BMC Structural Biology

Research article

Crystal structure of the YffB protein from *Pseudomonas aeruginosa* suggests a glutathione-dependent thiol reductase function

Alexey Teplyakov*[1], Sadhana Pullalarevu[1], Galina Obmolova[1], Victoria Doseeva[1], Andrey Galkin[1], Osnat Herzberg[1], Miroslawa Dauter[2], Zbigniew Dauter[2] and Gary L Gilliland*[1]

Address: [1]Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute and the National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, MD 20850, U.S.A and [2]National Cancer Institute, Brookhaven National Laboratory, Building 725A-X9, Upton, NY 11973, U.S.A

Email: Alexey Teplyakov* - teplyako@umbi.umd.edu; Sadhana Pullalarevu - pullalar@umbi.umd.edu; Galina Obmolova - galinao@niddk.nih.gov; Victoria Doseeva - doseeva@umbi.umd.edu; Andrey Galkin - galikin@umbi.umd.edu; Osnat Herzberg - herzberg@umbi.umd.edu; Miroslawa Dauter - dauter@bnl.gov; Zbigniew Dauter - dauter@bnl.gov; Gary L Gilliland* - gilliland@nist.gov

* Corresponding authors

Published: 08 March 2004

BMC Structural Biology 2004, **4**:5

Received: 03 January 2004
Accepted: 08 March 2004

This article is available from: http://www.biomedcentral.com/1472-6807/4/5

Abstract

Background: The *yffB* (PA3664) gene of *Pseudomonas aeruginosa* encodes an uncharacterized protein of 13 kDa molecular weight with a marginal sequence similarity to arsenate reductase from *Escherichia coli*. The crystal structure determination of YffB was undertaken as part of a structural genomics effort in order to assist with the functional assignment of the protein.

Results: The structure was determined at 1.0 Å resolution by single-wavelength anomalous diffraction. The fold is very similar to that of arsenate reductase, which is an extension of the thioredoxin fold.

Conclusion: Given the conservation of the functionally important residues and the ability to bind glutathione, YffB is likely to function as a GSH-dependent thiol reductase.

Background

The *yffB* (PA3664) gene of *Pseudomonas aeruginosa* encodes an uncharacterized protein of 13 kDa molecular weight. Based on the amino acid sequence analysis, YffB and its homologs have been assigned to the family of arsenate reductase (AR) and related proteins (Pfam entry PF03960) [1]. AR participates in arsenic detoxification by catalyzing the reduction of arsenate [the oxyanion of As(V)] to arsenite [the oxyanion of As(III)], which is then exported through a specific transport system [2]. There are two different types of bacterial ARs. AR of Gram-negative bacteria has a distinct HX_3CX_3R catalytic sequence motif, belongs to the thioredoxin (Trx) structural superfamily, and is coupled to the glutathione (GSH) and glutaredoxin (Grx) system for its enzyme activity [3,4]. AR of Gram-positive bacteria also has a redox active cysteine residue but within a CX_5R sequence motif embedded in a fold typical for low molecular weight protein tyrosine phosphatases, and requires Trx and Trx reductase for enzyme activity [5-7].

Besides the Trx-like AR, the PF03960 family includes a number of uncharacterized proteins whose function seems unlikely to be arsenate reductase. Members of the

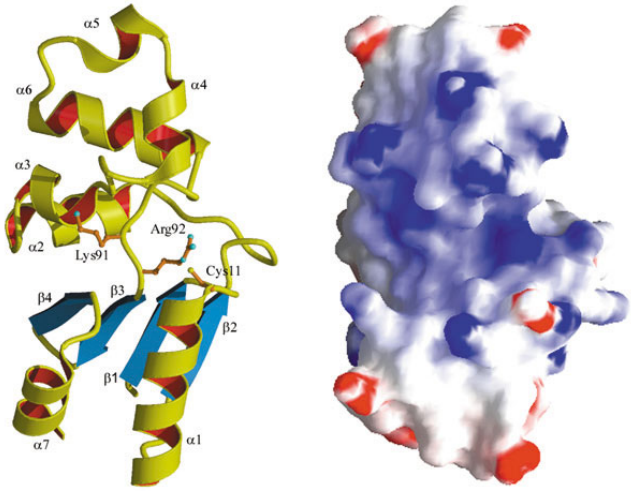

Figure 1
(Left) Ribbon presentation of the polypeptide fold of YffB. Active site residues are shown as ball-and-stick models (produced with MOLSCRIPT [19]). (Right) Electrostatic surface potential calculated with GRASP [20]. Positive charges are blue and negative are red.

family are widely represented in bacteria, but not in archaea or eukaryotes. The pattern of amino acid distribution along the polypeptide chain suggests that these proteins may have a common fold.

The crystal structure determination of YffB was undertaken as part of a structural genomics effort [8] in order to assist with the functional assignment of the protein. The project was focused on the so-called hypothetical proteins from *Haemophilus influenzae*. The YffB protein from *P. aeruginosa* has emerged as an ortholog of HI0103 to increase chances for successful crystallization. YffB was cloned, expressed, and the crystal structure determined at 1.0 Å resolution. The protein fold appeared to be similar to that of *Escherichia coli* AR. Analysis of the structure suggests that YffB may function as a thiol reductase.

Results and discussion

The atomic model of YffB contains all residues but the N-terminal methionine, which has probably been posttranslationally cleaved. The molecular weight of the SeMet protein measured by matrix-assisted laser-desorption ionization (MALDI) mass spectrometry was close to the calculated value of 13,121 Da for residues 2–115. An addition of one selenomethionine would have increased the molecular weight by 178 Da.

The structure of YffB consists of two domains (Fig. 1). One is formed by a four-stranded mixed β-sheet flanked by two α-helices on one side. The other domain is an α-helical bundle comprising residues 38–88. The overall fold is very similar to AR from *E. coli* encoded by the *arsC* gene [4]. The rms deviation between the structures is 2.3 Å for all 109 common Cα atoms, the Z-score calculated by DALI [9] is 12.6. Relative to YffB, ArsC has an additional C-terminal domain, a 3-strand meander, that covers helices α1 and α7.

The α/β-domain is characteristic of a superfamily of the Trx-like proteins [10]. Common to all of them, a *cis*-proline residue is located at the N-terminus of strand β3. This residue is important for the integrity of the active site, as the *cis*-peptide bond promotes a turn of the polypeptide chain. The corresponding Pro93 in YffB is also in the *cis*-conformation.

The overall structural similarity to the members of the Trx superfamily allows the identification of the putative active site of YffB in the loop between the first β-strand and the following α-helix. A cysteine residue at the N-terminus of the α-helix acts as a catalytic nucleophile [10]. Its pK is lowered by the basic environment and the dipole of the helix [11]. In ArsC the thiol group of the cysteine attacks the arsenic of the substrate to form a covalent intermediate [4]. The arsenite ion is released upon binding of GSH, which is then reduced by Grx [3]. Unlike many Trx-like proteins, the catalytic cysteine in ArsC does not form an internal disulfide in the oxidized state. YffB also has only one cysteine residue at the active site (Cys11), which further emphasizes its similarity to ArsC and suggests that it functions in a GSH-dependent manner.

Despite the structural similarity to ArsC, the two proteins share only 16% identical residues. However, most of the residues involved in substrate binding and catalysis are conserved in these proteins. Three invariant arginine residues, Arg60, Arg94, and Arg107, bind the substrate and stabilize the reaction intermediate in ArsC [12]. They may also enhance the nucleophilicity of the active cysteine. Arg94 together with the ensuing cis-Pro95 is conserved in all AR-related proteins. Lys91 of YffB, despite its different location in the sequence, is spatially equivalent to Arg60 of ArsC and therefore can take part in substrate binding. The most important difference with respect to ArsC is the substitution of Gly for Arg107 that leaves the binding site without the positively charged anchor and also makes Cys11 more accessible for bulky compounds. This substitution probably reflects a difference in substrate specificity between ArsC and YffB.

One particular consequence of this Arg to Gly replacement might be the ability of YffB to bind GSH, whereas ArsC cannot bind GSH in the absence of arsenate [13]. The binding of GSH to YffB was detected by MALDI mass

spectrometry using an oxidized form of glutathione (GSSG). The mass increase of about 300 Da indicated one GSH molecule bound per protomer. This result supports the contention that GSH is probably involved in the functional cycle of YffB.

YffB has a very polarized distribution of charges over the surface of the molecule. The active site area is thronged with basic residues, as it is in ArsC, whereas the opposite side of the molecule is predominantly negatively charged (Fig. 1B). The positive electrostatic potential would certainly favor binding of anions.

Conclusion

Given the structural similarity to the Trx-like proteins and particularly to ArsC, and the conservation of the functionally important residues, YffB is likely to function as a thiol reductase. The nature of the substrate remains to be established in further biochemical and biophysical studies. These studies will be facilitated by the three-dimensional structure of the protein.

Methods

Cloning, expression and purification

The *yffB* (PA3664) gene from *Pseudomonas aeruginosa* PAO1 was amplified using *PfuTurbo* DNA polymerase (Stratagene), genomic DNA (ATCC 47085D), and the following 5'- and 3'-end primers.

Forward: 5'-CACCCTGGTGCCGCGCGGCAGC*CATAT*-GACCTACGTTCTCTACGGCATCA-3'.

Reverse: 5'-TCAGGCCAGGGCGGC-3'.

The sequence encoding the thrombin cleavage site is underlined, and the *Nde*I restriction site is shown in italic. The PCR product was introduced into a pET100/D-TOPO expression vector by the TOPO directional cloning procedure (Invitrogen). Recombinant plasmids were isolated from the *E. coli* TOP10 strain. The expression construct for production of the native protein without a His-tag was prepared by digestion with *Nde*I and self ligation. For production of the selenomethionine (SeMet) protein, *E. coli* strain B834 (DE3) was transformed with the recombinant plasmid, and cells were grown in a minimal medium supplemented with 100 μg/mL ampicillin and 40 μg/mL SeMet until the A_{600} reached 0.8. At this point the cells were induced with 1 mM isopropyl β-D-thiogalactoside and harvested after 3 h.

The SeMet protein was purified by column chromatography in three steps. The cell extract was applied to a Q Sepharose HP (Pharmacia) column equilibrated with 20 mM HEPES (pH 6.7), 50 mM NaCl, and 0.5 mM EDTA. About half of the protein bound to the column and was eluted in a 50–500 mM NaCl gradient. After dialysis in 20 mM HEPES (pH 6.7) and 0.5 mM EDTA, the fractions containing the protein were applied to a Source 15S (Pharmacia) column, and eluted with a 0–450 mM NaCl gradient. The protein was concentrated to 4 mg/ml, applied to a Sephacryl S100 (Pharmacia) gel filtration column, and eluted in 20 mM HEPES (pH 7.5), 100 mM NaCl, and 0.25 mM EDTA. According to the SDS gel, the protein was at least 95% pure. For crystallization, the protein was concentrated to 15 mg/ml.

Crystallization and structure determination

YffB crystals were grown by the vapor diffusion hanging drop method at room temperature from 0.1 M CHES, pH 10, 26% polyethylene glycol 3350, and 5% isopropanol. They belong to the space group C2 with unit cell parameters: a = 87.45 Å, b = 43.25 Å, c = 29.06 Å, β = 93.5°. There is one protein molecule in the asymmetric unit with the solvent content of 40%. For X-ray data collection, the crystals were soaked in the mother liquor supplemented with 15% polyethylene glycol 400 and flash-frozen in liquid propane.

The structure was solved by using single-wavelength (0.9794 Å) anomalous X-ray diffraction data collected on the NCI-NIH beamline at the National Synchrotron Light Source (Upton, NY). The data (Table 1) were processed with HKL2000 [14]. Two selenium sites were located by SHELXD and were used for phasing with SHELXE [15]. The polypeptide chain was automatically traced with RESOLVE [16]. The atomic model was completed using O [17] and refined with REFMAC [18] using anisotropic B-factors. The model includes residues 2–115 of the protein, a molecule of isopropanol, and 220 water molecules. 93% residues have main-chain torsion angles in the most favored conformation.

The atomic coordinates of YffB and structure factors were deposited in the Protein Data Bank under the accession code 1RW1.

Glutathione binding

Binding of glutathione was detected by MALDI mass spectrometry using a Voyager spectrometer (Applied Biosystems, Foster City, CA). The SeMet protein (10 μM) was incubated for 1 h at room temperature with 1 mM GSSG in 20 mM HEPES buffer, pH 7.**T**he sample was mixed 1:1 with matrix solution (10 mg/mL 3,5-dimethoxy-4-hydroxycinnamic acid, 50% aqueous acetonitrile, and 0.2% trifluoroacetic acid), deposited onto a golden plate, and allowed to dry at room temperature. Bovine myoglobin was used for molecular mass calibration.

Table 1: X-ray data and refinement statistics.

Resolution (Å)	30-1.02 (1.04-1.02)[1]
Number of unique reflections[2]	98,866 (5,709)
Completeness (%)	91.2 (92.3)
Redundancy	3.9 (3.7)
R_{sym} ($\Sigma\|I-\langle I\rangle\|/\Sigma I$)	0.051 (0.419)
$\langle I/\sigma\rangle$	22.0 (3.6)
Fraction of refls with $I>3\sigma$ (%)	81.6 (50.2)
R_{cryst} ($\Sigma\|\|F_o\|-\|F_c\|\|/\Sigma\|F_o\|$)	0.129 (0.155)
R_{free} (2% data)	0.139 (0.202)
Number of protein non-hydrogen atoms	935
Number of water molecules	220
Mean B-factor from the model (Å^2)	10.4
Mean B-factor from Wilson plot (Å^2)	8.1
RMSD in bonds (Å)	0.011
RMSD in angles (°)	1.3
RMSD in main-chain B factors (Å^2)	2.1

[1]Data for the highest resolution shell are given in parentheses [2]Anomalous pairs not merged

Authors' contributions

AT modeled, refined and analyzed the structure, performed MALDI experiments, and drafted the manuscript. SP and GO purified and crystallized the protein. VD and AG cloned and expressed the protein. MD and ZD collected and processed the diffraction data and calculated electron density maps. OH and GLG coordinated the study and provided financial support.

Acknowledgement

This work was supported by the National Institutes of Health grant No. P01-GM57890.

Certain commercial materials, instruments, and equipment are identified in this manuscript in order to specify the experimental procedure as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the materials, instruments, or equipment identified is necessarily the best available for the purpose.

References

1. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL: **The Pfam protein families database.** *Nucleic Acids Res* 2000, **28:**263-266.
2. Mukhopadhyay R, Rosen BP, Phung le T, Silver S: **Microbial arsenic: from geocycles to genes and enzymes.** *FEMS Microbiol Rev* 2002, **26:**311-325.
3. Gladysheva TB, Oden KL, Rosen BP: **Properties of the arsenate reductase of plasmid R773.** *Biochemistry* 1994, **33:**7288-7293.
4. Martin P, DeMel S, Shi J, Gladysheva T, Gatti DL, Rosen BP, Edwards BF: **Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme.** *Structure* 2001, **9:**1071-1081.
5. Ji G, Garber EA, Armes LG, Chen CM, Fuchs JA, Silver S: **Arsenate reductase of *Staphylococcus aureus* plasmid pI258.** *Biochemistry* 1994, **33:**7294-7299.
6. Zegers I, Martins JC, Willem R, Wyns L, Messens J: **Arsenate reductase from *S. aureus* plasmid pI258 is a phosphatase drafted for redox duty.** *Nat Struct Biol* 2001, **8:**843-847.
7. Bennett MS, Guan Z, Laurberg M, Su XD: ***Bacillus subtilis* arsenate reductase is structurally and functionally similar to low molecular weight protein tyrosine phosphatases.** *Proc Natl Acad Sci USA* 2001, **98:**13577-13582.
8. Eisenstein E, Gilliland GL, Herzberg O, Moult J, Orban J, Poljak RJ, Banerjei L, Richardson D, Howard AJ: **Biological function made crystal clear – annotation of hypothetical proteins via structural genomics.** *Curr Opin Biotechnology* 2000, **11:**25-30.
9. Holm L, Sander C: **Touring protein fold space with Dali/FSSP.** *Nucleic Acids Res* 1998, **26:**316-319.
10. Martin JL: **Thioredoxin – a fold for all reasons.** *Structure* 1995, **3:**245-250.
11. Katti SK, LeMaster DM, Eklund H: **Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution.** *J Mol Biol* 1990, **212:**167-184.
12. Shi J, Mukhopadhyay R, Rosen BP: **Identification of a triad of arginine residues in the active site of the ArsC arsenate reductase of plasmid R773.** *FEMS Microbiol Lett* 2003, **227:**295-301.
13. Liu J, Rosen BP: **Ligand interactions of the ArsC arsenate reductase.** *J Biol Chem* 1997, **272:**21084-21089.
14. Otwinowski Z, Minor W: **Processing of X-ray diffraction data collected in oscillation mode.** *Methods Enzymol* 1997, **276:**307-326.
15. Schneider TR, Sheldrick GM: **Substructure solution with SHELXD.** *Acta Crystallogr D* 2002, **58:**1772-1779.
16. Terwilliger TC: **Automated structure solution, density modification and model building.** *Acta Crystallogr D* 2002, **58:**1937-1940.
17. Jones TA, Zou JY, Cowan SW, Kjeldgaard M: **Improved methods for building models in electron density maps and the location of errors in these models.** *Acta Crystallogr A* 1991, **47:**110-119.
18. Murshudov GN, Vagin AA, Dodson EJ: **Refinement of macromolecular structures by maximum-likelihood method.** *Acta Crystallogr D* 1997, **53:**240-255.
19. Kraulis PJ: **MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures.** *J Appl Crystallogr* 1991, **24:**946-950.
20. Nicholls A, Sharp KA, Honig B: **Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons.** *Proteins* 1991, **11:**281-296.