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

Crystallization and preliminary structure determination of *Escherichia coli* Mfd, the transcription-repair coupling factor

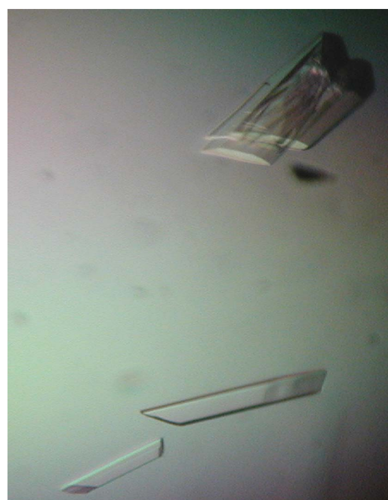
Alexandra M. Deaconescu and Seth A. Darst*

Laboratory of Molecular Biophysics, Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

Correspondence e-mail: darst@rockefeller.edu

Received 15 September 2005
Accepted 1 November 2005
Online 24 November 2005

Transcription-repair coupling factors (TRCFs) are SF2 ATPases that couple transcription to DNA-damage repair by recognizing and removing RNA polymerase-elongation complexes stalled at DNA lesions and recruiting the nucleotide excision-repair machinery to the damaged sites. As a first step towards understanding the TRCF mechanism, the 130 kDa *Escherichia coli* TRCF (the product of the *mfd* gene) has been overexpressed, purified and crystallized using an unusual precipitant, pentaerythritol ethoxylate. Initial phases were obtained using single-wavelength anomalous dispersion with a highly redundant 4 Å resolution data set collected from selenomethionyl-substituted crystals and dramatically improved by density modification and phase extension to 3.2 Å resolution. Model building and refinement, which are in progress, will provide insight into transcription-coupled DNA-repair pathways, as this represents the first TRCF to be crystallized to date.



© 2005 International Union of Crystallography
All rights reserved

1. Introduction

Transcription-repair coupling factor (TRCF, the product of the *mfd* gene) is a widely conserved bacterial protein that couples DNA repair with transcription (Selby & Sancar, 1993). TRCF recognizes RNA polymerase (RNAP) stalled at DNA lesions, disrupts the transcription complex to release the transcript and enzyme and recruits the nucleotide excision-repair machinery to the site. The mechanism of RNA release has been illuminated by the discovery that TRCF causes forward translocation of RNAP through an ATP-dependent motor that is homologous to that of the Holliday branch-migration protein RecG (Park *et al.*, 2002).

TRCF is a large (130 kDa) multi-functional protein with a complex structure–function relationship that is currently understood only from sequence analysis and genetic manipulation (Selby & Sancar, 1993, 1995*a*,*b*). A high-resolution crystal structure of TRCF would provide a firm structural basis to design more incisive experiments exploring the interactions between TRCF, the RNAP ternary elongation complex and the nucleotide excision-repair machinery. As a first step towards understanding the structural basis for transcription-coupled repair, *Escherichia coli* TRCF has been overexpressed, purified and crystallized. Initial crystallographic analysis has provided an excellent electron-density map at 3.2 Å resolution.

2. Materials and methods

2.1. Protein overexpression and purification

The *E. coli mfd* gene was amplified from plasmid pMFD19 and subcloned between the *Nhe*I/*Eco*RI sites into a pET28a-based vector as a hexahistidine-fusion protein to generate pAD6. Transformed Rosetta(DE3)pLysS cells were grown to an OD_{600} of ~0.6 in the presence of 50 µg ml^{-1} kanamycin at 310 K and expression was induced with 1 m*M* isopropyl β-D-thiogalactopyranoside. After induction for 4 h at 303 K, cells were harvested by centrifugation, resuspended in buffer IMAC-15 (50 m*M* Tris pH 8.0, 500 m*M* NaCl, 15 m*M* imidazole, 10% glycerol, 2 m*M* β-mercaptoethanol) supplemented with 1 m*M* PMSF and EDTA-free protease-inhibitor cocktail (Sigma) and lysed using a French press. For purification, the clarified lysate was loaded onto an Ni^{2+}-chelating HiTrap column (Amersham

Biosciences) and eluted using a 15–200 m*M* imidazole linear gradient. After overnight digestion of the His-tag with Prescission protease (Amersham Biosciences) and dialysis against buffer IMAC-15, a second subtractive Ni^{2+}-chelating chromatographic step was used to remove uncleaved protein and a tandem GST-affinity column (Amersham Biosciences) was used to remove the protease. The flowthrough was dialysed against low-salt buffer (20 m*M* Tris pH 8.0, 75 m*M* NaCl, 10% glycerol, 10 m*M* β-mercaptoethanol) supplemented with 5 m*M* EDTA to ensure removal of any residual bound nucleotide. The sample was further purified on a Heparin HiTrap column (Amersham Biosciences) using a linear gradient from 100 m*M* to 2 *M* NaCl and finally by size-exclusion chromatography on an SD200 column (Amersham Biosciences) in a buffer consisting of 20 m*M* Tris pH 8.0, 0.5 *M* NaCl and 10 m*M* DTT. For overproduction of selenomethionyl-substituted protein, the *E. coli* Met-auxotroph strain JB(DE3) (a gift from W. Hendrickson) was transformed with pAD6. Protein was overproduced and purified using the same protocol as for native TRCF. The final protein-storage buffer consisted of 20 m*M* Tris pH 8.0, 0.5 *M* NaCl and 20 m*M* DTT.

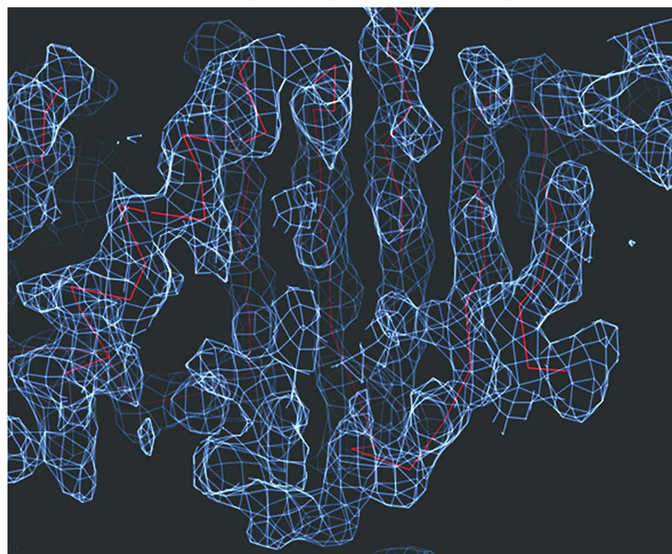
Table 1
Diffraction, data-collection and reduction statistics for TRCF selenomethionyl-substituted crystals.

Values in parentheses are for the highest resolution shell.

	SeMet1	SeMet2
Space group	*C*2	*C*2
Molecules per ASU	2	2
Unit-cell parameters (Å, °)	*a* = 151.7, *b* = 162.2, *c* = 162.1, β = 104.9	*a* = 151.9, *b* = 162.0, *c* = 161.7, β = 105.1
Wavelength (Å)	0.98166	0.97899
Resolution (Å)	40–4.0 (4.14–4.0)	40–3.2 (3.31–3.20)
Reflections measured	228334	300181
Unique reflections	62430	122388
Completeness (%)	99.6 (100)	99.9 (99.9)
Redundancy	3.7 (3.7)	2.5 (2.0)
$\langle I/\sigma(I)\rangle$	10.9 (2.9)	12.4 (2.5)
R_{sym}† (%)	12.9 (50.2)	0.08 (37.7)
SAD figure of merit	0.17	

† $R_{sym} = |I - \langle I\rangle|/\langle I\rangle$, where *I* is the observed intensity and $\langle I\rangle$ is the average intensity of multiple observations of symmetry-related reflections.

(*a*)



(*b*)

Figure 1
(*a*) Crystals of selenomethionyl-substituted TRCF. (*b*) Section of the 3.2 Å resolution electron-density map contoured at 1σ after solvent modification and phase extension using *SOLOMON* (Abrahams & Leslie, 1996). Preliminary model building is shown as a C^{α} trace in red.

2.2. Crystallization

Preliminary crystallization trials were conducted with Crystal Screens I and II, Natrix, PEG/Ion and Index Screens (Hampton Research), which resulted in the identification of a candidate condition in cocktail No. 57 of the Index Screen. After optimization, TRCF crystals suitable for structure determination were grown at 277 K using hanging-drop vapor diffusion against 1 ml 100 m*M* HEPES pH 7.5, 75 m*M* ammonium sulfate, 28% pentaerythritol ethoxylate with a protein concentration of 10 mg ml^{-1} and a 1:1 protein:crystallant ratio. Pentaerythritol ethoxylate, a branched polymer built on the pentaerythritol backbone, is a precipitating agent similar to the low-molecular-weight polyethylene glycols and is a relatively new addition to the battery of protein-crystallization reagents (Gulick *et al.*, 2002). With this precipitant, long clustered plates appeared after 4 d, grew to dimensions of 20 × 40 × 400 µm (Fig. 1*a*) and typically diffracted to ∼5 Å Bragg spacing on an in-house rotating-anode X-ray generator/area detector at cryogenic conditions. These crystals belonged to the monoclinic space group *C*2, had 68% solvent content and contained two molecules (2302 residues) per asymmetric unit. Standard derivatization methods with tantalum and tungsten heavy-metal clusters or with mercury compounds ($HgCl_2$, methylmercury, ethylmercury) destroyed crystal diffraction. Therefore, selenomethionyl-substituted protein was prepared and this yielded isomorphous crystals which were typically thicker and diffracted to higher resolution. Cryoprotection was achieved by transferring the crystals for a few seconds into 100 m*M* HEPES pH 7.5, 75 m*M* ammonium sulfate, 32% pentaerythritol ethoxylate supplemented with 7.5% glycerol and subsequent flash-cooling in liquid ethane.

3. Results

The TRCF structure was solved using single-wavelength anomalous dispersion (SAD), as radiation damage affected the quality of data collected at subsequent wavelengths. An initial 4 Å resolution data set was collected at Advanced Photon Source beamline NE-CAT-8BM (Argonne National Laboratory) at the white line of the selenium *K* absorption edge (Table 1, SeMet1).

Data were processed using *HKL*2000/*SCALEPACK* (Otwinowski & Minor, 1997) and 25 out of 60 possible Se sites were found using *SnB* (Weeks & Miller, 1999). Difference Fourier syntheses were used to locate an additional 23 sites, giving a total of 48 sites. Initial phases calculated from data cut off at 5 Å resolution with *MLPHARE* (Otwinowski, 1991) produced an uninterpretable electron-density map. Density modification as implemented in *DM* also failed to improve the map. However, density modification with solvent flipping and phase extension to 4 Å with *SOLOMON* (Abrahams & Leslie, 1996) yielded a map with excellent connectivity. Subsequently, a 3.2 Å resolution data set was collected at Advanced Photon Source beamline SBC-19ID (Table 1, SeMet2) and these amplitudes were used for density modification and phase extension using *SOLOMON* (Abrahams & Leslie, 1996). The resulting 3.2 Å resolution electron-density map (Fig. 1*b*) was of high quality and contained significant side-chain information. Manual building and refinement of the molecular model are now in progress.

We thank Dr S. K. Burley for the kind gift of plasmid pSKB2, Drs Aziz Sancar and Chris Selby for plasmid pMFD19, Drs A. Ferré-d'Amaré and A. Roll-Mecak for helpful discussions and Ms H. Kulik for technical assistance. We thank Malcolm Capel, Craig Ogata and N. Sukumar for support at Advanced Photon Source beamline NE-CAT-8BM. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. We thank Stephan L. Ginnell and Younchang Kim for support at Advanced Photon Source beamline SBC-19ID. Use of the Advanced Photon Source is supported by the US Department of Energy, Office of Basic Energy Sciences under contract No. W-31-109-ENG-38. DNA sequencing was performed at The Rockefeller University DNA Sequencing Center. This work was supported by NIH grant GM073829 to SAD.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Gulick, A. M., Horswill, A. R., Thoden, J. B., Escalante-Semerena, J. C. & Rayment, I. (2002). *Acta Cryst.* D**58**, 306–309.
Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Park, J.-S., Marr, M. T. & Roberts, J. W. (2002). *Cell*, **109**, 757–767.
Selby, C. P. & Sancar, A. (1993). *Science*, **260**, 53–58.
Selby, C. P. & Sancar, A. (1995*a*). *J. Biol. Chem.* **270**, 4882–4889.
Selby, C. P. & Sancar, A. (1995*b*). *J. Biol. Chem.* **270**, 4890–4895.
Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.