Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Escherichia coli stress protein YciF: expression, crystallization and preliminary crystallographic analysis

Deqian Liu,[a,b] Yonghong Zhao,[b,c] Xiuzhen Fan,[a,b] Yuan Sun[a,b] and Robert O. Fox[a,b,c]*

[a]The Department of Human Biological Chemistry and Genetics, The University of Texas Medical Branch at Galveston, TX 77555-0647, USA, [b]Sealy Centre for Structural Biology, The University of Texas Medical Branch at Galveston, TX 77555-0647, USA, and [c]Department of Physiology and Biophysics, The University of Texas Medical Branch at Galveston, TX 77555-0647, USA

Correspondence e-mail: fox@bloch.utmb.edu

Received 5 August 2004
Accepted 2 November 2004

The *Escherichia coli* stress protein YciF was overexpressed and purified by three chromatographic steps. Crystals were obtained using ammonium sulfate as a precipitant. The YciF protein crystals diffracted beyond 2.25 Å resolution using a rotating-anode X-ray source. The lattice type is rhombohedral, with unit-cell parameters $a = b = 80.0$, $c = 131.03$ Å, $\alpha = \beta = 90.0$, $\gamma = 120.0^\circ$. The crystal belongs to space group $R32$.

1. Introduction

Microorganisms survive under different environmental stress conditions by expressing sets of proteins that afford protection. Knowing the mechanisms by which stress proteins protect cells may be important for the development of new classes of antimicrobial therapeutics. The *Escherichia coli* YciF protein is encoded by the gene *yciF* (Blattner *et al.*, 1997). YciF, a 19 kDa protein, was accumulated in an H-NS deletion strain, a genetic background known to derepress stress-gene expression (Yoshida *et al.*, 1993). Primary sequence alignment indicated that the YciF protein is conserved across a wide range of bacterial genomes. Interestingly, there is sequence similarity with the *E. coli yciE* gene (Fig. 1), a member of the same operon that also encodes another heat-shock protein (HSP) of unknown function.

Most HSPs function as chaperones (GroEL, GroES *etc.*), proteases (ClpP, HslV *etc.*) or transcription regulators (PhoB, RseA *etc.*), although other enzymatic functions also occur. To obtain a better understanding of the possible molecular function of the YciF stress protein, it was expressed in large quantities, purified to homogeneity and crystallized in preparation for an X-ray crystallographic structure determination. Here, we report the crystallization and preliminary X-ray analysis of the *E. coli* YciF protein.

2. Materials and methods

2.1. Protein preparation

Genomic DNA was purified from *E. coli* K-12 strain MG1655 using a genomic DNA purification kit (Promega). The gene for the YciF protein was cloned by PCR using the oligonucleotide primers 5′-AGGGGCGCC-ATGAATATGAAGACCATTGAAGATG-3′ and 5′-CCGGAATTCTCAGGCTTTATTTT-CGGC-3′ (Sigma-Genosys). The PCR product was digested with the restriction enzymes *Kas*I and *Eco*RI (New England Biolabs) and ligated with pProEX-1 vector treated with the same enzymes. The recombinant plasmid encoded a fusion protein with an N-terminal hexahistidine tag separated from the YciF protein by a tobacco etch virus (TEV) protease-cleavage site and was designated p215 (Phan *et al.*, 2002). After confirmation of the plasmid's *yciF* DNA sequence at the UTMB recombinant DNA laboratory, p215 was transfected into DH10B cells for overexpression.

The cells were grown at 310 K in LB medium containing 50 μg ml^{-1} ampicillin. When the absorbance at 600 nm reached 0.6,

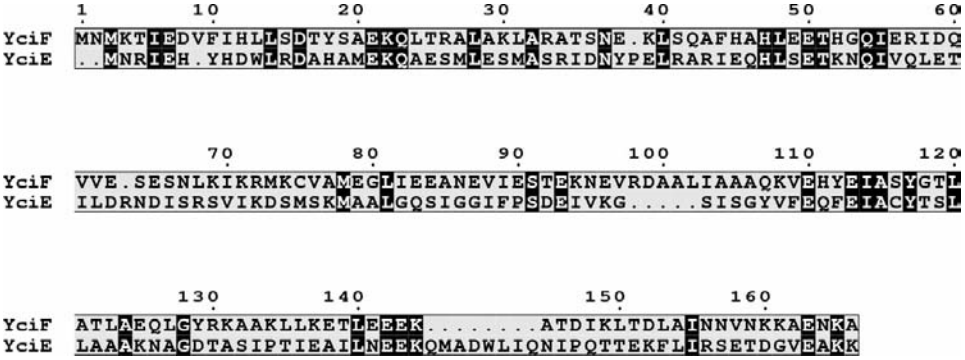

Figure 1
CLUSTALW sequence alignment between YciF and YciE. The two proteins share 22% sequence identity (solid boxes) and 29% similarity (grey boxes).

the culture was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.8 m*M*. 3 h after induction, cells were harvested and disrupted by sonication in buffer *A* (10 m*M* imidazole, 50 m*M* Tris pH 8.2, 300 m*M* NaCl, 2 m*M* 2-mercaptoethanol). Protein purification was performed at room temperature. After centrifugation of the homogenate at 20 000*g* for 30 min, the supernatant was applied onto a nickel–nitrilotriacetic acid (Ni^{2+}–NTA) column (Qiagen). His-tagged YciF was eluted from the column with a gradient of 0–60% buffer *B* (1 *M* imidazole, 50 m*M* Tris pH 8.2, 300 m*M* NaCl, 2 m*M* 2-mercaptoethanol). Protein was dialyzed in 3 × 500 ml buffer *C* (25 m*M* Tris pH 8.2, 100 m*M* NaCl, 2 m*M* 2-mercaptoethanol) to remove imidazole. The protein concentration was measured by absorbance at 280 nm using a specific absorbance of 7680 M^{-1} cm^{-1}, assuming that all Cys residues are reduced. The N-terminal hexahistidine tag was removed by overnight incubation with 0.35 mg of TEV protease per milligram of recombinant protein. The cleaved protein was loaded onto an Ni^{2+}–NTA column to remove the cleaved His tag and the TEV protease, which also contained a His tag (Kapust *et al.*, 2001). The flowthrough of the second Ni^{2+}–NTA column was concentrated using a Centricon YM-10 (Millipore) and further purified on a Superdex 200 column (Amersham Pharmacia Biotech) equilibrated with buffer *C*. The homogeneity of the purified preparation was judged by SDS–PAGE and mass spectrometry. The typical yield of cleaved and purified YciF protein was about 10 mg per litre of culture.

2.2. Crystallization and data collection

The purified YciF protein was concentrated to 10.0 mg ml^{-1} in sample buffer (25 m*M* Tris pH 8.2, 100 m*M* NaCl and 5 m*M* DTT) and crystallized using the hanging-drop vapor-diffusion technique. Preliminary screening was carried out using a HYDRA-PLUS-ONE Robotic System (Apogent Technologies Inc.) with Hampton Reaseach HT screen kits. A few conditions such as E12 (0.1 *M* sodium acetate, 0.1 *M* $CdCl_2$, 30% PEG 400) of Crystal Screen HT and A1 (0.1 *M* citric acid pH 3.5, 2.0 *M* ammonium sulfate) and F2 (0.2 *M* TMAO, 0.1 *M* Tris pH 8.5, 20% PEG MME 2000) of Index HT yielded small crystals. After optimization of crystallization conditions, crystals were obtained from 0.1 *M* citric acid pH 3.5–4.0, 1.9–2.1 *M* ammonium sulfate (Fig. 2). A crystal was step-soaked into a cryoprotectant consisting of 0.1 *M* citric acid pH 3.5, 2.0 *M* ammonium sulfate and 30%(*v*/*v*) glycerol, scooped up in a cryoloop, rapidly cooled in liquid nitrogen and mounted on the goniometer in a nitrogen stream at 100 K. X-ray diffraction was measured in-house on a MacScience X-ray rotating-anode generator with MSC Blue confocal multilayer optics, using a DIP2030 imaging plate for macromolecular crystallography with a crystal-to-detector distance of 150 mm. Diffraction data were collected from a YciF crystal to 2.25 Å resolution and were indexed, integrated and scaled with *DENZO* and *SCALEPACK* from the *HKL* package (Otwinowski & Minor, 1997).

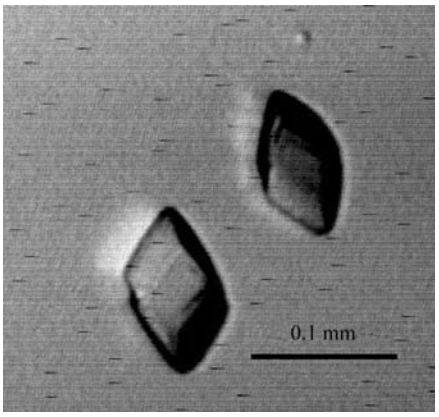

Figure 2
The typical size of diffraction-quality YciF crystals is approximately 0.1 × 0.07 mm.

3. Results and discussion

The TEV-cleaved recombinant protein left two vector-derived residues (Gly-Ala) at the N-terminus of the endogenous YciF protein, with a calculated molecular weight of 18 725.2 Da (168 residues). YciF protein eluted on a Superdex 200 size-exclusion column as a homodimer. The purified YciF protein migrated as a single 19 kDa band with >95% homogeneity on 15% SDS–PAGE gel.

The YciF protein crystals diffracted to beyond 2.25 Å resolution on a rotating-anode X-ray generator with confocal optics. The lattice type was rhombohedral, with unit-cell parameters $a = b = 80.01$, $c = 131.03$ Å, $\alpha = \beta = 90.0$, $\gamma = 120^\circ$. The crystal belongs to space group *R*32. Assuming the presence of one molecule per asymmetric unit, the calculated Matthews coefficient V_M value was 2.15 $Å^3$ Da^{-1}, with a solvent content of 42.9% (Matthews, 1968). Data-collection statistics are given in Table 1. Structure determination of the YciF protein is under way.

Table 1
Data-collection statistics.

Values in parentheses are for the last resolution shell (2.30–2.25 Å).

Space group	*R*32
Unit-cell parameters (Å)	
$a = b$	80.0
c	131.03
Temperature (K)	100
Wavelength (Å)	1.54178
Resolution (Å)	30.0–2.25
Total No. reflections	41134
No. unique reflections	7919
Completeness (%)	99.7 (99.4)
R_{merge}	0.076 (0.209)
Mean $I/\sigma(I)$	18.7 (12.0)

The sequence similarity of YciE and YciF (29%) suggests that the proteins may have the same overall fold but different functions. Further, the juxtaposition of the two genes within one operon suggests they may occur sequentially in a functional pathway. Comparison of the two structures will provide a powerful restraint on hypotheses concerning the function of these two proteins. The yciE crystals also occurred in the *R*32 space group, but with poorer diffracting power and somewhat different unit-cell parameters of $a = b = 64.2$, $c = 167.9$ Å, $\alpha = \beta = 90.0$, $\gamma = 120.0^\circ$ (Liu *et al.*, 2004).

We thank Dr Mark A. White for maintaining the X-ray diffraction facility and for helpful discussions. We thank Dr Yousif Shamoo for access to the HYDRA-PLUS-ONE Robotic System at Rice University. We thank Dr David Waugh for the TEV expression construct and Dr David Konkel for editing the manuscript. A Robert A. Welch Foundation award to ROF (H-1345) funded this work.

References

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997). *Science*, **277**, 1453–1474.

Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D. & Waugh, D. S. (2001). *Protein Eng.* **14**, 993–1000.

Liu, D., Zhao, Y. H., Fan, X., Sun, Y. & Fox, R. O. (2004). *Acta Cryst.* D**60**, 1888–1889.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Phan, J., Zdanov, A., Evdokimov, A. G., Tropea, J. E., Peters, H. K. III, Kapust, R. B., Li, M., Wlodawer, A. & Waugh, D. S. (2002). *J. Biol. Chem.* **277**, 50564–50572.

Yoshida, T., Ueguchi, C., Yamada, H. & Mizuno, T. (1993). *Mol. Gen. Genet.* **237**, 113–122.