Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of restriction endonuclease *Sfi*I in complex with DNA

Hector Viadiu,[a] Éva Scheuring Vanamee,[a] Eric M. Jacobson,[a] Ira Schildkraut[b] and Aneel K. Aggarwal[a]*

[a]Structural Biology Program, Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York, NY 10029, USA, and [b]New England Biolabs Inc., 32 Tozer Road, Beverly, MA 01915, USA

Correspondence e-mail: aggarwal@inka.mssm.edu

Received 19 March 2003
Accepted 30 May 2003

The *Sfi*I endonuclease from *Streptomyces fimbriatus* (EC 3.1.21.4) is a tetrameric enzyme that binds simultaneously to two recognition sites and cleaves both sites concertedly. It serves as a good model system for studying both specificity and cooperative DNA binding. Crystals of the enzyme were obtained by the hanging-drop vapor-diffusion method in complex with a 21-mer oligonucleotide. The crystals are trigonal, with unit-cell parameters $a = b = 85.7$, $c = 202.6$ Å, and diffract to 2.6 Å resolution on a rotating-anode X-ray generator. Preliminary X-ray analysis reveals the space group to be either $P3_121$ or $P3_221$. Interestingly, the crystals change to space group $P6_122$, with unit-cell parameters $a = b = 85.5$, $c = 419.6$ Å, when the selenomethionyl (SeMet) derivative of the enzyme is co-crystallized with the same DNA. Phase information is currently being derived from this SeMet *Sfi*I–DNA complex.

1. Introduction

Classical type II restriction endonucleases are dimeric and recognize short 4–8 base-pair palindromic DNA sequences and cleave both DNA strands within these sequences. To cleave another site, the dimer has to first dissociate and rebind at another recognition site. However, a growing number of restriction enzymes have been shown to bind to two DNA sites simultaneously (Halford *et al.*, 1999; Bath *et al.*, 2002). These endonucleases fall into several different subclasses, including the type IIE enzymes such as *Eco*RII and *Nae*I that are dimeric and cleave only one site during a single turnover but require a second DNA site for allosteric activation (Reuter *et al.*, 1998; Huai *et al.*, 2000) and the type IIF enzymes such as *Sfi*I and *Cfr*10I that are tetrameric and cleave both DNA sites concertedly (Wentzell *et al.*, 1995; Deibert *et al.*, 2000). Another example is the type IIS endonuclease *Fok*I that is monomeric in solution but forms an active complex consisting of two protein and two DNA molecules (Vanamee *et al.*, 2001). Together, these novel restriction endonucleases show similarities to various DNA-modification enzymes that bring distant DNA sites together. For example, *Eco*RII shows sequence homology to the integrase family of recombinases (Topal & Conrad, 1993), while *Nae*I has been shown to posses topoisomerase and recombinase activities (Jo & Topal, 1995). *Sfi*I bears a close relationship to a family of recombinases that simultaneously catalyze a four-strand DNA breakage and a *Fok*I-like fold has recently been identified in TnsA, one of the two proteins of the Tn7 transposase that mediates the release of the transposon (Hickman *et al.*, 2000).

The *Sfi*I endonuclease is a 31 kDa protein that recognizes an eight base-pair palindromic sequence with a five-base pair interruption, 5′-GGCC*NNNN*$^{\downarrow}$*N*GGCC-3′, where *N* is any nucleotide and the arrow indicates the site of cleavage. *Sfi*I, like other members of the type IIF family, is a homotetramer in which each monomer recognizes a half-site. Both copies of the recognition sequence have to be bound to the enzyme, leading to concerted cleavage of both sites. Interestingly, binding of *Sfi*I to a non-cognate site results in an inactive complex that cannot bind to another DNA site (Williams & Halford, 2002). The structures of three type IIF enzymes have been reported: *Cfr*10I (Bozic *et al.*, 1996), *Ngo*MIV in complex with cleaved DNA (Deibert *et al.*, 2000) and *Bse*634I, an isoschisomer of *Cfr*10I (Grazulis *et al.*, 2002), none of which are related in sequence to *Sfi*I.

Here, we report the crystallization and preliminary X-ray data analysis of *Sfi*I in complex with a 21-mer oligonucleotide containing a single recognition site. Elucidating the structure of *Sfi*I in complex with its cognate DNA will provide a better understanding of cooperative binding and concerted cleavage of two DNA sites. *Sfi*I is the best biochemically

characterized of the type IIF enzymes; its sequence is unrelated to that of other type II enzymes and its recognition sequence is unusually long.

2. Materials and methods

2.1. Crystallization and preliminary X-ray diffraction analysis

Production and purification of native *Sfi*I was carried out as described previously (Wentzell *et al.*, 1995). Recombinant SeMet *Sfi*I was prepared by inhibiting the methionine-synthetic pathway and expressing the protein in the presence of selenomethionine. The purified native and SeMet proteins were stored in 0.2 *M* KCl, 10 m*M* K_3PO_4 pH 7.3, 0.1 m*M* EDTA and 1 m*M* dithiothreitol. The protein samples were concentrated to a final concentration of 22 mg ml^{-1} and were stored in small aliquots at 193 K.

Single-stranded oligonucleotides were synthesized for HPLC purification by leaving the trityl group attached to the 3′ end. The trityl group was then removed directly on the HPLC column. After purification, the oligonucleotides were desalted prior to annealing. Equimolar amounts of complementary oligonucleotides were mixed and annealed at room temperature overnight. The final concentration of the double-stranded oligonucleotides was ∼10 mg ml^{-1}. Co-crystallization was achieved by the hanging-drop vapor-diffusion method at 293 K. Commercially available screens from Hampton Research Inc. were used to determine the initial crystallization conditions. A 5:4 molar ratio of protein solution:DNA solution was mixed and kept on ice for 30–60 min prior to setting up crystallization trays. 1 µl of protein/DNA solution was mixed with 1 µl of well solution. After the initial conditions were identified, the pH, salt and precipitant concentrations were adjusted to obtain better diffracting crystals. Crystals grew overnight and reached maximum size within 2–3 d. The largest native crystals were 0.8 × 0.2 × 0.2 mm in size; the SeMet-derivative crystals were somewhat smaller, reaching a maximum size of 0.5 × 0.2 × 0.1 mm. All crystallographic data were measured at cryogenic temperatures. In-house data were collected using a Rigaku rotating-anode generator and an R-AXIS IV imaging-plate detector. The data were indexed and integrated with *DENZO* and reduced with *SCALEPACK* (Otwinowski & Minor, 1997).

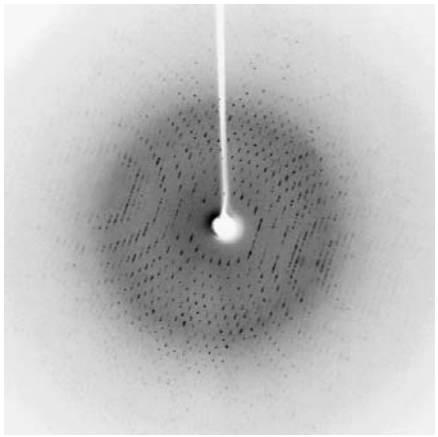

Figure 1
A 1° oscillation frame of a cryocooled crystal of native *Sfi*I in complex with an iodinated 21-mer oligonucleotide collected on a rotating-anode generator. Diffraction data are detectable to 2.6 Å resolution.

Table 1
Data-collection statistics for *Sfi*I–DNA complex.

Values in parentheses are for the highest resolution shell.

		SeMet derivative	
	Pseudo-native	Edge	Peak
Wavelength (Å)	1.54	0.97941	0.97931
Max. resolution (Å)	2.60	3.10	3.05
Total No. of reflections	955868	719313	709329
No. of unique reflections	27126 (2633)	17729 (1653)	18452 (1714)
R_{merge} (%)	9.6 (28.2)	11.0 (44.6)	11.2 (45.5)
Completeness (%)	99.6 (99.9)	99.6 (98.0)	99.7 (97.7)
$I/\sigma(I)$	34.0 (7.8)	19.9 (3.8)	18.0 (3.5)
No. of Se sites	N/A	8	8

3. Results and discussion

3.1. Crystallization

Several oligonucleotides ranging in size from 17 to 21 base pairs were tested for co-crystallization. The best crystals were obtained with a 21-mer oligonucleotide containing the sequence 5′-ATGT*GGCCAACAAGGCCT*ATT-3′ (top strand) and 5′-AAT*AGGCCTT*GTT*GGCCACAT-3′ (bottom strand). Initially, small crystals grew from several Hampton Crystal Screen conditions that contained MPD. After optimization, the best diffracting crystals grew from 0.1 *M* sodium acetate pH 4.6–5.0, 30–32.5% MPD and 5–10 m*M* $CaCl_2$ and diffracted to 3.0 Å on the home source (Fig. 1). Interestingly, replacement of the thymidine bases with iodouracil derivatives (indicated by stars in the sequence) lead to an increase in resolution from 3.0 to 2.6 Å. Crystals of the SeMet-derivative protein–DNA complex grew under identical conditions.

3.2. Preliminary X-ray diffraction analysis

Data of the pseudo-native crystals were collected on the home source. Preliminary X-ray diffraction analysis reveals the crystals to belong to the trigonal crystal system, with unit-cell parameters $a = b = 85.6$, $c = 202.6$ Å. The space group is either $P3_121$ or $P3_221$. Assuming one *Sfi*I dimer and one DNA molecule per asymmetric unit results in a V_M value of 2.4 Å^3 Da^{-1}. The other half of the *Sfi*I tetramer can be generated by the crystallographic twofold symmetry of the space group.

Crystals of the SeMet-derivative complex diffracted much more poorly than the native crystals: they diffracted to 3.0 Å at beamline 19-ID at the Advanced Photon Source. Moreover, the space group changes to $P6_122$ or $P6_522$, with unit-cell parameters $a = b = 85.7$, $c = 420.0$ Å. The basis for this change in space group and doubling of the c axis is unclear, although it is possible that one of the methionines in the native crystal is involved in lattice contacts and its replacement by SeMet destabilizes those contacts. We have found no evidence for merohedral twinning in the intensity data of the pseudo-native or the SeMet-derivative crystals. Like the native crystals, the asymmetric unit is likely to contain one *Sfi*I dimer and one DNA molecule, corresponding to a V_M of 2.5 Å^3 Da^{-1}. This was confirmed by the identification of a total of eight selenium sites (four per monomer) in *CNS* (Brünger *et al.*, 1998). Data-collection statistics for the pseudo-native and SeMet-derivative data are listed in Table 1. Because the native and SeMet-derivative crystals belong to different space groups, a medium-resolution structure of the *Sfi*I–DNA complex determined by SeMet MAD phasing will be used as a starting model for molecular replacement against the high-resolution native data. The structure will provide a framework for understanding the unique DNA-recognition properties of *Sfi*I.

We thank the staff at beamline 19-ID at the APS for help with data collection. This work was supported by NIH grants

GM44006 (AKA) and GM20015 (ÉSV). HV was supported by a Fulbright/CONACYT scholarship.

References

Bath, A. J., Milsom, S. E., Gormley, N. A. & Halford, S. E. (2002). *J. Biol. Chem.* **277**, 4024–4033.

Bozic, D., Grazulis, S., Siksnys, V. & Huber, R. (1996). *J. Mol. Biol.* **255**, 176–186.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Deibert, M., Grazulis, S., Sasnauskas, G., Siksnys, V. & Huber, R. (2000). *Nature Struct. Biol.* **7**, 792–799.

Grazulis, S., Deibert, M., Rimseliene, R., Skirgaila, R., Sasnauskas, G., Lagunavicius, A., Repin, V., Urbanke, C., Huber, R. & Siksnys, V. (2002). *Nucleic Acids Res.* **30**, 876–885.

Halford, S. E., Bilcock, D. T., Stanford, N. P., Williams, S. A., Milsom, S. E., Gormley, N. A., Watson, M. A., Bath, A. J., Embleton, M. L., Gowers, D. M., Daniels, L. E., Parry, S. H. & Szczelkun, M. D. (1999). *Biochem. Soc. Trans.* **27**, 696–699.

Hickman, A. B., Li, Y., Mathew, S. V., May, E. W., Craig, N. L. & Dyda, F. (2000). *Mol. Cell*, **5**, 1025–1034.

Huai, Q., Colandene, J. D., Chen, Y., Luo, F., Zhao, Y., Topal, M. D. & Ke, H. (2000). *EMBO J.* **19**, 3110–3118.

Jo, K. & Topal, M. D. (1995). *Science*, **267**, 1817–1820.

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

Reuter, M., Kupper, D., Meisel, A., Schroeder, C. & Kruger, D. H. (1998). *J. Biol. Chem.* **273**, 8294–8300.

Topal, M. D. & Conrad, M. (1993). *Nucleic Acids Res.* **21**, 2599–2603.

Vanamee, E. S., Santagata, S. & Aggarwal, A. K. (2001). *J. Mol. Biol.* **309**, 69–78.

Wentzell, L. M., Nobbs, T. J. & Halford, S. E. (1995). *J. Mol. Biol.* **248**, 581–595.

Williams, S. A. & Halford, S. E. (2002). *J. Mol. Biol.* **318**, 387–394.