The 1.9Å X-ray structure of *Escherichia coli* Bacterioferritin.

Simon C. Willies¹, Michail N. Isupov¹, Elspeth F. Garman², and Jennifer A. Littlechild¹✉

¹Henry Wellcome Building for Biocatalysis, School of Biosciences, University of Exeter, Stocker Road, Exeter, EX4 4QD, UK.

²Laboratory of Molecular Biophysics, Department of Biochemistry, Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK.

j.a.littlechild@exeter.ac.uk

01392263434

Abstract

Whilst attempting to crystallise, and solve the structure of a thermophilic Aldo-Keto Reductase an unidentified host *Escherichia coli* protein was crystallised. Through the use of XAFS, and MAD, the protein was identified as *E. coli* Bacterioferritin. The crystal structure has been solved to 1.9Å, and shows the symmetrical binding of a heme molecule on the local two-fold axis between subunits and a pair of metal atoms bound to each subunit at the ferroxidase centre. These metals have been identified as zinc by the use of microPIXE.

Introduction

Escherichia coli Bacterioferritin is an iron storage protein consisting of 24 identical subunits (18KDa each) which form a roughly spherical structure approximately 120Å in diameter, with an 80Å hollow core. Within the core, iron is deposited as either a crystalline or amorphous ferrihydrite core. The X-ray structure of *E. coli* Bacterioferritin has previously been solved [1] to 2.9Å (PDB: 1BFR, 1BCF). The arrangement of the subunits exhibits 432 point group symmetry, with channels at the

4-fold and 3-fold axis connecting the inner core to the outside environment. Previous structures showed the heme binding in a symmetrical hydrophobic pocket located on the 2-fold axes between two symmetry related subunits, also two manganese ions are shown binding to each monomer at the putative ferroxidase centre.

As well as the heme binding bacterioferritin, *E. coli* also possess another iron storage protein in the form of the non-heme binding ferritin. Ferritin is an iron storage protein with a similar overall architecture to bacterioferritin but lacks heme, it is made up of 24 identical subunits of approximately 18KDa, binds two non-heme metals per subunit and also exhibits 432 point group symmetry. It has been suggested that the ferritin and bacterioferritin monomers are the bacterial equivalents of the H and L chains of the mammalian ferritin system [2].

Here *E. coli* BL21 expressing a gene for a thermophilic Aldo-Keto Reductase (AKR), also produced moderate amounts of bacterioferritin which was not separated from the protein of interest until the final stage of purification. The then unidentified protein was crystallised and by solving the crystal structure identified as *E. coli* bacterioferritin. This structure shows zinc binding in the ferroxidase centre, which has been shown to act as an inhibitor [3].

Materials and Methods

Expression and purification

The gene encoding the thermophilic AKR was cloned into the expression vector pTRC99 [4], and expressed in *E. coli* BL21. Expression was carried out by inoculating 1litre of liquid broth with 1ml of a starter culture and grown at 310K to an optical density of approximately 1.2AU prior to addition of IPTG and further

incubation at 310K for 16hours. Cells were harvested by centrifugation at 277K (12000g, 20mins).

All purification steps were performed at 277K unless stated. Cells were resuspended 10%(w/v) in 20mM Tris-HCl pH7.2 (Buffer A) and disrupted by sonication on ice using a Soniprep 150 (Sanyo) and the cell debris removed by centrifugation (12000g, 20mins). The supernatant was incubated with Benzonase (1U/ml) at 310K for 30mins. Protamine sulfate (1mg/ml) was added and incubated at 277K for 30mins, prior to centrifugation (12000g, 20mins). The protein was heated to 353K for 20mins, and centrifuged (12000g, 20mins), to precipitate and remove a large proportion of the E. coli protein.

The heat treated supernatant had KCl added to give a concentration of 2M and was purified using a Phenyl-Sepharose column (GE Healthcare, UK) the protein of interest eluted with a linear gradient from 2 to 0 M KCl in buffer A. Fractions containing the recombinant AKR were pooled together. The pooled protein was then purified by a Fast Flow Q anion exchange column (GE Healthcare, UK) and eluted with a gradient of 0-0.5M NaCl in buffer A. A final step of gel-filtration chromatography (Superdex 200; GE Healthcare, UK) was carried out. The purity of the sample was assessed through-out the procedure by SDS-PAGE [5], and activity assays.

Activity assays for the AKR were performed spectrophotometrically, by following the change in absorbance at 340nm due to the consumption of NADPH cofactor. Each reaction contained 0.9ml 20mM Tris-HCl pH7.2 with 25mM 1,2,3,6-tetrahydrobenzaldehyde, pre-heated to 323K, 20 μ l NADPH (1.3mM final concentration) was added, and finally 100 μ l of enzyme was added to start the reaction. The change in absorbance at 340nm was observed over a 5 minute period using a Shimadzu UV-2100 double-beam spectrophotometer.

Protein concentrations were estimated spectrophotometrically using the extinction coefficient for the AKR as calculated by the ExPASy database (Bairoch & Apweiler, 2000) using the amino-acid sequence.

Crystallisation

The purified protein was concentrated to 10mg ml⁻¹ in buffer A containing 0.1M NaCl using a Vivaspin centrifugal concentrator (Sartorius, UK) with a 30KDa molecular weight cutoff. Crystallisation was carried out using the sitting-drop vapour-diffusion method at 293K. 2µl of protein was mixed with 2µl of the reservoir solution to form the droplet.

Crystallographic data collection and processing

X-ray data were collected at beamline 10.1 at the SRS at Daresbury, using a single cryocooled crystal at 100K and a MAR225 CCD detector. The crystal diffracted to 1.9Å and belonged to space group P4₂2₁2, with unit cell parameters a=208.1 b=208.1 c=142.8Å and $\alpha=\beta=\gamma=90.00^\circ$. An XAFS scan at the Fe K edge was performed to confirm the presence of iron and allow data collection for MAD phasing. Data were collected at a peak wavelength of 1.739Å and a remote wavelength of 1.729Å. Both datasets were processed with DENZO and SCALEPACK [7]. The program TRUNCATE [8] was used to calculate structure factors and MLPHARE [9] for phase calculation. Averaging and phase combination was performed with DM [10].

MicroPIXE

MicroPIXE measurements were performed at the Surrey Ion Beam Centre at the university of Surrey. A 2.5MeV 1.5x1.5µm proton beam was used to induce x-ray emission from 0.2µl liquid protein samples dried onto 2µm thick mylar film in a vacuum. X-rays were detected by a **?solid-state lithium-drifted silicon detector?**. The proton beam was scanned across the drop and x-rays sorted into elemental maps, and

point spectra measured. **program?** was used to process the data and analyse the elemental content of the sample.

Results and Discussion

Purification and crystallisation

Throughout the purification the AKR co-purified with the bacterioferritin, and were only separated by gel filtration chromatography. The AKR eluted in a symmetrical peak corresponding to a molecular weight of 112KDa, and the bacterioferritin eluted at a volume corresponding to approximately 460KDa. SDS-PAGE analysis showed that the bacterioferritin peak contained a significant proportion of the AKR and also exhibited significant activity when assayed. The bacterioferritin fractions were concentrated to 10mg ml⁻¹ and exhibited a bright red colour.

The concentrated protein was used in crystallisation trials with varying amounts of ammonium sulfate, and yielded crystals overnight. The best crystals with the dimensions 0.12 x 0.12 x 0.12mm were obtained in 20mM Tris-HCl pH7.2, 60% ammonium sulfate. The crystals were transferred into cryogenic liquor containing 30% glycerol, 47% ammonium sulfate, 100mM NaCl in 20mM Tris-HCl pH7.2 and flash frozen in liquid N₂.

Structure solution and refinement

Iron atom positions were identified using SHELXD of the SHELX package [11] and used to create MAD phases. A partial model was built into the density using COOT [12]; refinement of partial model phases and combination with MAD phases was performed using the program REFMAC [13]. Several rounds of model building and refinement were carried out until sidechains could be identified. The putative amino

acid sequence was then input into BLAST [14] identifying the protein to be *E. coli* bacterioferritin.

A dimer of the bacterioferritin (PDB code 1BFR) was used as a model for molecular replacement with the remote wavelength data using MOLREP [15]. Further dimers, to produce a total of 12 monomers, were found neighbouring the first using molecular replacement and the symmetry operators from the iron positions using LSQKAB [16]. This model was used for further refinement using COOT and REFMAC. Processing statistics are presented in table 1.

MicroPIXE

The presence of iron in the Bacterioferritin was determined by XAFS prior to data collection, but the structure revealed other metal sites in the protein not occupied by iron. MicroPIXE analysis showed only iron and zinc present in the sample (Fig 1), unfortunately due to the presence of chloride in the buffer masking the sulphur signal it was not possible to quantify the amounts of iron and zinc present. Since all the iron positions were identified from MAD phasing and no other metal atoms were found in the PIXE analysis we conclude that the metal sites are occupied by zinc.

Final structure

Overall architecture

Since the structure of bacterioferritin has previously been solved [1] only a brief summary of the structure and its main features are provided here. Each subunit consists of a four helix bundle, with a short fifth helix. The biological unit is composed of 24 subunits in a near spherical arrangement, 120Å in diameter with an 80Å hollow core. The protein shows 432 point group symmetry, the four-fold and three-fold axes are located at what is expected to be channels linking the core to the external environment. The protein can be viewed as being made up of 12 dimers.

These dimers are arranged head to tail with a local two-fold axis located between them. (Fig. 2)

Heme

The iron located in the structure is present within a heme molecule situated on the local 2-fold axis within a symmetrical hydrophobic pocket created by a dimer, and is coordinated by two equivalent Met52 residues from neighbouring subunits. In the previous structures (PDB codes 1BFR & 1BCF), due to limited resolution it was not possible to show the heme binding in two orientations, although the authors expected this to be the case due to the symmetrical nature of the binding site [1]. At the resolution obtained here it has been possible to show that the heme binds in both orientations (Fig. 3).

Dinuclear metal site

Also in the previous structure two bivalent metal atoms were found in each subunit. These were assigned as manganese, due to the presence of manganese in their crystallisation buffer, but were not conclusively identified. These metals are coordinated by Glu 18, Glu 51, His 54, Glu 94, Glu 127, and His 130 in the previous structures and the one reported here. The metal atoms in the structure presented here have been identified as zinc which has previously been shown to act as an inhibitor of the ferroxidase centre.

Figure 1. MicroPIXE point spectra for bacterioferritin, showing peaks for iron and zinc.

Figure 2. Spherical structure of bacterioferritin viewed down 4-fold axis. Picture produced using PyMol (Delano Scientific, USA).

Figure 3. 2Fo-Fc map contoured at 1sigma for heme, showing one of the two orientations it can adopt. The other orientation is a 180° rotation about the y-axis round the iron. Picture produced using PyMOL (Delano Scientific, USA).

	Remote	Peak
Wavelength (Å)	1.729	1.739
Resolution	147.44	- 25 - 1.9
range (Å)	1.91 (1.96- 1.91)	(1.93- 1.90)
Completeness (%)	98.68 (99.34)	97.7 (78.9)
No. of reflections	231295 (15365)	460546 (18605)
No. of atoms	17869	
R value	0.197 (0.368)	
Free R value	0.224 (0.436)	
Mean B value	23.34	
Wilson B value	15.4	18.8

Table 1. Processing statistics for peak and remote data sets. Values in parentheses are for the highest resolution shell.

Acknowledgements

We would like to thank J.E. Guy for cloning of the Aldo-Keto Reductase, and the Surrey Ion Beam Centre for performing the microPIXE experiments.

References

1. Frolov F, Kalb A, Yariv J (1994) *Nat Struct Biol* 1:453-460
2. Stillman T J, Hempstead P D, Artymiuk P J, Andrews S C, Hudson A J, Treffry A, Guest J R, Harrison P M (2001) *J Mol Biol* 307, 587-603
3. Le Brun NE, Andrews SC, Guest JR, Harrison PM, Moore GR, Thomson AJ (1995) *Biochem J* 312:385-392
4. Guy J (2004) PhD thesis, University of Exeter
5. Laemmli U (1970) *Nature* 227:680-685
6. Bairoch A, Apweiler R (2000) *Nucleic Acids Res* 28:45-48
7. Otwinowski Z, Minor W (1997) *Methods Enzymol* 267:21839-21843
8. French GS, Wilson KS (1978) *Acta Cryst* A34:517
9. Collaborative Computational Project, Number 4 (1994) *Acta Cryst* D50:760-763
10. Cowtan K (1994) *Jnt CCP4/ESF-EACBM Newsl protein Crystallogr* 31:34-38
- 11.
12. Emsley P, Cowtan K (2004) *Acta Cryst* D60:2126-2132
13. Murshudov G, Vagin A, Dodson E (1997) *Acta Cryst.* D53:240-255
14. Altschul SF, Gish W, Miller W, Myers EW, Lipman D J (1990) *J Mol Biol* 215:403-410
15. Vagin A, Teplyakov A (1997) *J Appl Cryst* 30:1022-1025
16. Kabsch W (1976) *Acta Cryst* A32:922-923