Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary diffraction studies of TraF, a component of the *Escherichia coli* type IV secretory system

Gerald F. Audette,[a] Samantha J. Holland,[b] Trevor C. Elton,[b] Jan Manchak,[b] Koto Hayakawa,[a] Laura S. Frost[b] and Bart Hazes[a]*

[a]Department of Medical Microbiology and Immunology, University of Alberta, Edmonton, Alberta, Canada T6G 2H7, and [b]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E9

Correspondence e-mail: bart.hazes@ualberta.ca

Received 27 July 2004
Accepted 23 August 2004

TraF, a component of the *Escherichia coli* type IV secretory system, has been crystallized and preliminary X-ray diffraction data have been collected. TraF is a 26 kDa protein encoded by the *E. coli* F plasmid and is required for conjugative plasmid transfer and the formation of sex pili. The N-terminal domain of TraF has no recognizable sequence features, whereas the C-terminal domain is believed to adopt a thioredoxin fold. However, since the active-site cysteines of thioredoxin-like proteins are not conserved in TraF, its biochemical role remains unclear. TraF crystallizes in space group $C2$, with unit-cell parameters $a = 119.87$, $b = 34.36$, $c = 46.21$ Å, $\beta = 90.40^\circ$, and crystals diffract to 2.3 Å resolution.

1. Introduction

Conjugation is a plasmid-driven process that involves intimate association between a donor and recipient cell and the transfer of donor plasmid DNA through a supramolecular structure known as the mating-pair formation (Mpf) complex (Christie, 2001). Bacterial conjugation by F-like plasmids allows traits of medical importance such as antibiotic resistance to spread by horizontal gene transfer throughout a bacterial population. The F plasmid of *Escherichia coli* was the first discovered conjugative plasmid (Lederberg & Tatum, 1946) and remains a paradigm for the study of bacterial conjugation. The 100 kbp plasmid encodes all the genes necessary for conjugation in a 33.3 kbp DNA segment referred to as the transfer (*tra*) region. More recently, conjugative systems have been considered to be a subgroup of type IV secretion systems (T4SS) owing to the considerable similarity between the genes involved in the two processes (Salmond, 1994; Christie, 2001; Lawley *et al.*, 2003).

TraF is one of 12 proteins encoded by the F plasmid *tra* region that are essential for assembly of the conjugative F pilus required for conjugation. Null mutations in *traF* allow pilus-tip formation but block the outgrowth of the pilus beyond the cell surface (Wu *et al.*, 1988; Anthony *et al.*, 1999). *traF* is one of six genes that are hallmarks of F-like T4SS, as there are no orthologues in the P- or I-like T4SS subfamilies (Lawley *et al.*, 2003). TraF is also one of a number of plasmid-encoded periplasmic proteins that are related to the thioredoxin superfamily. Thioredoxins share a structural motif, the thioredoxin fold, consisting of a minimum of three α-helices flanking a four-stranded antiparallel β-sheet and a C*XX*C active site (Raina & Missiakas, 1997; Collet & Bardwell, 2002; Kinch *et al.*, 2003). In TraF the thioredoxin fold forms a C-terminal domain preceded by a domain of unknown function. Surprisingly, TraF does not contain the active-site cysteines in the C*XX*C motif characteristic of thioredoxin-like proteins. The fact that TraF is absolutely required for conjugation indicates that it retains an as yet unknown function independent of thioredoxin redox activity.

In our ongoing investigations into the molecular details of the T4SS process, we have identified several protein targets for structural characterization. The present report details the expression, purification, crystallization and initial X-ray diffraction analysis of TraF. These and future studies should provide insights into the role of TraF in pilus assembly during the conjugative process.

2. Methods and results

2.1. Cloning, expression and purification

The *traF* gene minus the leader sequence was PCR amplified from the pRS29 plasmid (Skurray *et al.*, 1978) using Vent DNA Polymerase (New England Biolabs) and a forward primer (GGATCCAAAGATGCAGGCTGGCAG) containing a 5′ *Bam*HI restriction endonuclease site and a reverse primer (CAGAATTCCTCAGAAAAGAAATAACCGG) containing a 3′ *Eco*RI site. The PCR-amplified *traF* DNA was isolated from a 1.0% agarose gel using standard procedures and ligated in-frame into the pGEX6P-2 expression vector (Amersham/GE Healthcare) for expression of TraF as a glutathione *S*-transferase (GST)-TraF fusion protein (plasmid pTEFX6). *E. coli* strain BL21-DE3 was trans-

 DOI: 10.1107/S0907444904020724

formed with the GST-TraF expression plasmid pTEFX6 and transformed cells were cultured at 310 K with shaking in Luria–Bertani medium containing ampicillin (50 µg ml^{-1}) to mid-log phase (OD_{600} = 0.5–0.7). The culture was then transferred to 303 K and protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside to a final concentration of 0.2 m*M* and allowed to proceed for 3 h. Induced cells were harvested by centrifugation at 4000*g* and stored frozen prior to protein purification. Cell pellets were lysed using the B-PER (Bacterial Protein Extraction Reagent) lysis reagent (Pierce) according to the manufacturer's protocol with addition of Complete-Mini Protease Inhibitor (Roche) and Benzonase Nuclease (Novagen). GST-TraF was separated from cleared lysates by affinity binding to glutathione agarose beads (Sigma) and the GST-TraF-bound beads were washed with 50 m*M* Tris–HCl pH 7.5, 150 m*M* sodium chloride, 1 m*M* EDTA, 1 m*M* DTT and 0.01% Nonidet P-40 (NP40). TraF was released from GST by incubation with PreScission Protease (Amersham/GE Healthcare) according to the manufacturer's protocol for on-column cleavage. The cleaved TraF was loaded onto a HiLoad 16/60 Superdex 75 (prep-grade) column (Amersham/GE Healthcare) and eluted with 50 m*M* Tris–HCl pH 7.5, 150 m*M* sodium chloride and 0.01% NP40 at a flow rate of 0.5 ml min^{-1}. Purified TraF was dialyzed against 50 m*M* Tris–HCl pH 7.5 and 100 m*M* sodium chloride and finally concentrated to 1.2–1.4 mg ml^{-1} as determined by a bicinchoninic acid (BCA) assay (Pierce). Concentrations greater than 1.5 mg ml^{-1} resulted in precipitation of the protein.

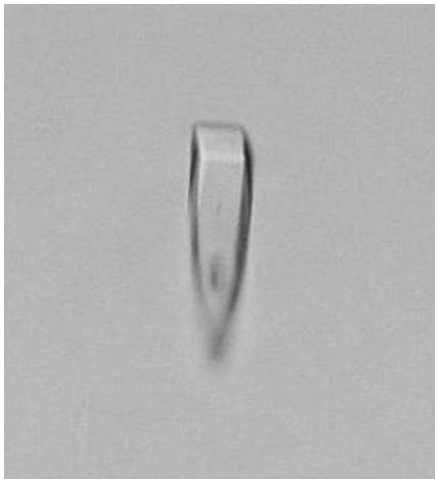

Figure 1
A crystal of TraF. Typical crystal dimensions are 0.05 × 0.03 ×0.02 mm.

The purity and identity of TraF were assessed by SDS–PAGE and MALDI–TOF mass-spectrometric (MS) analysis. Samples for MALDI–ToF MS analysis were prepared by adding 1 µl protein sample and 2 µl matrix solution (saturated sinapinic acid in 50:50 water:acetonitrile) directly to a MALDI target plate pretreated with 2 µl matrix solution and allowed to dry. A 3 µl drop containing a 2:1 mixture of matrix solution and 1 mg ml^{-1} horse heart myoglobin (Sigma) was employed for calibration. MALDI–TOF MS analysis was performed on a Voyager-DE PRO (Applied Biosystems). Mass calibration, baseline correction, noise filtering and de-isotoping procedures were carried out using *Data Explorer* v.4.0 (Applied Biosystems). TraF separates as a single band on SDS–PAGE; however MALDI–TOF MS analysis of purified TraF revealed two main peaks at 26 340 and 52 778 Da, respectively. The lower peak closely matches the theoretical molecular weight of 26 351 Da for cleaved TraF including five N-terminal residues from the expression construct (GPLGS). The second, higher molecular weight, MS peak exceeds the expected weight of a dimer by only 72 Da and suggests that TraF may dimerize in solution. The source of the 72 Da excess weight is currently unknown, but may result from tightly bound ion(s) in the dimer interface.

2.2. Crystallization

All crystallization experiments were performed using the hanging-drop vapour-diffusion method at ambient temperature. Initial crystallization conditions for TraF were identified from the PEG/Ion Screen (Hampton Research). Crystals suitable for X-ray diffraction analysis (typical dimensions of 0.05 × 0.03 × 0.02 mm; Fig. 1) were grown in 4–6 weeks from 1 µl drops containing equal volumes of protein (1.2 mg ml^{-1} in 50 m*M* Tris pH 7.5, 100 m*M* sodium chloride) and reservoir solution (25% PEG 3350, 100 m*M* HEPES pH 7.3, 200 m*M* diammonium citrate). Crystals were soaked in reservoir solution supplemented with 20% glycerol for 1–2 min prior to flash-cooling in 20 µm mounted cryoloops (0.05–0.1 mm diameter, Hampton Resaerch) by direct immersion into liquid nitrogen.

2.3. Data collection and reduction

X-ray diffraction data were collected on beamline 8.3.1 at the Advanced Light Source (Berkeley, CA, USA). Data were collected as 180 images on an ADSC Quantum 210 CCD detector with a crystal-to-detector distance of 125 mm. Intensity data were collected at 100 K using a wavelength of 1.12 Å with 1° oscillation per image. Diffraction data were processed using *MOSFLM* (Leslie, 1992) with an initial orientation matrix obtained from *LABELIT* (Sauter *et al.*, 2004) and the programs *SCALA* and *TRUNCATE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). TraF crystallizes in space group *C*2, with unit-cell parameters a = 119.87, b = 34.36, c = 46.21 Å, β = 90.40°. The crystal mosaicity refined to 1.5°, but because of the small unit cell and moderate resolution the reflections remained well separated. Radiation damage is only apparent in the second half of the data set, when data-quality statistics slowly deteriorated (not shown). This may explain the elevated $R_{\rm sym}$ value at high resolution (Table 1). Optimization of crystal quality, size and cryoconditions may further improve the diffraction potential of these crystals. Data-collection statistics are summarized in Table 1.

Table 1
Summary of data-collection statistics.

Values in parentheses are for the highest resolution shell.

Resolution (Å)	26.8–2.30 (2.36–2.30)
Total measured observations	27788
Unique reflections	8571 (583)
Completeness (%)	100 (99.9)
Mosaicity (°)	1.5
Multiplicity	3.2 (3.3)
$\langle I/\sigma(I)\rangle$	11.4 (3.0)
$R_{\rm sym}$†	0.130 (0.538)
$B_{\rm Wilson}$ (Å^2)	27.95

† $R_{\rm sym} = \sum_h \sum_i |I_{hi} - \langle I_h\rangle| / \sum_h \sum_i I_h$ for all reflections h each measured i times.

DsbC and other periplasmic disulfide-bond isomerases with a C-terminal thioredoxin fold function as homodimers (Raina & Missiakas, 1997; Collet & Bardwell, 2002). Mass-spectrometric analysis suggested that TraF may also form a dimer, although gel-filtration chromatography in the presence of the detergent NP40 was consistent with a TraF monomer. With a TraF dimer in the asymmetric unit, the $V_{\rm M}$ (Matthews, 1968) and solvent content would be 1.8 Å^3 Da^{-1} and 29.4%, respectively. However, analysis of the self-rotation function does not detect non-crystallographic rotational symmetry and no significant non-crystallographic translational symmetry peaks were observed in a Patterson map. Therefore, a single TraF monomer is expected in the asymmetric unit, with $V_{\rm M}$ = 3.6 Å^3 Da^{-1} and a solvent content of 64.4%. The high mosaicity of the crystal and moderate diffraction resolution is also consistent with an elevated solvent content. It is still possible that a TraF dimer is present

in the unit cell, with the dimer twofold axis coinciding with the crystallographic symmetry axis. However, this awaits full structure solution and refinement. The most closely related thioredoxin of known structure (PDB code 1thx; thioredoxin from *Anabaena* sp.; Saarinen *et al.*, 1995) has only 17% sequence identity and represents less than half the scattering mass of TraF. This is unlikely to be sufficient to phase the TraF data by molecular replacement, but will facilitate structure solution once intial phases have been obtained. Experimental phasing by heavy-atom and SeMet derivitization is currently under way.

This work was supported by operating grants from the Canadian Institutes for Health Research to LSF (MOP-55843) and the Alberta Heritage Foundation for Medical Research to BH. BH is an AHFMR scholar. X-ray diffraction data were collected at beamline 8.3.1 of the Advanced Light Source (ALS) at Lawrence Berkeley Laboratory under an agreement with the Alberta Synchrotron Institute (ASI). The ALS is operated by the Department of Energy and supported by the National Institute of Health. Beamline 8.3.1 was funded by the National Science Foundation, the University of California and Henry Wheeler. The ASI synchrotron access program is supported by grants from the Alberta Science and Research Authority (ASRA) and the AHFMR.

References

Anthony, K. G., Klimke, W. A., Manchak, J. & Frost, L. S. (1999). *J. Bacteriol.* **181**, 5149–5159.
Christie, P. J. (2001). *Mol. Microbiol.* **40**, 294–305.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Collet, J.-F. & Bardwell, J. C. A. (2002). *Mol. Microbiol.* **44**, 1–8.
Kinch, L. N., Baker, D. & Grishin, N. V. (2003). *Proteins Struct. Funct. Genet.* **52**, 323–331.
Lawley, T. D., Klimke, W. A., Gubbins, M. J. & Frost, L. S. (2003). *FEMS Microbiol. Lett.* **224**, 1–15.
Lederberg, J. & Tatum, E. L. (1946). *Nature (London)*, **158**, 558.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Raina, S. & Missiakas, D. (1997). *Annu. Rev. Microbiol.* **51**, 179–202.
Saarinen, M., Gleason, F. K. & Eklund, H. (1995). *Structure*, **3**, 1097–1108.
Salmond, G. P. C. (1994). *Annu. Rev. Phytopathol.* **32**, 181–200.
Sauter, N. K., Grosse-Kunstleve, R. W. & Adams, P. D. (2004). *J. Appl. Cryst.* **37**, 399–409.
Skurray, R. A., Nagaishe, H. & Clark, A. J. (1978). *Plasmid*, **1**, 174–186.
Wu, J. H., Kathir, P. & Ippen-Ihler, K. (1988). *J. Bacteriol.* **170**, 3633–3639.