker, E. N. & Jameson, G. B. (2001). *Biochemistry*, **40**, 15–27.
Flohe, L. & Otting, F. (1984). *Methods Enzymol.* **105**, 93–104.
Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Henriques, A. O., Melsen, L. R. & Moran, C. P. Jr (1998). *J. Bacteriol.* **180**, 2285–2291.
Holm, R. H., Kennepohl, P. & Solomon, E. I. (1996). *Chem. Rev.* **96**, 2239–2314.
Krishnakumar, R., Craig, M., Imlay, J. A. & Slauch, J. M. (2004). *J. Bacteriol.* **186**, 5230–5238.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Liao, J., Liu, M.-Y., Chang, T., Li, M., Le Gall, J., Gui, L.-L., Zhang, J.-P., Jiang, T., Liang, D.-C. & Chang, W.-R. (2002). *J. Struct. Biol.* **139**, 171–180.
Miller, A. F. (2004). *Curr. Opin. Chem. Biol.* **8**, 162–168.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Ramilo, C. A., Leveque, V., Guan, Y., Lepock, J. R., Tainer, J. A., Nick, H. S. & Silverman, D. N. (1999). *J. Biol. Chem.* **274**, 27711–27716.
Stallings, W. C., Pattridge, K. A., Strong, R. K. & Ludwig, M. L. (1984). *J. Biol. Chem.* **259**, 10695–10699.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.