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

Purification, crystallization and preliminary crystallographic analysis of a thermostable endonuclease IV from *Thermotoga maritima*

Ronny C. Hughes,[a,b] Stephen J. Tomanicek,[a] Joseph D. Ng[b] and Leighton Coates[a]*

[a]Oak Ridge National Laboratory, Neutron Scattering Science Division, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA, and [b]Laboratory for Structural Biology, Department of Biological Sciences, University of Alabama in Huntsville, Huntsville, AL 35899, USA

Correspondence e-mail: coatesl@ornl.gov

Received 25 October 2009
Accepted 9 November 2009

The DNA-repair enzyme endonuclease IV from the thermophilic bacterium *Thermotoga maritima* MSB8 (reference sequence NC_000853) has been expressed in *Escherichia coli* and crystallized for X-ray analysis. *T. maritima* endonuclease IV is a 287-amino-acid protein with 32% sequence identity to *E. coli* endonuclease IV. The protein was purified to homogeneity and was crystallized using the sitting-drop vapor-diffusion method. The protein crystallized in space group $P6_1$, with one biological molecule in the asymmetric unit, corresponding to a Matthews coefficient of 2.39 Å^3 Da^{-1} and 47% solvent content. The unit-cell parameters of the crystals were $a = b = 123.2$, $c = 35.6$ Å. Microseeding and further optimization yielded crystals with an X-ray diffraction limit of 2.36 Å. A single 70° data set was collected and processed, resulting in an overall $R_{\rm merge}$ and a completeness of 9.5% and 99.3%, respectively.

1. Introduction

The endonuclease IV family is a major component of the cellular DNA-repair pathway that includes apurinic/apyrimidinic endonuclease activity and repair activities on 3′-phosphates, 3′-phosphoglycolates and 3′-*trans*-4-hydroxy-2-pentenal-5-phosphates. *Thermotoga maritima* endonuclease IV exhibits enzymatic activity at both low and high temperatures and is thermally more stable than that of *Escherichia coli*, with denaturation temperatures that approach 363 K for the *T. maritima* enzyme (Haas *et al.*, 1999). The most common lesion that occurs in DNA is at apurinic/apyrimidinic (AP) sites. These AP sites are generated by the action of DNA glycosylases, which remove mismatched and modified bases from damaged DNA (Krokan *et al.*, 1997; Cunningham, 1997). AP sites can also be generated by direct interactions between DNA and reactive oxygen species (Hutchinson, 1985). AP sites are non-instructive to DNA polymerases and can be mutagenic if left unrepaired (Loeb & Preston, 1986). The repair of AP sites is initiated by a class of enzymes referred to as AP endonucleases. These enzymes hydrolytically cleave the 5′-phosphodiester bond at an AP site to generate a 3′-hydroxyl group and a 5′-terminal sugar phosphate. A deoxyribophosphodiesterase enzyme can then remove the terminal 5′ sugar phosphate to leave a single nucleotide gap that can be repaired by a DNA polymerase and a DNA ligase to restore the proper genetic information. It has been shown that *T. maritima* endonuclease IV possesses both phosphomonoesterase and phosphodiesterase activity that is utilized to repair broken DNA strands (Haas *et al.*, 1999). These reactions are used to repair the oxidative DNA damage that can take place when reactive oxygen species (hydroxyl radicals, hydrogen peroxide) interact with DNA. Common sources for the generation of these radicals include ionizing radiation and aerobic metabolism and they can lead to the induction of single-stranded breaks in DNA (Halliwell & Aruoma, 1991). These strand breaks contain 3′-blocking groups such as phosphates and phosphoglycolates that inhibit DNA replication (Demple *et al.*, 1986). Removal of these 3′-blocking groups has been demonstrated using *T. maritima* endonuclease IV (Haas *et al.*, 1999).

2. Materials and methods

2.1. Protein expression and purification

The recombinant protein was purified using an ÄKTA Purifier system (GE Healthcare). All steps were performed at room temperature (295 K). The synthetic gene encoding the full-length endonuclease IV protein (287 amino acids) from *T. maritima* (Q9WYJ7) was codon-optimized by iXpressGenes Inc. for *E. coli* expression and was inserted into a pET3a expression plasmid (Novagen) at the *Bam*HI and *Nde*I restriction sites by *in vivo* homologous recombination in DH5α competent *E. coli* (Hughes & Ng, 2007). This plasmid was then transformed into competent BL21 expression cells (Invitrogen) for protein overexpression. The cells were grown in selective Luria–Bertani (LB) Miller broth to an OD_{595} of 0.6 before being induced by the addition of 1 *M* isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 m*M*. The cells were grown for 16 h at 298 K before being harvested by centrifugation. Protein purification was performed using established protocols (Haas *et al.*, 1999). In brief, the cells were lysed using ultrasonification in 50 m*M* 3-(*N*-morpholino)propanesulfonic acid (MOPS) buffer pH 8.0 containing 50 m*M* NaCl and 1 m*M* ethylenediaminetetraacetic acid (EDTA). The lysed cells were then centrifuged at 15 000*g* for 30 min to remove cellular debris and the supernatant was then subjected to a heat cut at 341 K for 20 min. The resulting suspension was centrifuged again at 15 000*g* to remove denatured thermally labile proteins. The clarified supernatant was then loaded onto a Q-Sepharose ion-exchange column (GE Healthcare) pre-equilibrated with the previously described lysis buffer. The column was then washed with a linear gradient of 0–1 *M* NaCl in the same buffer. The endonuclease IV protein eluted in the fractions containing 300–350 m*M* NaCl. The protein composition of these fractions was analyzed by SDS–PAGE. Polyethyleneimine precipitation was then used to remove unwanted nucleic acid contaminants to a final percentage fraction of less than 1% as judged by the UV absorptivity at 280 and 260 nm. The resulting precipitate was removed by centrifugation as described above and the resulting supernatant was concentrated to 2 ml prior to injection onto a Superdex S200 gel-filtration column (GE Healthcare) to remove EDTA and other low-molecular-weight contaminants. The buffer used for gel filtration was 50 m*M* MOPS buffer pH 8.0 containing 50 m*M* NaCl. Fractions from gel filtration with appropriate purity as judged by SDS–PAGE were pooled and concentrated to 10 mg ml^{-1} for crystallization trials. The homogeneity of the purified protein was confirmed by SDS–PAGE using Coomassie staining as shown in Fig. 1.

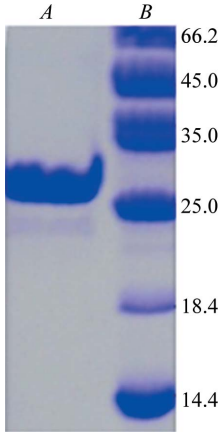

Figure 1
SDS–PAGE of purified *T. maritima* endonuclease IV protein (32.4 kDa). Lane *A*, protein sample (15 µg loaded); lane *B*, low-molecular-weight makers. SDS–PAGE was performed in a 12%(*w*/*v*) polyacrylamide gel. EZ-run low-molecular-weight standards (Fisher Bioreagents) were used as molecular-weight standards (labeled in kDa) and the proteins were detected by Coomassie Brilliant Blue staining.

2.2. Crystallization

Crystallization was performed using the sitting-drop vapor-diffusion technique at 295 K. A 4 µl protein droplet was set to equilibrate against 30 µl precipitating agent in a 96 × 2-well IntelliPlate (Art Robbins) sealed with transparent adhesive tape. The protein droplet was prepared by mixing equal volumes of the protein solution (10 mg ml^{-1}) and precipitating agent. Initial crystallization screens were explored using the Hampton Screen High-Throughput (HSHT) reagents (Hampton Research). Varying protein concentrations and buffers were used in screens corresponding to 10 and 20 mg ml^{-1} protein in three alternate buffers (sodium acetate pH 4.6, Tris pH 8.5 and sodium citrate pH 3.5) prepared at 50 m*M* concentration and containing 50 m*M* NaCl using 1:1, 1:2 and 2:1 protein:precipitant ratios.

Microcrystals were observed after 4 d of equilibration against HSHT condition G10 (0.05 *M* cadmium sulfate hydrate, 0.1 *M* HEPES pH 7.5, 1.0 *M* sodium acetate trihydrate). The microcrystalline material was crushed by physical disruption using a pipette tip and resuspended in 1 µl seed-stock solution (0.05 *M* cadmium sulfate hydrate, 0.1 *M* HEPES pH 7.5, 1.0 *M* sodium acetate trihydrate) for subsequent macroseeding experiments as described by D'Arcy *et al.* (2007). The seeded solution was further mixed with 99 µl purified protein solution at a concentration of 10 mg ml^{-1} in 50 m*M* MOPS buffer pH 8.0 containing 50 m*M* NaCl. The resulting protein solution was again screened against the HSHT reagents as described previously. Single prismatic crystals were then obtained directly from HSHT screen condition H10 [0.1 *M* sodium chloride, 0.1 *M* Bicine pH 9.0, 20%(*v*/*v*) polyethylene glycol monomethyl ether 550]. The crystal used for data collection is shown in Fig. 2.

2.3. X-ray diffraction and initial phasing

The crystal diffracted to 2.36 Å resolution as determined by data completeness. A single crystal cryoprotected with mother liquor containing 30% glycerol was used to collect X-ray diffraction data over a range of 70° (1° oscillation) on the SER-CAT 22-BM beamline (λ = 1.00 Å) at the Advanced Photon Source (Chicago, USA). Diffraction images were recorded with a MAR Research MX-300 detector and the data were indexed and scaled in space group $P6_1$ using *HKL*-2000 (Otwinowski & Minor, 1997). The data-collection statistics are shown in Table 1.

Figure 2
The endonuclease crystal used for data collection is shown mounted in a nylon loop. The longest dimension was measured to be 0.1 mm.

Table 1
X-ray data-collection statistics.

Values in parentheses are for the outer shell.

X-ray source, beamline	Synchrotron, APS 22-BM
Crystal volume (mm^3)	5×10^{-4}
Crystallization method	Vapor diffusion
Data-collection temperature (K)	100
Wavelength (Å)	1.0
No. of frames	70
Oscillation range (°)	1
Exposure time (s)	10
Beam transmission (%)	100
Space group	$P6_1$
Unit-cell parameters (Å)	a = 123.2, b = 123.2, c = 35.6
Molecules per ASU	1
Solvent content (%)	46.7
Resolution (Å)	50.0–2.36 (2.44–2.36)
No. of reflections	55429
No. of unique reflections	12873
Multiplicity	4.3 (4.1)
Crystal mosaicity (°)	0.19
Data completeness (%)	99.3 (95.6)
$R_{\rm merge}$† (%)	9.5 (17.4)
$\langle I/\sigma(I)\rangle$	13.1 (6.19)
Wilson plot B factor (Å^2)	19.6

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$.

The protein structure was solved using molecular replacement with the *Phaser* program (McCoy *et al.*, 2007) from the *PHENIX* suite (Adams *et al.*, 2002). The known structure of *E. coli* endonuclease IV (PDB code 1qum; Hosfield *et al.*, 1999), which has 32% sequence homology to *T. maritima* endonuclease IV, was used as the search model. Further work to model and refine the structure using *PHENIX* (Adams *et al.*, 2002) is in progress.

3. Results and discussion

A thermally stable homologue of the DNA-repair enzyme endonuclease IV from the thermophilic bacterium *T. maritima* was expressed in *E. coli* and purified to greater than 95% homogeneity as determined by SDS–PAGE. No significant nucleic acid contamination was apparent; the final UV absorbance ratio A_{280}/A_{260} was measured to be 1.92. The resulting full-length protein was crystallized and a single crystal was used for X-ray analysis, yielding a complete data set with a completeness of 99.3% and an overall $R_{\rm merge}$ of 9.5%. The crystal belonged to space group $P6_1$, with unit-cell parameters a = 123.2, b = 123.2, c = 35.6 Å. Based on diffraction data analysis using the program *XPREP* (Bruker AXS Inc.), space groups $P6_1$ and $P6_5$ were equally likely candidates. The $P6_1$ space group was distinguished as a result of subsequent phasing by molecular replacement. The asymmetric unit was determined to contain a single biological molecule and 46.7% solvent based on a Matthews coefficient of 2.39 Å^3 Da^{-1}. Despite a low sequence identity of only 32% between the *E. coli* structure (PDB code 1qum; Hosfield *et al.*, 1999) and that of *T. maritima*, a molecular-replacement solution was obtained using *Phaser* (McCoy *et al.*, 2007) with a log-likelihood gain of 111.34, indicating a clear solution.

This research was sponsored in part by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory (ORNL), managed by UT-Battelle LLC for the US Department of Energy under Contract No. DE-AC05-00OR22725. Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID (or 22-BM) beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at http://www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. This research at Oak Ridge National Laboratory's Center for Structural Molecular Biology (CSMB) was supported by the Office of Biological and Environmental Research, using facilities supported by the US Department of Energy, managed by UT-Battelle, LLC under contract No. DE-AC05-00OR22725.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Cunningham, R. P. (1997). *Mutat. Res.* **383**, 189–196.
D'Arcy, A., Villard, F. & Marsh, M. (2007). *Acta Cryst.* D**63**, 550–554.
Demple, B., Johnson, A. & Fung, D. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 7731–7735.
Haas, B. J., Sandigursky, M., Tainer, J. A., Franklin, W. A. & Cunningham, R. P. (1999). *J. Bacteriol.* **181**, 2834–2839.
Halliwell, B. & Aruoma, O. I. (1991). *FEBS Lett.* **281**, 9–19.
Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P. & Tainer, J. A. (1999). *Cell*, **98**, 397–408.
Hughes, R. C. & Ng, J. D. (2007). *Cryst. Growth Des.* **7**, 2226–2238.
Hutchinson, F. (1985). *Prog. Nucleic Acid Res. Mol. Biol.* **32**, 115–154.
Krokan, H. E., Standal, R. & Slupphaug, R. (1997). *Biochem. J.* **325**, 1–16.
Loeb, L. A. & Preston, B. D. (1986). *Annu. Rev. Genet.* **20**, 201–230.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.