Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic analysis of the surE protein from *Thermotoga maritima*

Jae Eun Kwak,[a]† Kook Sun Ha,[a]† Jae Young Lee,[a] Young Jun Im,[b] Seong Ho Park,[b] Soo Hyun Eom[b] and Se Won Suh[a]*

[a]School of Chemistry and Molecular Engineering, Seoul National University, Seoul 151-742, South Korea, and [b]Department of Life Science, Kwangju Institute of Science and Technology, Kwangju 500-712, South Korea

† These authors contributed equally to this work.

Correspondence e-mail: sewonsuh@snu.ac.kr

© 2001 International Union of Crystallography
Printed in Denmark – all rights reserved

The surE protein from *Thermotoga maritima* is a 247-residue protein of unknown function. Its homologues are well conserved among both the eubacteria and the archaea. It has been overexpressed in soluble form in *Escherichia coli*. The protein has been crystallized at 296 K using 2-propanol as a precipitant. X-ray diffraction data have been collected to 1.9 Å resolution using synchrotron radiation. The crystals belong to the trigonal space group $P3_121$ (or $P3_221$), with unit-cell parameters $a = b = 115.96$, $c = 78.60$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. The asymmetric unit contains two monomers of the surE protein, with a corresponding V_M of 2.72 Å^3 Da^{-1} and a solvent content of 54.7%.

Received 16 October 2000
Accepted 31 January 2001

1. Introduction

A new open reading frame (named *surE*) was discovered in *E. coli* by sequencing the upstream region of the *pcm* gene, which is located at 59 min on the *E. coli* chromosome (Li *et al.*, 1994, 1997). The *E. coli surE* gene encodes a protein of 253 amino-acid residues and overlaps the 5′-end of the *pcm* gene by four base pairs (Li *et al.*, 1994). The *pcm* gene product, protein-L-isoaspartyl O-methyltransferase, plays an essential role in the repair of damaged proteins by catalyzing the transfer of a methyl group from *S*-adenosylmethionine to atypical L-isoaspartyl but not to normal L-aspartyl residues in proteins (Clarke, 1985). On the other hand, no biochemical function or physiological role has yet been identified for the surE protein. It was proposed, however, that a functional relationship exists between the *surE* and *pcm* gene products (Visick *et al.*, 1998). Isoaspartyl accumulation during the stationary phase was much higher in a *pcm surE* double mutant than in either single mutant, suggesting that the two genes may represent two parallel pathways by which *E. coli* can respond to protein damage (Visick *et al.*, 1998).

SurE, like *pcm*, is an ancient gene and is well conserved among both the eubacteria and the archaea, with the notable exception of the Gram-positive bacteria and the mycoplasmas (Visick *et al.*, 1998). Only one eukaryotic sequence with significant similarity to *surE* has been identified, in *Arabidopsis thaliana* (Visick *et al.*, 1998). Since possible clues to the biochemical function of hypothetical proteins with unknown functions may come from the three-dimensional structure as exemplified by the work of Hwang *et al.* (1999), we have initiated the structure determination of the surE homolog in *T. maritima*. It comprises 247 amino-acid residues and the amino-acid sequence is 39% identical to that of *E. coli*. As a first step toward its structure determination, we report here the overexpression, crystallization and preliminary X-ray crystallographic data of the recombinant surE protein from *T. maritima*.

2. Experimental

2.1. Protein expression and purification

The *surE* gene was amplified by the polymerase chain reaction using the *T. maritima* genomic DNA as template. The amplified DNA was inserted into the *Nde*I/*Bam*HI-digested expression vector pET-21a. The protein was overexpressed in soluble form in C41(DE3) cells by induction with 0.5 m*M* isopropyl β-D-thiogalactopyranoside (IPTG) at 303 K. Cells were grown in Luria–Bertani medium for 6 h after IPTG induction and were harvested by centrifugation at 4200*g* (Sorvall GS3 rotor) for 10 min at 277 K. The cell pellet was resuspended in ice-cold lysis buffer (50 m*M* Tris–HCl pH 7.0, 0.20 *M* sodium chloride and 1 m*M* β-mercaptoethanol) and was then homogenized using a French press. The crude lysate was centrifuged at 36 000*g* (18 000 rev min^{-1}; Hanil Supra 21K rotor) for 50 min at 277 K. The supernatant fraction was heated and kept between 348 and 353 K for 10 min and then placed in an ice bath for 10 min. The cell extract was centrifuged at 36 000*g* (18 000 rev min^{-1}; Hanil Supra 21K rotor) for 50 min at 277 K. The supernatant was loaded onto a Q-Sepharose anion-exchange column (Amersham-Pharmacia), which was previously equilibrated with buffer *A* (50 m*M* Tris–HCl pH 7.0, 2 m*M* β-mercaptoethanol) and the protein was eluted with a linear gradient of 0–0.5 *M* NaCl. Next, gel

filtration was performed on a HiLoad XK 16 Superdex 75 prep-grade gel-filtration column (Amersham Pharmacia), which was previously equilibrated with buffer *A* containing 200 m*M* sodium chloride. The purified protein was homogeneous as assessed by polyacrylamide gel electrophoresis in the presence of 0.1%(*w*/*v*) sodium dodecyl sulfate (Laemmli, 1970). This procedure yielded approximately 17 mg of homogeneous surE protein from a 2 l culture. The protein solution was concentrated using a YM10 ultrafiltration membrane (Amicon) to about 15 mg ml^{-1}. The protein concentration was estimated by measuring the absorbance at 280 nm using the calculated molar extinction coefficient of 34 400 M^{-1} cm^{-1} (SWISS-PROT; http://www.expasy.ch/).

2.2. Crystallization and X-ray diffraction experiment

Crystallization was achieved by the hanging-drop vapour-diffusion method at 296 K using 24-well tissue-culture plates (Hampton Research). Each hanging drop was prepared by mixing 2 µl of the reservoir solution and 2 µl of the protein solution (15 mg ml^{-1}). Each hanging drop was placed over 1.0 ml of the reservoir solution. Initial crystallization conditions were established by sparse-matrix sampling (Jancarik & Kim, 1991).

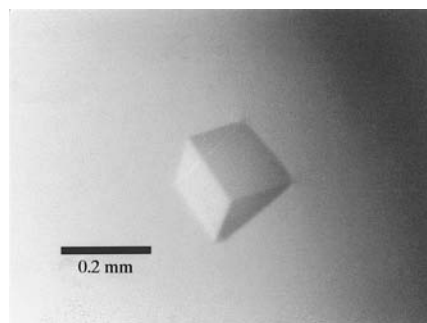

Figure 1
A trigonal crystal of surE protein from *T. maritima*. Its approximate dimensions are 0.3 × 0.2 × 0.2 mm.

A crystal of surE protein was transferred to a solution of 100 m*M* HEPES pH 7.5, 200 m*M* $MgCl_2$, 30%(*v*/*v*) 2-propanol and 30%(*v*/*v*) glycerol within a minute in two steps, thus increasing the concentration of glycerol prior to being flash-frozen. X-ray diffraction data were collected at 100 K with an ADSC Quantum 4R CCD detector at the X8-C experimental station of the National Synchrotron Light Source, Brookhaven National Laboratory. The wavelength of the synchrotron radiation was 1.000 Å. The crystal was rotated through a total of 120°, with a 1.0° oscillation range per frame. The raw data were processed and scaled using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

3. Conclusions

The recombinant surE protein from *T. maritima* was overproduced in soluble form with a yield of ~17 mg of homogeneous protein from a 2 l culture. The trigonal crystals grew to dimensions of 0.3 × 0.2 × 0.2 mm within 3 d from a reservoir solution containing 100 m*M* HEPES pH 7.5, 200 m*M* $MgCl_2$ and 30%(*v*/*v*) 2-propanol (Fig. 1). Without $MgCl_2$, the protein crystallized into extremely thin needle-like crystals. With other cations, such as Ca^{2+}, Mn^{2+} and Zn^{2+}, the protein gave only precipitates. The native diffraction data were collected to 1.9 Å resolution at 100 K using synchrotron radiation. A total of 470 499 measured reflections were merged into 47 996 unique reflections with an R_{merge} (on intensity) of 9.3%. The merged data set is 99.8% complete to 1.9 Å resolution. The crystals belong to the trigonal space group $P3_121$ (or $P3_221$), with unit-cell parameters $a = b = 115.96$ (0.17), $c = 78.60$ (0.19) Å, $\alpha = \beta = 90$, $\gamma = 120°$, where estimated standard deviations are given in parentheses. The presence of two monomers of the surE protein in the asymmetric unit gives a crystal volume per protein mass (V_M) of 2.72 $Å^3$ Da^{-1}, with a corresponding solvent content of 54.7% (Matthews, 1968). Table 1 summarizes the statistics for the synchrotron data collection.

Table 1
Synchrotron data-collection statistics.

Values in parentheses refer to the highest resolution shell (1.97–1.90 Å).

Space group	$P3_121$ (or $P3_221$)
Unit-cell parameters (Å,°)	$a = b = 115.96$ (0.17), $c = 78.60$ (0.19), $\alpha = \beta = 90$, $\gamma = 120$
Resolution range (Å)	50.0–1.90
No. of observed reflections	470499
No. of unique reflections	47996
Data completeness (%)	99.8 (99.9)
R_{merge}† (%)	9.3 (34.4)
Average $I/\sigma(I)$	7.6 (3.3)

† $R_{merge} = \sum_h \sum_i |I(h)_i - \langle I(h) \rangle| / \sum_h \sum_i I(h)_i$, where $I(h)$ is the intensity of reflection h, $\sum_h$ is the sum over all reflections and $\sum_i$ is the sum over i measurements of reflection h.

This work was supported by grant No. 305-20000004 from the Basic Research Program of the Korea Science and Engineering Foundation and the Center for Molecular Catalysis at Seoul National University. SWS is supported by the BK21 program. JEK and KSH are recipients of a BK21 fellowship.

References

Clarke, S. (1985). *Annu. Rev. Biochem.* **54**, 479–506.
Hwang, K. Y., Chung, J. H., Kim, S.-H., Han, Y. S. & Cho, Y. (1999). *Nature Struct. Biol.* **6**, 691–696.
Li, C., Ichikawa, J. K., Ravetto, J. J., Kuo, H.-C., Fu, J. C. & Clarke, S. (1994). *J. Bacteriol.* **176**, 6015–6022.
Li, C., Wu, P.-Y. & Hsieh, M. (1997). *Microbiology*, **143**, 3513–3520.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–493.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Visick, J. E., Ichikawa, J. K. & Clarke, S. (1998). *FEMS Microbiol. Lett.* **167**, 19–25.