

Crystallization and preliminary X-ray analysis of restriction endonuclease *FokI* bound to DNA

Joel A. Hirsch^{a,b}, David A. Wah^{a,b}, Lydia F. Dorner^{a,b}, Ira Schildkraut^{a,b},
Aneel K. Aggarwal^{a,b,*}

^aDepartment of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, 630 W. 168th St., New York, NY 10032, USA

^bNew England Biolabs, 32 Tozer Road, Beverly, MA 01915, USA

Received 16 December 1996

Abstract *FokI* is a type IIs restriction endonuclease which recognizes an asymmetric DNA sequence and cleaves DNA a short distance away from the sequence. The enzyme is bipartite in nature with its DNA recognition and cleavage functions located on distinct domains. We report here cocrystals of the complete *FokI* enzyme (579 amino acids) bound to a 20-bp DNA fragment containing its recognition sequence. The complex is amongst the largest protein-DNA complexes to be crystallized, and required macroseeding techniques for optimal crystal growth. The cocrystals diffract to at least 2.8 Å in resolution and belong to space group $P2_1$ with unit cell dimensions of $a = 67.9$ Å, $b = 119.8$ Å, $c = 69.1$ Å, $\beta = 96.6^\circ$. Using specific amino acid analysis we show that asymmetric unit contains a single *FokI* molecule bound to the 20-bp DNA fragment. This paper reports the first cocrystals of a type IIs restriction endonuclease.

© 1997 Federation of European Biochemical Societies.

Key words: *FokI*; Crystallization; Restriction endonuclease, type IIs; X-ray diffraction analysis

1. Introduction

FokI from *Flavobacterium okeanoikoites* is a member of type IIs subclass of restriction enzymes. Unlike the more common type II enzymes such as *BamHI* and *EcoRI* which recognize palindromic DNA sequences and cleave within those sequences, *FokI* recognizes an asymmetric DNA sequence 5'-GGATG-3' and cleaves 9 and 13 base pairs away from the recognition sequence [1] (Fig. 1). Furthermore, *FokI* exists as a monomer in solution while *BamHI* and *EcoRI* exist as dimers [2]. Biochemical studies show that *FokI* has a bipartite structure, containing two functionally independent domains, one which serves as the DNA binding domain (41 kDa), and one which serves as the non-specific cleavage domain (25 kDa) [3]. Due to its bipartite nature, *FokI* has been used to create artificial enzymes by attaching its cleavage domain to the DNA binding domains of transcription factors [4,5]. *FokI* can also be used to cleave single stranded DNA at predetermined sites by using suitably tailored oligonucleotide adapters [6]. To understand how the bipartite enzyme recognizes and cleaves DNA, we have cocrystallized the intact enzyme with a 20-bp DNA fragment containing its recognition sequence.

2. Materials and methods

The *FokI* endonuclease gene [7,8] was expressed under the control of P_{tac} promoter in *E. coli* with a copy of the *FokI* methylase gene residing on the chromosome. The cells were grown at 37°C to mid-log phase and then induced with 0.3 mM isopropyl β -thiogalactoside (IPTG). 3 h after induction the cells were harvested by centrifugation at 4°C, suspended in 4 vols. of buffer A (10 mM potassium phosphate (pH 6.5), 1 mM EDTA, 1 mM EGTA, and 1 mM DTT, 0.2 M NaCl). The cells were disrupted by sonication, and the cellular debris was removed by centrifugation. The resulting supernatant was applied to a phosphocellulose (P-11, Whatman) column and eluted with a 0.2–1.0 M NaCl gradient. The fractions which contained active *FokI* endonuclease were identified by incubating a sample with DNA followed by analysis with agarose gel electrophoresis. The active fractions were pooled and dialyzed against buffer B (10 mM potassium phosphate (pH 7.3), 0.1 M NaCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, and 10% glycerol). The dialysate was then applied to a CM Sepharose column and eluted with a 0.1–0.8 M NaCl gradient. In this and the following chromatographic steps, the *FokI* endonuclease was the major protein species and was followed as a ~66 kDa band by SDS PAGE. The pool of *FokI* endonuclease protein was diluted 7-fold with buffer C (10 mM potassium phosphate (pH 6.8), 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 10% glycerol), and then applied to a heparin Sepharose column and eluted with a 0.0–0.8 M KCl gradient developed in buffer C. The eluted endonuclease was diluted with 3 vols. of buffer C and then extensively dialyzed against buffer D (10 mM Tris-HCl (pH 7.8), 50 mM KCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 10% glycerol). The dialysate was then applied to a DEAE Sepharose column and eluted with a 50 mM to 0.8 M KCl gradient. The resulting *FokI* endonuclease had a specific activity of ~30 000 units per mg of protein (one unit is defined as the amount of enzyme required to completely digest 1 μ g of lambda DNA in a 50 μ l reaction mixture of 50 mM potassium acetate, 20 mM Tris-acetate (pH 7.9), 10 mM magnesium acetate, 1 mM DTT). A single liter of bacterial cells yields about 3 mg of purified protein (Fig. 2). For crystallization the enzyme is concentrated to 10 mg/ml by ultrafiltration, aliquoted and stored at –70°C.

DNA was synthesized on an Applied Biosystems 391 PCRmate automated synthesizer using phosphoramidite chemistry. We retained the 5' trityl groups to facilitate later HPLC purification. Derivatized oligonucleotides were synthesized by substituting iodouracil phosphoramidites (Peninsula Laboratories) for the thymidine phosphoramidites. Oligonucleotides were cleaved and deprotected with NH_4OH as per manufacturer's protocols. The DNA was purified by HPLC with a Rainin PureDNA column which facilitates removal of failed sequences and then detritylation with trifluoroacetic acid [9]. In some instances, oligonucleotides were purified with denaturing PAGE, extracted, desalted and further purified by HPLC. After HPLC, samples were desalted, resuspended in Bis-Tris buffer (25 mM pH 6.8) with 0.1 M KCl and annealed together by heating to 85°C for 5 min, followed by slow cooling.

3. Results and discussion

For cocrystallization trials we purified two 20-bp DNA oligomers, one with blunt ends and the other with a single AT base pair overhang (Table 1). Both oligomers encom-

*Corresponding author: Fax: (1) (212) 305-1019.
E-mail: aggarwal@convex.hhmi.columbia.edu

Abbreviations: DTT, dithiothreitol; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; HPLC, high-pressure liquid chromatography; MES, 2-(*N*-morpholino)ethanesulfonic acid

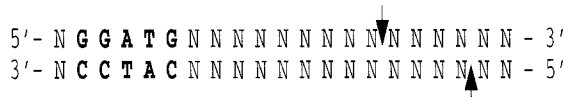


Fig. 1. DNA recognition and cleavage by *FokI*. The 5 base pair asymmetric recognition sequence is shown in bold. Arrows indicate the sites of cleavage.

passed the *FokI* recognition sequence GGATG at one end and the potential cleavage sites 9 and 13 base pairs away at the other end. The best cocrystals were obtained with the overhanging 20-mer, using macroseeding techniques [10]. The procedure for growing these cocrystals is as follows: Freshly thawed *FokI* is mixed with the DNA at a molar ratio of 1:1 and allowed to equilibrate on ice for 20 min. The mixture is set up as 2 μ l hanging drops containing 0.15 mM protein, 0.75 mM DNA, 1.1 M ammonium sulfate, 0.5 M MES (pH 6.0), 0.2 M KCl, 10 mM potassium phosphate, 0.5 mM DTT, 0.5 mM EDTA, 0.5 mM EGTA, and 5% glycerol against reservoir solutions containing 2.2 M ammonium sulfate and 1.0 M MES (pH 6.0). Small cocrystals, measuring ~ 0.03 mm $\times$ 0.2 mm $\times$ 0.15 mm, are observed within 5 days at 20°C. The small cocrystals are macroseeded by first washing them in a solution containing 1.4 M ammonium sulfate and 0.1 M MES (pH 6.0), and then transferring a single crystal in 1 μ l of wash solution to 1 μ l of fresh protein/DNA mixture. The cocrystals grow optimally with reservoir solutions containing 1.75 M ammonium sulfate and 1.0 M MES (pH 6.0). The size, quality, and reproducibility of the cocrystals are much improved with sitting drop geometry as opposed to hanging drops [11]. After 2–3 days at 20°C, the seeded cocrystals grow to 1.0 mm $\times$ 0.2 mm $\times$ 0.15 mm in dimensions and resemble parallelepipeds in shape (Fig. 3). The cocrystals are highly birefringent and stable in their sitting drops for many months. The presence of protein and DNA in the cocrystals was confirmed by PAGE after thoroughly washing crystals in their reservoir solution.

Cocrystals were analyzed by precession photography on a Rigaku rotating anode X-ray source operating at 100 mA and 50 kV. Cocrystals obtained with the blunt-ended oligomer diffracted weakly to a resolution of ~ 4 Å and were not further characterized. Cocrystals obtained with the overhanging oligomer exhibited much stronger diffraction. These cocrystals belong to space group $P2_1$ with unit cell dimensions of $a = 67.9$ Å, $b = 119.8$ Å, $c = 69.1$ Å, and $\beta = 96.6^\circ$. Very strong reflections at 3.4 Å along the a axis, characteristic of DNA, suggest that the DNA axis is parallel to the a axis. In accordance, the length of the axis is approximately the length of a 20-bp B-form DNA oligomer.

Assuming a model in which a single *FokI* molecule is bound to the 20-bp DNA duplex, V_m [12] calculations left an ambiguity as to the number of protein/DNA complexes contained in the crystallographic asymmetric unit. One protein/DNA complex per asymmetric unit (asu) gives a V_m of 3.5 Å³/Da,

Table 1
DNA oligomers used in cocrystallization studies with *FokI*

Blunt-ended ^a	5'- <u>CGGATG</u> GATAACGCTAGTCAC -3'
	3'- GCCTACTATTGCGATCAGTG -5'
TA overhang ^a	5'- TCGGATGATAACGCTAGTCA -3'
	3'- GCCTACTATTGCGATCAGTA -5'

^aUnderlines indicate the *FokI* recognition sequence.

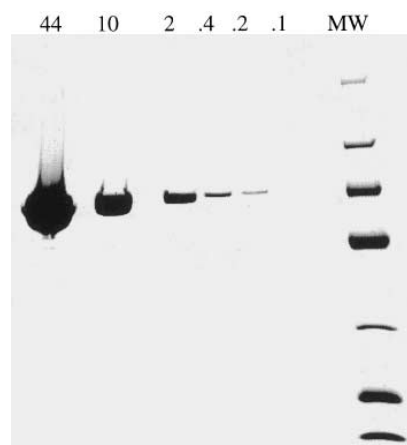


Fig. 2. An SDS-polyacrylamide gel electrophoretogram of final *FokI* endonuclease preparation stained with Coomassie blue. The numbers at the top of each lane indicate the amount of protein that was loaded in μ g. MW is the lane that contains molecular weight standards (top to bottom, in kDa: 200, 97, 68, 43, 29, 18 and 14).

while two complexes give a V_m of 1.75 Å³/Da. Both of these values are within the range found in macromolecular crystals [13]. Therefore, to determine Z_a (the number of molecules per asu) we used volume-specific amino acid analysis [14]. Two large cocrystals were carefully measured to determine their volume and then dissolved and subjected to amino acid analysis. The results from this analysis gave Z_a values of 0.91 and 1.07, strongly suggesting that the asymmetric unit contains a single protein/DNA complex.

For data collection, we developed a procedure to flash freeze the cocrystals at 110 K by first soaking them for 8 h in a stabilization solution (2.3 M ammonium sulfate, 1.0 M MES (pH 6.0)) containing 20% glycerol. The fragility of the crystals required transient addition of 0.125% glutaraldehyde [15]. The unit cell dimensions were found to change during freezing to $a = 65.59$, $b = 119.34$, $c = 71.52$ Å, and $\beta = 101.4^\circ$. We have measured native data at the Cornell High Energy Synchrotron Source (Beamline A1) from a single flash frozen crystal, utilizing the charge coupled device (CCD) detector. The data extends to 2.8 Å in resolution and merges with an overall R_{sym} of 7.5%. We have also measured data from heavy atom derivatives including an iodinated derivative in which contains iodouracil residues in place of thymine residues along the DNA duplex. The structure of *FokI*/DNA complex prom-

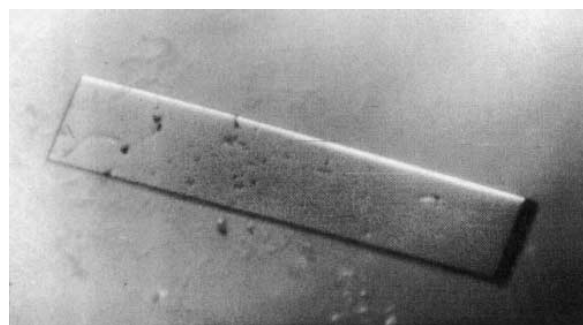


Fig. 3. A macroseeded cocrystal of *FokI* endonuclease with the 20-bp overhang DNA fragment (cf. Table 1) measuring 1.0 mm $\times$ 0.2 mm $\times$ 0.15 mm.

ises to provide a framework for the design of chimeric enzymes with novel specificities.

Acknowledgements: We thank W. Jack for providing the overexpressing construct of *FokI*, R. Knott for assistance with DNA synthesis, and the Protein Core laboratory (Columbia University) for help with amino acid analysis. We thank the staff at CHESS for help with synchrotron radiation, and members of the Aggarwal laboratory for help with data collection. This work was supported an NIH GM-44006 grant (A.K.A), an NIH Biophysics training grant (J.A.H.), and an NIH Eye Research training grant (D.A.W.).

References

- [1] Roberts, R.J. and Halford S.E. (1993) in: *Nucleases* (Linn, S.M., Lloyd, R.S. and Roberts, R.J. eds.) pp. 35–88, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [2] Skowron, P., Kaczorowski, T., Tucholski, J. and Podhajska, A.J. (1993) *Gene* 125, 1–10.
- [3] Lin, L., Wu, L.P. and Chandrasegaran S. (1992) *Proc. Natl. Acad. Sci. USA* 89, 4275–4279.
- [4] Kim, Y.G. and Chandrasegaran S. (1994) *Proc. Natl. Acad. Sci. USA* 91, 883–887.
- [5] Kim, Y.G., Cha, J. and Chandrasegaran, S. (1996) *Proc. Natl. Acad. Sci. USA* 93, 1156–1160.
- [6] Kim, S.C., Podhajska, A.J. and Szybalski, W. (1988) *Science* 240, 504–506.
- [7] Kita, K., Kotani, H., Sugisaki, H. and Takanami, M. (1989) *J. Biol. Chem.* 264, 5751–5756.
- [8] Looney, M.C., Moran, L.S., Jack, W.E., Feehery, G.R., Benner, J.S., Slatko, B.E. and Wilson G.G. (1989) *Gene* 80, 193–208.
- [9] Aggarwal, A.K. (1990) *Methods* 1, 83–90.
- [10] Stura, E.A. and Wilson, I.A. (1990) *Methods* 1, 38–49.
- [11] McPherson, A. (1990). *Eur. J. Biochem.* 189, 1–23.
- [12] Matthews, B.W. (1968) *J. Mol. Biol.* 33, 491–497.
- [13] Blundell, T.L. and Johnson, L.N. (1976) *Protein Crystallography*, Academic Press, London.
- [14] Kwong, P.D., Pound, A., Hendrickson, W.A. (1994) *J. Appl. Crystallogr.* 27, 504–509.
- [15] Petsko, G.A. (1985) *Methods Enzymol.* 114, 147–156.