Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray characterization of the relaxase domain of F factor TraI

Chris Larkin,† Saumen Datta,† Azin Nezami, Julie A. Dohm and Joel F. Schildbach*

Department of Biology, Johns Hopkins University, Baltimore, MD 21218, USA

† These authors contributed equally to the work.

Correspondence e-mail: joel@jhu.edu

Received 22 April 2003
Accepted 10 June 2003

Conjugative plasmids are capable of transferring a copy of themselves in single-stranded form from donor to recipient bacteria. Prior to transfer, one plasmid strand must be cleaved in a sequence-specific manner by a relaxase or mobilization protein. TraI is the relaxase for the conjugative plasmid F factor. A 36 kDa N-terminal fragment of TraI possesses the single-stranded DNA-binding and cleavage activity of the protein. Crystals of the 36 kDa TraI fragment in native and selenomethionine-labeled forms were grown by sitting-drop vapor-diffusion methods using PEG 1000 as the precipitant. Crystallization in the presence of chloride salts of magnesium and strontium was required to obtain crystals yielding high-resolution diffraction. To maintain high-resolution diffraction upon freezing, crystals had to be soaked in crystallization buffer with stepwise increases of ethylene glycol. The resulting crystals were trigonal and diffracted to a resolution of 3.1 Å or better using synchrotron radiation.

1. Introduction

Bacterial conjugation is the transfer of a copy of a conjugative plasmid between bacteria (reviewed in Frost *et al.*, 1994; Zechner *et al.*, 2000). The process is of great scientific and medical interest because of its important role in the transfer of antibiotic resistance genes between bacteria and in diversification of prokaryotic genomes. Although conjugative plasmids show considerable diversity, there are similarities in their transfer processes, including the DNA being transferred unidirectionally in single-stranded form. Proteins known as relaxases, nickases, mobilization proteins or transesterases are responsible for binding and cleaving single-stranded plasmid DNA (ssDNA) in a sequence-specific manner prior to transfer. These proteins belong to a broad family of proteins involved in DNA cleavage and ligation during plasmid transfer or rolling-circle replication of plasmids, bacteriophages or viruses (Ilyina & Koonin, 1992). This classification is based in part on the presence of a conserved 'H*U*H*UUU*' motif, where *U* is a hydrophobic residue. Family members also have at least one Tyr residue involved in DNA cleavage *via* nucleophilic attack of the Tyr hydroxyl oxygen on a DNA phosphate. The 'two-His' motif has been proposed to play a role in coordination of divalent cations (Ilyina & Koonin, 1992) or in abstracting the tyrosyl hydroxyl proton to assist the catalytic Tyr in its cleavage activity (Pansegrau *et al.*, 1994).

Relaxases from F factor and F-like plasmids are unique in that they also possess helicase activity, with the relaxase located to the N-terminal end of the protein while the helicase is located at C-terminal end (Abdel-Monem *et al.*, 1983; Grandoso *et al.*, 1994; Matson & Morton, 1991; Reygers *et al.*, 1991; Sherman & Matson, 1994; Traxler & Minkley, 1988). These two activities must be physically linked to permit efficient conjugative transfer (Llosa *et al.*, 1996; Matson *et al.*, 2001). The proteins are large, with F TraI consisting of 1756 amino acids and having a molecular weight of 192 kDa.

The size of F TraI complicates expression and purification and the presence of the helicase activity potentially interferes with measurements of relaxase DNA binding and function. We identified an N-terminal 330 amino-acid fragment of TraI, referred to as TraI36 or the F relaxase domain, that possesses *in vitro* relaxase function (Street *et al.*, 2003). We have shown that TraI36 has a subnanomolar K_d and a high level of sequence specificity for an ssDNA oligonucleotide having a sequence from the F factor origin of transfer (Stern & Schildbach, 2001). Single base changes over an 11-base region of this oligonucleotide can reduce affinity by up to three orders of magnitude.

To gain insight into the structural basis of the function of relaxases, we crystallized TraI36 in native and selenomethionine (SeMet) labelled forms. We have identified crystallization conditions and a cryoprotectant and

have collected complete data sets from crystals of both forms of the protein.

2. Materials and methods

2.1. Materials, overexpression and purification

Crystal Screen, Crystal Screen 2, Crystal Screen Lite, Additive Screen 1, 2 and 3, and CryoPro kits and Cryschem sitting-drop plates were purchased from Hampton Research. Corning 96-well sitting-drop plates were purchased from Fisher Scientific. Wizard I and II matrix screens were purchased from Emerald Biosciences. The cloning of the gene fragment encoding the F factor TraI relaxase domain and the expression and purification of the protein were as described in Street *et al.* (2003). Selenomethionine was purchased from Sigma and was incorporated into the protein as described previously (Doublie, 1997; Van Duyne *et al.*, 1993). Labeling was verified by MALDI–TOF mass spectrometry performed at the AB Mass Spectrometry Facility at the Johns Hopkins School of Medicine.

2.2. Crystallization and data collection

Crystals were obtained using hanging-drop and sitting-drop vapor-diffusion methods by combining protein with precipitant solution in a 1:1, 2:1 or 3:1 ratio. Corning 96-well plates or Cryschem 24-well plates were used. Crystallization was tested at 277, 290 and 296 K.

Crystal diffraction and cryoprotection conditions were screened using a Rigaku RU-H3RHB generator equipped with an MSC Max-Flux confocal optical system, an R-AXIS IV++ imaging-plate system and an X-Stream 2000 cryogenic system. Diffraction data sets were collected at CHESS F-2 beamline (Cornell High Energy Synchrotron Source, Ithaca, NY) using an Area Detector Systems CCD Quantum-210 detector (Szebenyi *et al.*, 1997) and an Oxford Cryosystems cryostat. Data were indexed, processed and scaled using the *HKL* processing package (Otwinowski & Minor, 1997).

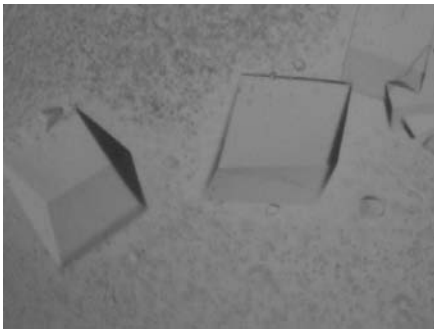

Figure 1
Crystals of the native TraI36 protein. The larger crystals shown are approximately 0.2 × 0.2 × 0.1 mm in size.

Table 1
Data-collection details and statistics.

Values in parentheses are for the highest resolution shell.

	Native	SeMet (remote)	SeMet (peak)	SeMet (inflection)
Space group	$P3_121$	$P3_121$	$P3_121$	$P3_121$
Unit-cell parameters				
$a = b$ (Å)	128.0	27.5	127.2	127.3
c (Å)	120.7	127.5	119.4	119.5
$\alpha = \beta$ (°)	90	90	90	90
γ (°)	120	120	120	120
No. of observed reflections	39965	20514	20417	20522
Average redundancy	11.2	10.9	10.9	10.9
Wavelength used (Å)	0.9713	0.9713	0.9789	0.9793
Resolution range (Å)	30.0–2.5 (2.59–2.50)	25.0–3.1 (3.2–3.1)	25.0–3.1 (3.2–3.1)	25.0–3.1 (3.2–3.1)
Completeness (%)	99 (100)	99.6 (100)	99.6 (100)	99.6 (100)
Mosaicity	0.354	0.630	0.648	0.632
$R_{\rm merge}$† (%)	5.7 (48.2)	9.0 (73.4)	10.8 (75.3)	9.5 (72.0)
$I/\sigma(I)$	43.3 (6.4)	18.0 (3.0)	17.6 (3.0)	18.0 (3.2)
$I > 3\sigma$ (%)	79	63	62	63

† $R_{\rm merge} = \sum |I - \langle I \rangle| / \sum I$.

3. Results and discussion

Hampton Research Crystal Screen 1, 2 and Lite kits and Emerald Biostructures Wizard kits I and II were used to screen for crystallization conditions for native protein (14 mg ml^{-1} in 25 m*M* Tris pH 8.25, 0.1 m*M* EDTA, 0.5 m*M* β-mercaptoethanol, 50 m*M* NaCl). For the native protein, the precipitants yielding best results were PEGs in the molecular-weight range 1000–4000, with the size of the crystals being increased in the presence of chloride salts of monovalent or divalent cations. Use of $MgCl_2$ yielded the best results. Crystal growth was best at pH 7.5 using either Tris or HEPES, with increasing precipitation with increasing pH. The optimal temperature was 296 K. Additive Screens 1, 2 and 3 were tested and addition of small concentrations of Sr, Ba and Cd to precipitant solutions increased the crystal size.

To improve the relatively small size of the crystals, macroseeding was attempted. Although the effort was complicated by considerable spontaneous crystal formation in the drops, some crystals increased in volume several-fold relative to the seed. However, the resulting crystals produced diffraction patterns with split reflections, indicating cracks or other imperfections in the crystals, and were unusable. Use of high concentrations of protein (20–35 mg ml^{-1}) with PEG 1000 (20–28%) and magnesium chloride, plus chloride salts of strontium, cadmium or barium generated crystals of approximately 0.2 × 0.2 × 0.2 mm (Fig. 1). The screen was repeated for SeMet-labelled protein. Crystallization conditions for the SeMet-labelled protein were similar to the native protein, although crystals tended to be smaller and crystallization was less reproducible.

During data collection at room temperature, crystals diffracted to approximately 3.5 Å using a rotating Cu-anode source, but were highly susceptible to radiation damage, with substantial loss of resolution after fewer than 60 min of exposure to X-rays. In an attempt to identify a cryoprotectant, the stability of the crystals in various cryoprotectants over time was examined. Effects were assessed visually. The crystals dissolved in the presence of a number of reagents, even when PEG concentrations were increased to offset the effect. Of those reagents that had no obvious effect, most, including lower molecular-weight PEGs, were insufficient cryoprotectants or reduced the resolution of the diffraction. Ethylene glycol proved to be the only effective cryoprotectant identified of the several screened. Passing the crystals through mother liquor plus 25 or 35% ethylene glycol immediately prior to freezing gave inconsistent results. Optimal results were achieved by soaking the crystals for 5 min in increasing percentages of ethylene glycol (5% increments) to 20% and then passing the crystals through 25% ethylene glycol and freezing in liquid nitrogen. Extending the soak periods beyond 5 min or soaking for any length of time in 30% ethylene glycol reduced the resolution of diffraction.

Once a cryoprotection procedure was identified, a narrow search of crystallization

conditions was conducted, with specific attention paid to the effects of cations and anions. The resulting crystals were screened for diffraction. The results indicated that $MgCl_2$ was required for optimal crystal size and for high-resolution diffraction following freezing. Including chloride salts of strontium or cadmium improved diffraction over the use of $MgCl_2$ alone. The best-quality diffraction was obtained with crystals grown from 100 m*M* Tris pH 7.5, 220 m*M* $MgCl_2$, 10 m*M* $SrCl_2$ and 20–28% PEG 1000. Attempts at co-crystallization with various oligonucleotides have so far been unsuccessful.

Using a rotating-anode source, a partial native data set to approximately 2.9 Å was collected from crystals of the native protein. In the hope of obtaining higher resolution data and of obtaining phases from MAD using data from the SeMet-labelled crystals, data were collected at the F-2 beamline at CHESS. The statistics for the data sets are listed in Table 1. Essentially complete data sets from the native (to 2.5 Å resolution) and the SeMet-labelled protein (to 3.1 Å) were collected. The high-resolution limit is considered to be where the highest resolution bin has $I/\sigma(I) \geq 3$. The $I/\sigma(I)$ cutoff used in scaling was set at -3 (the default value for *SCALEPACK*). We are currently attempting to obtain phasing information from the data sets collected from the SeMet-labeled protein. Based on the volume of the unit cell and the mass of the protein, we calculate a Matthews coefficient of 2.61 $Å^3\ Da^{-1}$, 52% solvent, three molecules in the asymmetric unit and 18 molecules in the unit cell. Each monomer has 13 methionines and we therefore expect 39 heavy atoms in the asymmetric unit. While the multiple molecules in the asymmetric unit suggests the possibility that the structure will reveal an oligomeric functional form for TraI36, we are skeptical given that the domain is monomeric under all conditions tested (Street *et al.*, 2003).

We are intrigued by the apparent requirement for $MgCl_2$ for achieving high-quality diffraction. Although $MgCl_2$ might influence crystallization in a number of ways, Mg^{2+} is required for DNA cleavage by TraI. Mg^{2+} is likely to be coordinated by multiple amino acids within the active site and these interactions might serve to stabilize a region of the protein or adopt a particular local conformation, both of which may assist crystallization.

We are grateful to Professors Ludwig Brand and Douglas Fambrough for use of equipment and to Apostolos Gittis for guidance and helpful discussions in the initial stages of this project. This work is based on research supported by National Institutes of Health grant GM61017 and National Science Foundation grant MCB9733655. The rotating-anode data were collected on a system purchased through support of National Science Foundation Multi-User Equipment Grant 9970207. We thank David Schuller, Bill Miller, Ernie Fontes and Tony Lloyd for assistance in data collection and analysis at CHESS. This work is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation under award DMR 97-13424, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by award RR-01646 from the National Institutes of Health through its National Center for Research Resources.

References

Abdel-Monem, M., Taucher-Scholz, G. & Klinkert, M. Q. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 4659–4663.

Doublie, S. (1997). *Methods Enzymol.* **276**, 523–530.

Frost, L. S., Ippen-Ihler, K. & Skurray, R. A. (1994). *Microbiol. Rev.* **58**, 162–210.

Grandoso, G., Llosa, M., Zabala, J. C. & de la Cruz, F. (1994). *Eur. J. Biochem.* **226**, 403–412.

Ilyina, T. V. & Koonin, E. V. (1992). *Nucleic Acids Res.* **20**, 3279–3285.

Llosa, M., Grandoso, G., Hernando, M. A. & de la Cruz, F. (1996). *J. Mol. Biol.* **264**, 56–67.

Matson, S. W. & Morton, B. S. (1991). *J. Biol. Chem.* **266**, 16232–16237.

Matson, S. W., Sampson, J. K. & Byrd, D. R. (2001). *J. Biol. Chem.* **276**, 2372–2379.

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

Pansegrau, W., Schroder, W. & Lanka, E. (1994). *J. Biol. Chem.* **269**, 2782–2789.

Reygers, U., Wessel, R., Muller, H. & Hoffmann-Berling, H. (1991). *EMBO J.* **10**, 2689–2694.

Sherman, J. A. & Matson, S. W. (1994). *J. Biol. Chem.* **269**, 26220–26226.

Stern, J. C. & Schildbach, J. F. (2001). *Biochemistry*, **40**, 11586–11595.

Street, L. M., Harley, M. J., Stern, J. C., Larkin, C., Williams, S. L., Miller, D. L., Dohm, J. A., Rodgers, M. E. & Schildbach, J. F. (2003). *Biochim. Biophys. Acta*, **1646**, 86–99.

Szebenyi, D. M. E., Arvai, A., Ealick, S., Laluppa, J. M. & Nielsen, C. (1997). *J. Synchrotron Rad.* **4**, 128–135.

Traxler, B. A. & Minkley, E. G. Jr (1988). *J. Mol. Biol.* **204**, 205–209.

Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.

Zechner, E. L., de la Cruz, F., Eisenbrandt, R., Grahn, A. M., Koraimann, G., Lanka, E., Muth, G., Pansegrau, W., Thomas, C. M., Wilkins, B. M. & Zatyka, M. (2000). *The Horizontal Gene Pool*, edited by C. M. Thomas, pp. 87–174. Amsterdam: Harwood Academic Publishers.