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

Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of DsrEFH from *Allochromatium vinosum*

Christiane Dahl,[a] Andrea Schulte[a] and Dong Hae Shin[b]*

[a]Institut für Mikrobiologie und Biotechnologie, Rheinische Friedrich-Wilhelms-Universität Bonn, Meckenheimer Allee 168, D-53115 Bonn, Germany, and [b]College of Pharmacy, Ewha Womans University, Seoul 120-750, South Korea

Correspondence e-mail: dhshin55@ewha.ac.kr

Received 10 July 2007
Accepted 21 August 2007

In purple sulfur bacteria, the proteins encoded by *dsr* genes play an essential role in the oxidation of intracellular sulfur, which is an obligate intermediate during the oxidation of sulfide and thiosulfate. One such gene product, DsrEFH from *Allochromatium vinosum*, has been cloned, expressed, purified and crystallized. Synchrotron data were collected to 2.5 Å from a crystal of selenomethionine-substituted DsrEFH. The crystal belongs to the primitive monoclinic space group $P2_1$, with unit-cell parameters a = 56.6, b = 183.1, c = 107.8 Å, β = 99.6°. A full structure determination is under way in order to provide insight into the structure–function relationships of this protein.

1. Introduction

Sulfur of oxidation state zero stored in intracellular sulfur globules is an obligate intermediate during the oxidation of sulfide and thiosulfate (Pott & Dahl, 1998). The proteins essential for the oxidation of the stored sulfur are encoded in the dissimilatory sulfite reductase (*dsr*) locus in the phototrophic sulfur bacterium *Allochromatium vinosum*. The *dsr* gene cluster includes the *dsrABEFHCMK* genes and the following *dsrLJOPNSR* genes (Dahl *et al.*, 2005). Among the products of these genes, DsrE, DsrF and DsrH are predicted to be soluble cytoplasmic proteins with apparent molecular weights of 14.6, 15.6 and 11.1 kDa. Interestingly, DsrE, DsrF and DsrH form a soluble multimeric protein DsrEFH, which is an $\alpha_2\beta_2\gamma_2$-structured holoprotein with a molecular weight of 75 kDa (Dahl *et al.*, 2005). The primary sequences of DsrE, DsrF and DsrH are homologous to each other (Pott & Dahl, 1998). Therefore, DsrE and DsrF belong to the same family of conserved domains (Pfam 02635.11; COG 1553, COG 2044, COG 2923). DsrH is the prototype of yet another family of conserved proteins found in bacteria and archaea (Pfam04077.6; COG 2168), although it also can be fitted into the DsrE/F family (Fig. 1).

The molecular function of DsrEFH is not known. Therefore, we have initiated the determination of its three-dimensional structure in order to obtain clues to deducing its molecular function. Here, we report the cloning, overexpression, purification, crystallization and preliminary X-ray study of DsrEFH from *A. vinosum*.

2. Experimental methods

2.1. Cloning of DsrEFH in *Escherichia coli*

Chromosomal DNA of *A. vinosum* was obtained as described previously (Pott & Dahl, 1998). PCR amplification of the *dsrEFH* genes was performed with *A. vinosum* DNA as the template using *Pfu* polymerase (following the protocol provided by Stratagene) and the primers 5′-CGAGGTCCATATGAAGTTCGCGCTTCAG-3′ and 5′-GTAAAGAAAACTCGAGAATTACAACCAG-3′, both of which were designed to introduce an *Nde*I restriction site. After digestion with *Nde*I, the PCR product was cloned into the *Nde*I site of plasmid pET15b (Novagen).

2.2. Overexpression and purification of recombinant DsrEFH

Overproduction of DsrEFH was performed in *E. coli* BL21 (DE3) and resulted in protein that carried an amino-terminal His tag on

```
DsrE   -----MKFALQINEGPYQHQASDSAYQFAKAALEKG-HEIFRVFFYHDGVNNSTRLTTPP-QDDRHIVNRWAELAEQYE 72
DsrF   MSEVVKKFMYLNRKAPYGTIYAWEALEVVLIG--AAFDQDVCVLFLDDGVYQLTRGQDTKGIGMKNFSPTYRTLGDYEV 77
DsrH   -----MSILHTVNKSPFERNSLESCLKFATEG--------ASVLLFEDGIYAALAGTRV--------ESQVTEALG--K 56
1JX7   ----MQKIVIVANGAPYGSESLFNSLRLAIALREQESNLDLRLFLMSDAVTAGLRGQKP---GEGYNIQQMLEILTAQN 72
1L1S   ----MVDYRVVFHIDEDDESRVLLLISNVRNLMADLESVRIEVVAYSMGVNVLRRDSEYS--------GDVSELTG-QG 66
1X9A   -----MALVLVK----YGTDHPVEKLKIRSAK------AEDKIVLIQNGVFWALEELET-------------------P 45
                                             :.    .:

                                                                          Identity
DsrE   LDMVVCVAAAQRRGIVDE---GEASRNGKDATNIHPKFRISGLGQLVEAAIQADRLVVFGD 130     100.0%
DsrF   RRIYVDRDSLEARGLTQDDLVEIAFEDMETEEEFDNIVEVIDSARVSELMNESDAVFSF-- 136      15.4%
DsrH   LKLYVLGPDLKARGFS--------------DERVIPGISVVDYAGFVDLTTECDTVQAWL- 102      13.8%
1JX7   VPVKLCKTCTDGRGIS--------------TLPLIDGVEIGTLVELAQWTLSADKVLTF-- 117      18.5%
1L1S   VRFCACSNTLRASGMD--------------GDDLLEGVDVVSSGVGHIVRRQTEGWAYIRP 113       9.2%
1X9A   AKVYAIKDDFLARGYS----------------EEDSKVPLITYSEFIDLLEGEEKFIG--- 87       6.9%
         .          *                        . :             :
```

Figure 1
Sequence comparison of DsrE, DsrF, DsrH and some of their homologues of known structure. Abbreviations are as follows: 1JX7, YchN from *E. coli*; 1X9A, Tm0979 from *Thermotoga maritima*; 1L1S, Mth1491 from *Methanobacterium thermoautotrophicum*. '–' represents a gap, '*' identical residues, ':' highly conserved residues and '.' less highly conserved residues.

DsrE. Growth, induction with IPTG and cell harvesting were performed as described in Dahl *et al.* (2005). Thawed cells were resuspended in 50 m*M* NaH_2PO_4, 300 m*M* NaCl, 10 m*M* imidazole pH 8.0, incubated with lysozyme (1 mg ml^{-1}) for 1 h on ice and disrupted by sonication (1.5 min ml^{-1}; Cell Disruptor BIS, Branson) followed by centrifugation (25 000*g* for 30 min at 277 K). The supernatant was chromatographed on an Ni–NTA column (Qiagen) as specified by the manufacturer. The column was washed with a stepwise gradient of imidazole in 50 m*M* NaH_2PO_4, 300 m*M* NaCl. DsrEFH eluted at 100 and 150 m*M* imidazole. The combined fractions were dialyzed against 10 m*M* Tris–HCl pH 7.5 and loaded onto a Mono Q HR5/5 column equilibrated with the same buffer. The column was washed with 10 m*M* Tris–HCl pH 7.5 containing 100 m*M* NaCl. The protein was eluted with a linear gradient from 100 to 500 m*M* NaCl in 10 m*M* Tris–HCl pH 7.5. Fractions containing recombinant DsrEFH were combined, dialyzed against 10 m*M* Tris–HCl pH 7.5 and concentrated to a final protein concentration of 40 mg ml^{-1} by ultrafiltration centrifugation (Centriplus YM10, Millipore).

A selenomethionine derivative of the protein was produced in a methionine auxotroph: *E. coli* strain B834(DE3). The cells were grown and induced in M9 minimal medium containing 50 mg l^{-1} selenomethionine together with the other 19 amino acids (Ramakrishnan *et al.*, 1993). Purification of the selenomethionine-containing DsrEFH was performed as described above, except that all buffers contained 2 m*M* TCEP in order to avoid potential oxidation of selenomethionine. After chromatography on MonoQ, fractions containing DsrEFH were dialyzed against 100 m*M* ADA pH 6.5 containing 2 m*M* TCEP and concentrated to a final protein concentration of 45 mg ml^{-1} by ultrafiltration centrifugation. During purification, recombinant DsrEFH was detected using specific antisera.

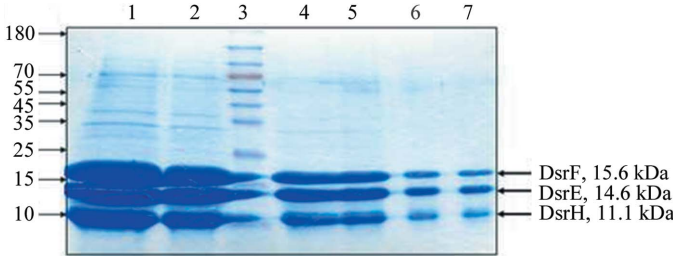

Figure 2
Coomassie-stained SDS–PAGE (15%) of DsrEFH after purification and concentration. Lane 1, 20 µl concentrated DsrEFH; lane 2, 10 µl concentrated DsrEFH; lane 3, prestained markers (labelled in kDa). Lanes 4, 5, 6 and 7 contain 5, 2, 1 and 0.5 µl concentrated DsrEFH, respectively. 1 µl of concentrated DsrEFH contains 16.8 µg protein.

2.3. Crystallization

The purified protein was concentrated to 20 mg ml^{-1} for crystallization. Screening for initial crystallization conditions was performed using the sparse-matrix method (Jancarik & Kim, 1991) with several screens from Hampton Research (Laguna Niquel, CA, USA) and from deCODE Genetics (Bainbridge Island, WA, USA). A Hydra-Plus-One crystallization robot (Matrix Technologies, Hudson, NH, USA) was used to set up screens using the sitting-drop vapour-diffusion method at room temperature. Since the first crystallization trial was not successful, optimum-solubility (OS) screening was performed to obtain biochemically pure and conformationally homogenous protein samples (Jancarik *et al.*, 2004). ADA buffer turned out to be the best buffer for the protein solution. In the optimized crystallization condition, 1 µl protein solution dialyzed against 0.1 *M* ADA pH 6.5 was mixed with 1 µl well solution containing 0.2 *M* Li_2SO_4, 0.1 *M* bis-Tris pH 5.5 and 25% PEG 3350 using the hanging-drop vapour-diffusion method.

2.4. Data collection and reduction

1 µl of reservoir solution in which the PEG 3350 concentration was increased to 30% was added to the hanging drop prior to flash-

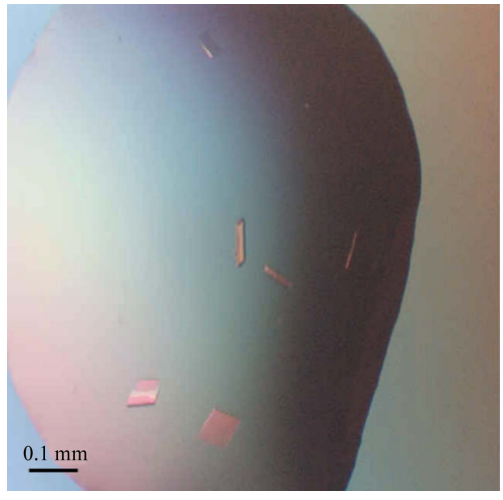

Figure 3
Crystals of DsrEFH.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

X-ray source	Advanced Light Source beamline 5.0.2
X-ray wavelength (Å)	0.9796
Temperature (K)	100
Space group	$P2_1$
Unit-cell parameters	
a (Å)	56.6
b (Å)	183.1
c (Å)	107.8
α (°)	90.0
β (°)	99.6
γ (°)	90.0
Resolution range (Å)	99–2.5 (2.54–2.50)
Total unique reflections	72592 (2988)
R_{sym}† (%)	12.1 (65.3)
Data completeness (%)	97.8 (81.1)
Average $I/\sigma(I)$	13.7 (1.7)
No. of hexamers per ASU	3

† $R_{sym} = \sum_{hkl}\sum_i |I_{hkl,i} - \langle I\rangle_{hkl}|/|I_{hkl}|$.

freezing in liquid nitrogen and exposure to X-rays. X-ray diffraction data sets were collected at a single wavelength at the Macromolecular Crystallography Facility beamline 5.0.2 at the Advanced Light Source at Lawrence Berkeley National Laboratory using a Quantum 4 CCD detector (Area Detector Systems Co., Poway, CA, USA) placed 250 mm from the sample. The oscillation range per image was 1.0°, with no overlap between two contiguous images.

3. Results and discussion

Expression of hexahistidine-tagged fusion protein in *E. coli* and purification by IMAC yielded ∼25 mg DsrEFH per litre of *E. coli* culture. After anion-exchange chromatography, DsrEFH appeared to be approximately 99% pure, with prominent protein bands at 14, 16 and 10 kDa on SDS–PAGE (Fig. 2). In the first crystallization trial, no crystals were observed using various screen solutions. Therefore, optimum-solubility (OS) screening was performed to find an additive to improve the conformational homogeneity of the protein solution. ADA buffer turned out to be the best buffer for this purpose. Various crystals appeared using ADA buffer under several conditions. The best crystal was obtained using PEG 3350 as a precipitant. Plate-shaped crystals grew in a week to approximate dimensions of 0.10 × 0.09 × 0.02 mm (Fig. 3).

Synchrotron data were collected to 2.5 Å. X-ray diffraction data were processed and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). The crystal belongs to the primitive monoclinic space group $P2_1$, with unit-cell parameters $a = 56.6$, $b = 183.1$, $c = 107.8$ Å, $\beta = 99.6°$, with a Matthews coefficient V_M of 2.23 Å^3 Da^{-1} and a solvent content of 42.6% (Matthews, 1968) assuming the asymmetric unit to contain three hexamers. Details of the data-collection statistics are presented in Table 1. A full structure determination using the single- or multi-wavelength anomalous dispersion method is under way in order to provide insight into the structure and possible molecular function of this protein.

The work described here was supported by the Ewha Womans University Research Grant of 2005 and by the Deutsche Forschungsgemeinschaft (grants Da 351/3-3, 3-4 and 3-5 to CD). Skilful technical assistance by Birgitt Hüttig and Jaru Jancarik is gratefully acknowledged.

References

Dahl, C., Engels, S., Pott-Sperling, A. S., Schulte, A., Sander, J., Lübbe, Y., Deuster, O. & Brune, D. C. (2005). *J. Bacteriol.* **187**, 1392–1404.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. (2004). *Acta Cryst.* D**60**, 1670–1673.
Matthews, B. W. (1968). *J. Mol. Biol.* **28**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pott, A. S. & Dahl, C. (1998). *Microbiology*, **144**, 1881–1894.
Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. & Sweet, R. M. (1993). *Nature (London)*, **362**, 219–223.