Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Expression, purification, crystallization and preliminary X-ray diffraction studies of recombinant class B non-specific acid phosphatase of *Salmonella typhimurium*

R. D. Makde,[a] Vinay Kumar,[a]* G. D. Gupta,[a] J. Jasti,[b] T. P. Singh[b] and S. K. Mahajan[c]†

[a]Synchrotron Radiation Section, Bhabha Atomic Research Centre, Mumbai 400 085, India, [b]Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110 029, India, and [c]Molecular Biology and Agriculture Division, Bhabha Atomic Research Centre, Mumbai 400 085, India

† Present address: 10-Rohini, Sector-9A, Vashi, Navi Mumbai 400 703, India.

Correspondence e-mail: vinay@magnum.barc.ernet.in

Received 6 May 2003
Accepted 12 August 2003

The *aphA* gene of *Salmonella enterica* sv. Typhimurium strain MD6001 was cloned in the multicopy plasmid pBluescript SK$^-$. The recombinant AphA protein was purified to homogeneity. The protein crystallized in the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 112.4, b = 130.2, c = 139.6 Å. Consistent with the self-rotation function, there are two tetramers in the asymmetric unit, indicating a solvent content of ∼54%. The crystals are composed of biologically active AphA molecules.

1. Introduction

Bacterial non-specific acid phosphatases (bNSAPs; EC 3.1.3.2) are secreted enzymes that function as soluble periplasmic proteins or as membrane-bound lipoproteins (Rossolini *et al.*, 1998). They are a diverse group of monomeric or oligomeric proteins with monomer molecular weight in the range 25–30 kDa. These enzymes do not exhibit marked substrate specificity and can dephosphorylate disparate phosphomonoesters at acidic to neutral pH. Based on their amino-acid sequence, the bNSAPs have been further classified into three classes: A, B and C. The amino-acid sequences of class A proteins are completely unrelated to those of class B and class C proteins. Class B, class C and some plant phosphatases (vegetative storage proteins) show amino-acid sequence similarity (identity 12–22%) and have been grouped into the DDDD family (Thaller *et al.*, 1998; Rossolini *et al.*, 1998). The four invariant aspartate residues of the DDDD family have been proposed to be activity-linked, as the phosphatase activity is observed to be abolished on mutation of any of these Asp residues (Reilly *et al.*, 2001; Godlewska *et al.*, 2002).

The *aphA* gene of *Salmonella enterica* sv. Typhimurium (*S. typhimurium*) codes for a 237-amino-acid polypeptide, 23 of which function as the signal peptide for periplasmic localization (Swiss-Prot accession No. P58683). The periplasmic AphA protein is a member of the class B bNSAPs and exists as homotetramer of molecular weight ∼100 kDa. The AphA protein possesses the consensus signature sequence motifs of the aspartyl-phosphate-utilizing phosphohydrolases/phosphotransferases (APUP; Collet *et al.*, 1998; Wang *et al.*, 2002) and the mitochondrial energy-transfer proteins (Prosite code PDOC00189). The physiological function of the AphA protein is largely unclear, although it is generally believed that bNSAPs function in the periplasmic space as organophosphate-degrading enzymes (Rossolini *et al.*, 1998). The hydrolysis of organophosphates in the periplasm enables the organic moieties to enter the cell. The inorganic phosphate released during the catalysis can be useful in bioremediation of nuclear waste (Macaskie *et al.*, 1996). The phosphotransferase activity of the AphA enzyme has been demonstrated by terminal phosphate transfer from hydrolysable phosphomonoesters to organic compounds with a free hydroxyl group (Uerkvitz, 1988). The AphA protein has also been thought to be involved in an apparently unrelated biological function, namely chromosomal replication, as it binds to the hemi-methylated DNA at *ori*C of *Escherichia coli* (Reshetnyak *et al.*, 1999). Purification and crystallization of AphA from *S. typhimurium* LT2 has been reported without characterization and X-ray analysis of the crystals (Uerkvitz, 1988). Here, we report the cloning of the *aphA* gene of a pathogenic strain (MD6001) of *S. typhimurium* and the purification, crystallization and preliminary X-ray analysis of the recombinant protein.

2. Cloning of the *aphA* gene

Unless otherwise specified, standard molecular-biology techniques (Sambrook *et al.*, 1989) were used. The pathogenic *S. typhimurium* strain MD6001 obtained from the Central Research Institute (CRI), Kasuali, India (CRI No. 5710) had previously been identified in our laboratory based on the 16S and 23S rRNA gene sequences (Makde *et al.*, 2003). The primers Apha3 (5′-CCGGCCTGACAATAGCTGTTGTGAT-3′) and Apha4 (5′-CGTCTCCTGAATAGTGTGCAGCAAG-3′), designed on the basis of the published sequence of the *aphA* gene and its flanking regions in the *S. typhimurium* LT2 genome (McClelland *et al.*, 2001; GenBank accession Nos. NC_003197 and AE008898),

were employed to amplify a 973 bp *aphA* fragment from the genomic DNA of MD6001. The amplified DNA fragment contained the *aphA* structural gene along with 155 bp upstream of the start codon and 104 bp downstream of the stop codon. The amplified product was blunt-ended using T4 DNA polymerase and ligated with *Hin*cII-restricted pBluescript pSK$^-$ vector. The ligation mixture was used to transform the DH5α strain of *E. coli* and the transformants were selected on LB plates supplemented with ampicillin, X-Gal and IPTG. The white recombinant clones were monitored on phosphatase-indicator plates as described in Makde *et al.* (2003). The phosphatase-positive clones were checked for the proper inserts by PCR and restriction-enzyme analysis. All recombinant clones showed the presence of a properly oriented *aphA* gene in the plasmid. The clones were checked for phosphatase activity using pNPP (*p*-nitrophenyl phosphate) as the substrate (Kier *et al.*, 1977). The clone (named pSK$^-$-*aphA*) showing the highest activity was selected for purification of the AphA protein. The sequences of the PCR product of the chromosomal *aphA* of MD6001 (GenBank accession No. AY125468) and the *aphA* insert in the recombinant plasmid (pSK$^-$-*aphA*) were determined using an ABI automated DNA sequencer.

3. Protein purification and characterization

The protein was purified using a low-pressure liquid-chromatography system (Bio-Rad Biologic Lp) with manually packed columns at 297 K, except where stated otherwise. The purification progress was monitored by pNPP phosphatase assay and SDS–PAGE. The recombinant DH5α clone was grown to 1.2 OD_{600} in LB media containing 100 µg ml^{-1} ampicillin. The cells (25 g) harvested from a 7 l culture by centrifugation at 4000*g* at 277 K were suspended in 100 ml ice-cold lysis buffer [50 m*M* Tris–HCl pH 7.6, 200 m*M* NaCl, 1 m*M* DTT and one tablet of protease-inhibitor cocktail (Roche) per 100 ml]. The suspended cells were disrupted by sonication for 3 min in 30 ml portions maintained on ice. The lysate was cleared of cellular debris by centrifugation at 20 000*g* for 1 h and the supernatant was subjected to 60% ammonium sulfate precipitation. The precipitate collected by centrifugation at 12 000*g* for 30 min was suspended in buffer *A*1 (10 m*M* Tris–HCl pH 7.6, 1 m*M* $MgCl_2$) and extensively dialyzed against the same buffer, followed by loading onto a pre-equilibrated DEAE-Sepharose CL6B column (Amersham-Pharmacia; 2.5 × 18 cm). The bound proteins were eluted with a linear gradient of 0–300 m*M* NaCl over six column volumes in buffer *A*1 at a flow rate of 1.0 ml min^{-1}. The peak phosphatase activity was eluted at approximately 80 m*M* NaCl. The fractions showing phosphatase activity were pooled and dialyzed against buffer *A*2 (10 m*M* sodium acetate buffer pH 5.4, 1 m*M* $MgCl_2$, 0.5% PEG 6000), which unexpectedly resulted in nearly twofold purification of the AphA protein owing to precipitation of contaminant proteins. The supernatant was loaded onto a CM-Sepharose FF column (Amersham-Pharmacia; 1.5 × 15 cm) and elution of the bound proteins was achieved with a linear gradient of 0–1.0 *M* NaCl in buffer *A*2. A single peak of phosphatase activity was eluted at approximately 315 m*M* NaCl. The AphA protein was further purified by the chromatofocusing method. The active fractions were dialyzed against the start buffer (25 m*M* imidazole–HCl pH 7.6, 1 m*M* $MgCl_2$) and loaded onto polybuffer exchanger PBE94 (Amersham-Pharmacia) with elution of the proteins achieved with eluent buffer [polybuffer 74 (1:8 dilution) pH 4]. The phosphatase activity was recovered at pH 6.6. The active fractions were subjected to 37% ammonium sulfate saturation by adding solid ammonium sulfate. The clear solution was loaded onto a Phenyl-Sepharose HP column (Amersham-Pharmacia; 5 ml) and the bound proteins were eluted with a decreasing concentration of ammonium sulfate from 1.5 to 0 *M* in buffer *A*3 (50 m*M* sodium phosphate pH 7.0). The single peak of phosphatase activity eluted at ∼660 m*M* ammonium sulfate in buffer *A*3 showed a specific pNPP phosphatase activity (98 U mg^{-1}) that was very similar to that reported for AphA protein purified from *S. typhimurium* LT2 (Uerkvitz, 1988). Significant enhancement in the pNPP phosphatase activity was observed in the presence of Mg^{2+} and Zn^{2+} ions (data not shown). A single band on SDS–PAGE showed that the purified protein was a homogenous preparation with a molecular weight of ∼24 kDa (Fig. 1). The mobility of the partially denatured protein (unboiled in 2% SDS sample buffer) on SDS–PAGE indicated that the active AphA protein exists as a tetramer (Fig. 1).

The N-terminal amino-acid residues of the purified protein were sequenced using a PPSQ-10 protein sequencer (Shimadzu). The 15-residue sequence exactly matched residues 24–38 of the AphA amino-acid sequence. The amino-acid sequence and the phosphatase activity observed with the intact cells confirm the processing of the 23-residue signal peptide and the localization of the 214-residue mature protein in the periplasm.

We also purified to homogeneity a totally inactive protein product which was generated on storage of the partially purified AphA enzyme. The inactive protein stored at 277 K for a month was observed on SDS–PAGE to be of reduced molecular weight (∼22 kDa) (Fig. 1). The presence of the amino-acid sequence LAEQAPV... at the N-terminus of this protein revealed the cleavage of an additional 16 amino-acid residues from the active periplasmic protein. The truncated AphA was observed as a monomer on SDS–PAGE, in contrast to the tetrameric active AphA protein (Fig. 1).

4. Crystallization

The protein was equilibrated against buffer *A*4 (25 m*M* sodium acetate pH 5.6) and concentrated to 6 mg ml^{-1} using a Centricon-30 (Millipore). The initial crystallization conditions were screened by means of the hanging-drop vapour-diffusion method at 283 K using conditions similar to the sparse-matrix conditions (Jancarik & Kim, 1991) and the PEG 6000 grid of JBScreen 4 (Jena Bioscience GmbH). Typically, 2 µl protein solution and 2 µl precipitation solution were mixed on a siliconized cover slip and were equilibrated against 500 µl of the same precipitation solution in the reservoir. Crystals with dimensions of less than 30 µm were observed in crystallization conditions with a PEG 6000 concentration varying between 6 and 12% over the pH range 4.4–5.6. Improvements to the size of the crystals were achieved using

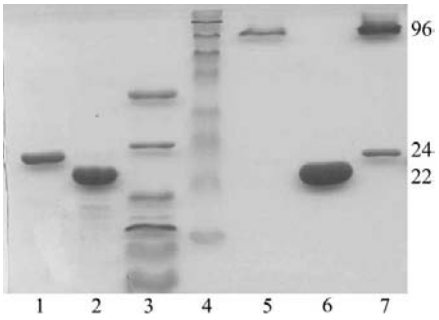

Figure 1
12% SDS–PAGE with protein samples prepared in standard buffer: active AphA protein (lane 1), truncated AphA (lane 2), partially denatured (unboiled in 2% SDS sample buffer) AphA crystals (lane 5), partially denatured truncated AphA (lane 6) and partially denatured active AphA (lane 7). Standard low- and high-molecular-weight markers were loaded in lanes 3 and 4, respectively. Molecular weights are in kDa.

the sitting-drop vapour-diffusion method. During the trials, 5 µl protein solution and 5 µl precipitation solution mixed on a Hampton microbridge were equilibrated against the precipitating solution. Elongated thin needles of dimensions 200 × 50 × 50 µm grew at 295 K from condition *C*1 (10% PEG 6000, 25 m*M* sodium acetate pH 4.7, 5 m*M* $MgCl_2$). Crystals of dimensions 800 × 200 × 150 µm were grown under condition *C*1 by macroseeding these needles into sitting drops at 295 K. The macroseeds were introduced after pre-equilibrating the sitting drops against 500 µl of the precipitating solution *C*1 for 2 d. SDS–PAGE of the crystals confirmed the protein integrity (Fig. 1). The crystals are composed of the functionally active form of the protein; addition of the substrate pNPP turned the thoroughly washed crystals yellow, presumably owing to the formation of *p*-nitrophenylate.

5. Data collection and crystallographic analysis

Single crystals were mounted in Lindemann thin glass capillaries for X-ray diffraction. The diffraction-intensity data were collected at room temperature using a MAR 345dtb imaging-plate system (MAR Research) and Cu $K\alpha$ radiation from a Rigaku RU-300 rotating-anode X-ray generator equipped with Osmic mirror optics (Molecular Structure Corporation). The crystals diffracted to 2.9 Å resolution. The unit-cell parameters were determined by careful inspection of graphical images of the diffraction and were confirmed by autoindexing using *DENZO*/*SCALEPACK* (Otwinowski & Minor, 1997). The integrated intensity data were scaled and merged using *SCALEPACK* (Table 1). The crystals are orthorhombic, space group $P2_12_12_1$, with unit-cell parameters a = 112.4, b = 130.2, c = 139.6 Å. One monomer with a molecular weight of 24 kDa (as estimated from the amino-acid sequence) per asymmetric unit of the $P2_12_12_1$ cell corresponds to a Matthews coefficient (V_M; Matthews, 1968) of 21.7 $Å^3\ Da^{-1}$. Since the AphA protein has been demonstrated to exist as a tetramer of nearly 100 kDa, two tetramers of the AphA protein can be expected to be present in the crystallographic asymmetric unit of the $P2_12_12_1$ unit cell, corresponding to a V_M and solvent content of 2.67 $Å^3\ Da^{-1}$ and 54%, respectively. The presence of eight monomers of the AphA protein in the asymmetric unit is also consistent with the self-rotation function. The program *POLARRFN* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994) was employed to compute the self-rotation function using a radius of integration of 18 Å and data in the resolution range 10–4 Å. The self-rotation function clearly revealed the presence of a fourfold and a twofold non-crystallographic symmetry (NCS) axes. The fourfold NCS axis is parallel to the *b* crystallographic axis of $P2_12_12_1$, corresponding to $(\omega, \varphi, \kappa) = (90, 90, 90^\circ)$ (Fig. 2). The twofold NCS axis lies in the plane parallel to the *ac* crystallographic plane at an inclination of ∼45° with respect to the *c* crystallographic axis, corresponding to $(\omega, \varphi, \kappa) = (45, 0, 180^\circ)$. The intensity of the NCS peaks relative to the crystallographic twofold axes was observed to be ∼70%. A head-to-head or tail-to-tail packing of the two tetramers in the crystallographic asymmetric unit is suggested by the mutually orthogonal fourfold and twofold non-crystallographic symmetry axes.

Table 1
Data-collection statistics.

Values in parentheses refer to the highest resolution shell (3.7–3.6 Å).

X-ray source	Rotating-anode X-ray generator
Detector	MAR 345dtb imaging plate
Temperature (K)	298
Wavelength (Å)	1.5418
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 112.4, b = 130.2, c = 139.6
Monomers in asymmetric unit	8
V_M ($Å^3\ Da^{-1}$)	2.67
Solvent content (%)	54
No. of measured reflections	99078
No. of unique reflections	24431
Resolution range (Å)	20–3.6
Completeness (%)	99.9 (99.7)
Average $I/\sigma(I)$	10.3 (2.6)
Multiplicity	4.0
R_{merge}† (%)	11.7 (45.5)

† $R_{merge} = \sum_{hkl} |I_{hkl} - \langle I_{hkl}\rangle| / \sum_{hkl} I_{hkl}$.

6. Three-dimensional structure and biological significance

The biological role of the class B bacterial NSAPs is not well understood. The phosphohydrolase/phosphotransferase activity of the AphA protein has been demonstrated for a number of substrates differing in the structure of the organic moiety and the protein has also been shown to bind to hemimethylated GATC of *ori*C DNA (Reshetnyak *et al.*, 1999). The phosphatase activity of class B bacterial NSAPs has been demonstrated against organophosphates which are hydrolysable by the structurally characterized class A bacterial non-specific acid phosphatase (Pho of *E. blattae*; Ishikawa *et al.*, 2000). Although class B NSAPs share functional similarity with class A NSAPs, the chemistry of the reaction catalysed is expected to be different owing to the lack of sequence similarity and the absence of any shared catalytic motifs. A BLAST (Altschul *et al.*, 1997) search with the AphA amino-acid sequence did not reveal any close homologue in the PDB (Berman *et al.*, 2000). Fold prediction using the Structure Prediction MetaServer (http://bioinfo.pl/meta; Bujnicki *et al.*, 2001) assigned an L-2-haloacid dehalogenase-like (HAD-like) hydrolase superfamily fold (SCOP c1.108.1) for the AphA protein. The HAD-like superfamily is characterized by two domains: an α/β core domain resembling the NAD(P)-binding Rossmann fold and a cap domain inserted between β1 and α5 of the core domain (Hisano *et al.*, 1996) (Fig. 3).

The members of the HAD-like superfamily show poor sequence identity of less than 15% with the AphA sequence, with the major differences at the N-terminal (∼38 residues) and the cap domain (residues 55–87 of the AphA protein; Fig. 3). A partial structural model of the AphA monomer for residues 39–54 and 90–214, defining the structurally conserved regions (SCR) of the core domain, has been constructed (Vinay Kumar *et al.*, unpublished results) using the homology-modelling software *MODELLER* (Sali & Blundell, 1993). The atomic coordinates of the SCR, however, can only be derived to low resolution in view of the inconclusive poor sequence identity with any of the identified template structures. Also, the structure and function of

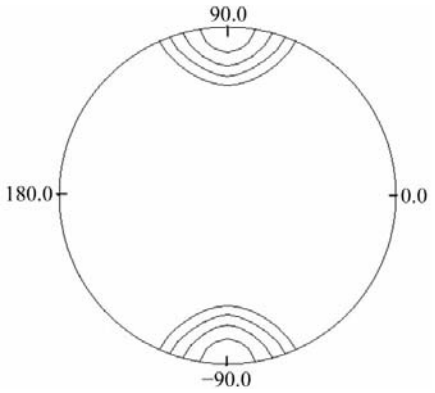

Figure 2
Self-rotation function, $\kappa = 90^\circ$ section. The ω angle varies in the radial direction and the angle φ varies along the circumference. The radius of integration was 18 Å and the resolution limits were 10–4 Å. A peak arising from the fourfold non-crystallographic symmetry axis is observed at $(\omega, \varphi, \kappa) = (90, 90, 90^\circ)$. The non-crystallographic fourfold axis is thus parallel to the crystallographic *b* axis. Contours are shown in steps of 10% starting from 30% of the maximum.

```
Stm_pcr  LVSSPSTLNPGTNVAKLAEQAPVHWVSVAQIENSLTGR---PPMAVGFDIDDTVLFSSPG 57
d1k1ea_  ------------------------------------KLENIKFVITDVDGVLTDG--- 19
d1feza_  -----------------------------------------KIEAVIFDWAGTTVDYGCF 19
d1j97a_  ------------------------------------EKKKKLILFDFDSTLVNNETI 21
d1qq5a_  ---------------------------------------mikavvfdaygtlfdvQSV 19

Stm_pcr  FWRGKKTYSPDSDDYLKNPAFWEKMN---------------------------------- 83
d1k1ea_  ------------------------------------------------------------ 19
d1feza_  APLEVFMEIFHKRGVAITAEEARKPMGLLKIDHVRALTEMPRIASEWNRVFRQLPTEADI 79
d1j97a_  DEIAREAGVEEEVKKITKEAMEGKLNFEQSLRKRVSLLKDLPIEKVEKA---------- 70
d1qq5a_  ADATERAYPGRGEYITQVWRQKQLEYSWLRALMGRYADFWSVTREALAYTLGTLGLEPDE 79

Stm_pcr  ---------------NGWDE--FSIPKEAARQLIDMHVRRGDSIYFVTGRSQTKTETVSK 126
d1k1ea_  ----------QLHYDANGEA--IKSFHVRDGLGIKMLMDADIQVAVLSGRDSPILRRRIA 67
d1feza_  QEMYEEF---EEILFAILPR--YASPINAVKEVIASLRERGIKIGSTTGYTREMMDIVAK 134
d1j97a_  ------------------IK--RITPTEGAEETIKELKNRGYVVAVVSGGFDIAVNKIKE 110
d1qq5a_  SFLADMAQA----------ynrltpypdaaqclaelap--lkrailsngapdmlqalva 126

Stm_pcr  TLADNFHIPAANMNPVIFAGD---------KPEQNTKVQWLQEKNM-----RIFYGDSDN 172
d1k1ea_  DLGIKLFFLGK-------------------LEKETACFDLMKQAGVT-AEQTAYIGDDSV 107
d1feza_  EAALQGYKPDFLVTPDDVPAG---------RPYPWMCYKNAMELGVYPMNHMIKVGDTVS 185
d1j97a_  KLGLDYAFANRLIVKDGKLTGDVEGEVLKENAKGEILEKIAKIEGIN-LEDTVAVGDGAN 169
d1qq5a_  nagltdsfdavisvdakrvf----------kphpdsyalveevlgvt-paevlfvssngf 175

Stm_pcr  DITAARDCGIR--GIRIL----RAAN---------------------------------- 192
d1k1ea_  DLPAFAACGT---SFAVA----DAPIYVKNAVDHVLSTHGGKGAFREMSDMILQAQGKSS 160
d1feza_  DMKEGRNAGMWTVGVILGSSELGLTEEEVENMDSVELREKIEVVRNR------------- 232
d1j97a_  DISMFKKAGL---KIAF-----CAKPILKEKADICIEKRDLREILKYIK----------- 210
d1qq5a_  dvggaknfgf---svarvarl-sqealarelvsgtiapltmfkal--------------- 216

Stm_pcr  -----------------STYKPLPQAGAFGEEVIVNSEY------- 214
d1k1ea_  VFDTAQGFLKSVKSMGQ---------------------------- 177
d1feza_  -----------------FVENGAHFTIETMQELESVMEHIE----- 256
d1j97a_  ---------------------------------------------- 210
d1qq5a_  -----------------rmreetyaeapdfvvpalgdlprlvrgma 245
```

Figure 3
An alignment of the AphA sequence (Stm_pcr) with representative sequences of the HAD-like superfamily. The alignment of the sequences was performed using the superfamily server (http://supfam.org; Gough & Chothia, 2002). The aligned protein sequences can be identified by the PDB codes (four middle letters): phosphatase YrbI, d1k1ea; phosphonoacetaldehyde hydrolase, d1feza; phosphate-L-serine phosphatase, d1j97; L-2-haloacid dehalogenase, d1qq5. The amino acids forming the core domain of the L-2-haloacid dehalogenase are shown in lower case. Insertions in the sequence alignment are indicated by -.

non-homologous polypeptide regions constituting the N-terminal polypeptide and the cap domain of AphA cannot be elucidated using bioinformatics approaches. The partial tertiary structure (structural model of the core domain) of the AphA protein is insufficient to understand the exact mechanistic details, as the active site in the HAD-like fold is situated between the core and the cap domains.

The tetrameric assembly of the AphA protein has been reported to be essential for phosphatase activity (Uerkvitz, 1988). We also observed a total loss of phosphatase activity in the truncated AphA protein lacking 16 N-terminal residues. The inactive protein product was observed to be a monomer, suggesting a role for the N-terminal residues in constituting the active tetramer of the AphA protein. However, owing to the non-availability of any correct template structure, the quaternary structure of the protein and its functional role cannot be elucidated from the model of the monomer (Fig. 3), further emphasizing the need for the experimental structure of the active AphA protein.

In summary, we report here cloning, purification and crystallization of the AphA protein from *S. typhimurium* with a view to elucidating the structure of the protein. The crystals diffract to 2.9 Å resolution using a laboratory X-ray source. The presence of two biologically active tetramers in the crystallographic asymmetric unit is of considerable significance. This is the first report of the characterization of the crystals of a bacterial class B non-specific acid phosphatase. We propose to explore the use of the low-resolution partial structural model that we have constructed using bioinformatics and fold-prediction algorithms in elucidating the experimental structure of the AphA protein.

Note added in proof: Crystallization and preliminary X-ray characterization of AphA from *E. coli* has been published after the submission of this manuscript (Forleo *et al.*, 2003). The point symmetry of the tetramer of AphA from *E. coli* is expected to be 222.

We are deeply indebted to Dr S. K. Sikka for constant interest in the work and for many useful discussions, Dr Anil Lala of the National Facility for Photolabelling and Peptide Sequencing in Biomolecular systems, Indian Institute of Technology, Mumbai for the N-terminal sequencing of the protein and Dr R. Mukhopadhyaya for help in DNA sequencing.

References

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bujnicki, J. M., Elofsson, A., Fischer, D. & Rychlewski, L. (2001). *Protein Sci.* **10**, 352–361.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Collet, J. F., Strooban, V., Pirard, M., Delpierre, G. & van Schaftingen, E. (1998). *J. Biol. Chem.* **273**, 14107–14112.
Forleo, C., Benvenuti, M., Calderone, V., Schippa, S., Docquier, J.-D., Thaller, M. C., Rossolini, G. M. & Mangani, S. (2003). *Acta Cryst.* D**59**, 1058–1060.
Godlewska, R., Bujnicki, J. M., Osterowski, J. & Jagusztyn Krynicka, K. (2002). *FEBS Lett.* **525**, 39–42.
Gough, J. & Chothia, C. (2002). *Nucleic Acids Res.* **30**, 268–272.
Hisano, T., Hata, Y., Fujii, T., Liu, J.-Q., Kurihara, T., Esaki, N. & Soda, K. (1996). *J. Biol. Chem.* **271**, 20322–20330.
Ishikawa, K., Mihara, Y., Gondoh, K., Suzuki, E. & Asano, Y. (2000). *EMBO J.* **19**, 2412–2423.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kier, L. D., Weppelman, R. & Ames, B. N. (1977). *J. Bacteriol.* **130**, 399–410.
Macaskie, L. E., Lloyd, J. R., Thomas R. A. P. & Tolley, M. R. (1996). *Nucl. Energy*, **35**, 257–271.
McClelland, M. *et al.* (2001). *Nature (London)*, **413**, 852–856.
Makde, R. D., Kumar, V., Rao, A. S., Yadava, V. S. & Mahajan, S. K. (2003). *Acta Cryst.* D**59**, 515–518.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Reshetnyak, M., Alencon, E., Kern, R., Taghbalout, A., Guillaud, P. & Kohiyama, M. (1999). *Mol. Microbiol.* **31**, 167–175.
Reilly, T. J., Green, B. A., Zlotnick, G. W. & Smith, A. L. (2001). *FEBS Lett.* **494**, 19–23.
Rossolini, G. M., Schippa, S., Riccio, M. L., Berlutti, F., Makaskie, L. E. & Thaller, M. C. (1998). *Cell Mol. Life Sci.* **54**, 833–850.
Sali, A. & Blundell, T. L. (1993). *J. Mol. Biol.* **234**, 779–815.
Sambrook, J., Fritch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press.
Thaller, M. C., Schippa, S. & Rossolini, G. M. (1998). *Protein Sci.* **7**, 1647–1652.
Uerkvitz, W. (1988). *J. Biol. Chem.* **263**, 15823–15830.
Wang, W., Cho, H. S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H. H., Grigoriev, I. V., Wemmer, D. E. & Kim, S.-H. (2002). *J. Mol. Biol.* **319**, 421–431.