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

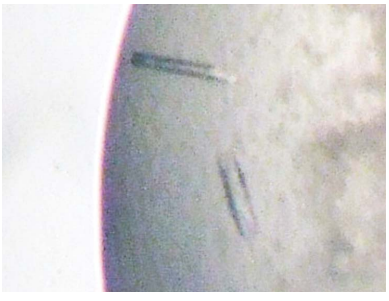
Crystallization and X-ray diffraction studies of DNA-free and DNA-bound forms of *Eco*O109I DNA methyltransferase

Makoto Iwamoto,[a] Asami Hishiki,[a] Takashi Shimada,[a] Tsuyoshi Imasaki,[a] Junko Tsuda,[b] Keiko Kita,[b,c] Toshiyuki Shimizu,[d] Mamoru Sato[a] and Hiroshi Hashimoto[a]*

[a]Graduate School of Nanobioscience, Yokohama City University, Japan, [b]Department of Biotechnology, Tottori University, Japan, [c]Graduate School of Agriculture, Kyoto University, Japan, and [d]Graduate School of Pharmaceutical Science, The University of Tokyo, Japan

Correspondence e-mail: hash@tsurumi.yokohama-cu.ac.jp

Received 26 August 2010
Accepted 20 September 2010



*Eco*O109I DNA methyltransferase (M.*Eco*O109I) is a type II modification enzyme from the *Eco*O109I restriction-modification system identified in *Escherichia coli* strain H709c. M.*Eco*O109I recognizes double-stranded *R*GG*N*CC*Y* (where *R* = A or G, *Y* = T or C and *N* is any base) and transfers a methyl group to the C5 of the inner cytosines from *S*-adenosylmethionine. To reveal the mechanism of substrate recognition by M.*Eco*O109I, DNA-free and DNA-bound forms of M.*Eco*O109I were successfully crystallized. Crystals of the DNA-free and DNA-bound forms belonged to space groups $P4_22_12$, with unit-cell parameters $a = b = 120.5$, $c = 79.8$ Å, and $P2_1$, with unit-cell parameters $a = 55.8$, $b = 77.4$, $c = 117.4$ Å, $\beta = 93.5^\circ$, respectively.

1. Introduction

The restriction-modification (RM) system is a defensive strategy used by bacteria to protect themselves from foreign DNAs from viruses such as bacteriophages. The RM system is basically composed of two enzymatic activities: restriction endonuclease (REase) and modificaton DNA methyltransferase (MTase). REases recognize a specific base sequence of four to eight bases in double-stranded DNA and cleave both strands of the duplex. MTases methylate a specific base (at the amino group of an adenine or either the 5 position or the amino group of a cytosine) in the same base sequence as recognized by the cognate REase. REases do not cleave the corresponding methylated DNA. Therefore, the RM system protects the bacteria against viral invasion by cutting foreign DNAs. RM systems are divided into four classes: types I, II, III and IV (Roberts *et al.*, 2003). Of those classes, the type II RM system has been the most well studied because it is a useful tool in gene-manipulation technology. To date, over 250 types of recognition sequences have been identified in the REBASE database (http://rebase.neb.com). Therefore, the structures of RM enzymes in complex with DNA that includes the recognition sequence will provide a comprehensive structural basis for sequence-specific interactions between protein and DNA.

*Eco*O109I DNA methyltransferase (M.*Eco*O109I) is a type II MTase from the *Eco*O109I RM system identified in *Escherichia coli* strain H709c (Kita *et al.*, 1999). M.*Eco*O109I is composed of 414 amino acids with a molecular mass of 46 kDa and exists as a monomer in solution. M.*Eco*O109I recognizes double-stranded *R*GG*N*CC*Y* (where *R* = A or G, *Y* = T or C and *N* is any base) and transfers a methyl group to the C5 of the inner cytosines of both strands from *S*-adenosylmethionine (SAM; Kita *et al.*, 2001). To date, several structures of MTases in complex with DNA that includes their recognition sequence have been determined (Klimasauskas *et al.*, 1994; Reinisch *et al.*, 1995; Goedecke *et al.*, 2001; Horton *et al.*, 2006). These MTases recognize a 4 bp palindromic sequence, but no three-dimensional structures of MTases that recognize a discontinuous and degenerate 7 bp sequence are available. We have previously determined the crystal structures of DNA-free and DNA-bound forms of an REase from the *Eco*O109I RM system (R.*Eco*O109I) and revealed the mechanisms of the base-recognition and hydrolysis reaction (Hashimoto *et al.*, 2005). R.*Eco*O109I forms a dimer and recognizes the pseudo-palindromic sequence *R*GG*N*CC*Y*. In contrast, M.*Eco*O109I is a monomeric protein and recognizes this sequence. Thus, comparison of the base-recognition mechanisms

between R.*Eco*O109I and M.*Eco*O109I is of great interest. To date, DNA-bound structures of both the REase and MTase from an RM system have not been determined. Here, we report the purification, crystallization and initial X-ray diffraction study of the DNA-free and DNA-bound forms of M.*Eco*O109I.

2. Materials and results

2.1. Expression and purification of recombinant proteins

The amplified cDNA of M.*Eco*O109I was inserted into the *Bam*HI–*Xho*I site of the pGEX6P-1 vector (GE Healthcare). Wild-type M.*Eco*O109I was overexpressed as a GST-fusion protein in *E. coli* strain HB101 in LB medium containing 50 µg ml^{-1} ampicillin. Harvested cells were suspended in buffer *A* (50 m*M* HEPES–NaOH pH 7.4, 200 m*M* NaCl and 10 m*M* β-mercaptoethanol) and lysed by sonication. The cell lysate was clarified by centrifugation. The supernatant was applied onto glutathione Sepharose 4B resin (GE Healthcare) equilibrated with buffer *A*. The resin was washed with buffer *B* (50 m*M* HEPES–NaOH pH 7.4, 1.0 *M* NaCl and 10 m*M* β-mercaptoethanol). The GST-fusion protein was eluted with buffer *C* (50 m*M* Tris–HCl pH 9.0, 200 m*M* NaCl, 10 m*M* β-mercaptoethanol and 50 m*M* reduced glutathione). PreScission protease (GE Healthcare) was added to the eluted solution to cleave the GST tag. Complete digestion of the GST tag was confirmed by SDS–PAGE with Coomassie Brilliant Blue stain and the reaction mixture was applied onto a HiTrap Heparin (GE Healthcare) column equilibrated with buffer *A*. The bound protein was eluted with a linear gradient to buffer *B*. The eluted solution was applied onto a HiLoad Superdex 200 (GE Healthcare) column equilibrated with buffer *D* [10 m*M* HEPES–NaOH pH 7.4, 100 m*M* NaCl, 10 m*M* β-mercaptoethanol and 5%(*v*/*v*) glycerol]. The fractions containing M.*Eco*O109I were concentrated to 20 mg ml^{-1} using a Centricon YM10 (Millipore). The homogeneity of the purified protein was evaluated by SDS–PAGE with Coomassie Brilliant Blue stain. The concentrated protein solution was frozen in liquid nitrogen and stored at 193 K.

Selenomethionyl protein with the C84S substitution (SeMet-M.*Eco*109I^{C84S}) for structure determination using the MAD or SAD method was overexpressed in the methionine-auxotrophic *E. coli* strain B834(DE3)pLysS using a minimal medium containing 50 µg ml^{-1} ampicillin, 34 µg ml^{-1} chloramphenicol and 25 mg ml^{-1} seleno-L-methionine. The C84S substitution, which produces catalytically inactive M.*Eco*O109I, was introduced using the QuikChange protocol (Stratagene) because this strain possesses the *Mcr*BC restriction system that digests cytosine-methylated DNA. Therefore, to overexpress M.*Eco*O109I in this strain, inactivation of M.*Eco*O109I was required. SeMet-M.*Eco*O109I^{C84S} was purified using a procedure similar to that used for the wild-type protein.

2.2. Crystallization and X-ray diffraction studies

Crystals suitable for X-ray diffraction experiments were obtained using the hanging-drop vapour-diffusion method. The typical hanging drop was prepared by mixing 2.0 µl each of protein and reservoir solution, and vapour diffusion was then started against 200 µl reservoir solution. Crystals of the DNA-free form of M.*Eco*O109I were obtained in a few weeks in a hanging drop prepared by mixing protein solution [200 µ*M* M.*Eco*O109I and 1.0 m*M* *S*-adenosylhomocysteine (SAH) in buffer *D*] with reservoir solution (100 m*M* sodium citrate pH 5.6, 300 m*M* lithium sulfate and 800 m*M* ammonium sulfate) at 293 K (Fig. 1). For crystallization of the DNA-bound form of M.*Eco*O109I, the commercially synthesized single strands 5′-TGG-CAGGGCCCGGT-3′ and 5′-AACCGGGCCCTGCC-3′ (Invitrogen) were annealed to yield a double-stranded DNA (dsDNA) with 5′-overhangs. The DNA complex was formed in a solution consisting of 200 µ*M* M.*Eco*O109I, 1.0 m*M* SAH and 200 µ*M* dsDNA on ice for 30 min. Formation of the DNA complex was validated using an electrophoretic mobility-shift assay (EMSA; Fig. 2*a*). The mixture was separated by electrophoresis at 277 K on a 1% agarose gel at 100 V for 25 min in a conventional TAE buffer. The gel was stained with ethidium bromide and bands were detected using a UV transilluminator. Crystals of the DNA-bound form were obtained in a hanging drop prepared by mixing the above protein–DNA solution

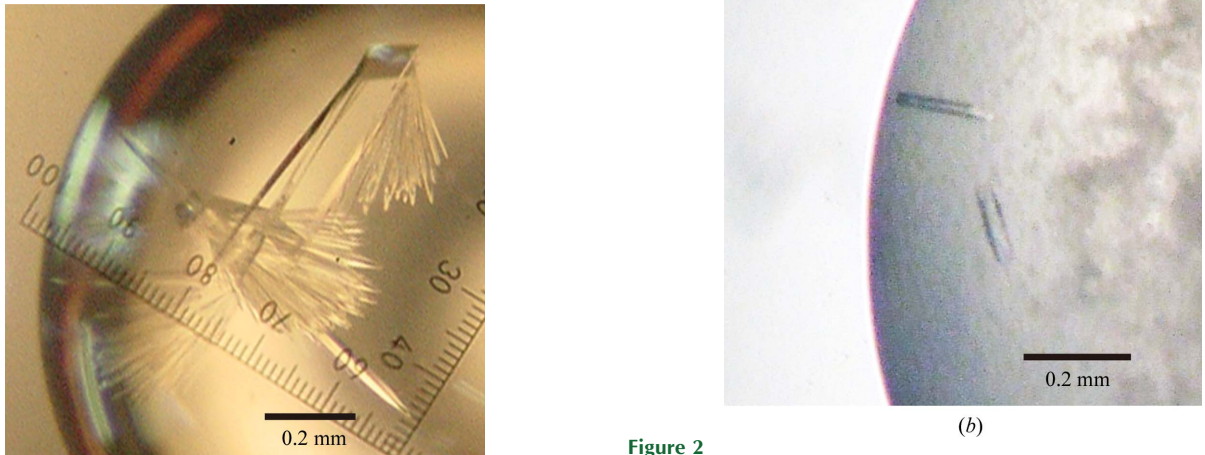


Figure 1
Crystals of the DNA-free form of M.*Eco*O109I.

Figure 2
(*a*) Electrophoretic mobility-shift assay using dsDNA. Lanes 1 and 2 contain dsDNA and M.*Eco*O109I only, respectively. (*b*) Crystals of the DNA-bound form of M.*Eco*O109I.

Table 1
Crystallographic data of M.*Eco*109I.

Values in parentheses are for the highest resolution shell.

	DNA-free form	DNA-bound form
Space group	$P4_22_12$	$P2_1$
Unit-cell parameters		
a (Å)	120.5	55.8
b (Å)	120.5	77.4
c (Å)	79.8	117.4
β (°)		93.5
V_M (Å^3 Da^{-1})	3.16	2.33
Solvent content (%)	61.1	47.2
Wavelength (Å)	1.0000	1.0000
Resolution range (Å)	50.00–2.00 (2.07–2.00)	50.00–2.00 (2.07–2.00)
Measured reflections	515682 (43329)	236429 (21444)
Unique reflections	40345 (3939)	66911 (6498)
Completeness (%)	99.9 (100)	98.7 (96.4)
Mean $I/\sigma(I)$	16.6 (9.8)	12.2 (4.5)
R_{merge}† (%)	9.8 (35.9)	7.0 (30.5)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

with reservoir solution [50 m*M* HEPES–NaOH buffer pH 7.0, 80 m*M* magnesium acetate and 30%(*w*/*v*) PEG 8K] at 293 K after a few weeks (Fig. 2*b*). All crystals for X-ray diffraction experiments were cryoprotected in reservoir solution containing 20–30% ethylene glycol and were then cooled in a nitrogen-gas stream at 100 K.

X-ray diffraction data were collected using an ADSC Q210 (Area Detector Systems Corp.) detector on NW12A at KEK PF-AR and an ADSC Q315 (Area Detector Systems Corp.) detector on BL41XU at SPring-8. All X-ray diffraction data were indexed, integrated, scaled and averaged using *HKL*-2000 (Otwinowski & Minor, 1997). Crystals of the DNA-free and DNA-bound forms belonged to space groups $P4_22_12$, with unit-cell parameters $a = b = 120.5$, $c = 79.8$ Å, and $P2_1$, with unit-cell parameters $a = 55.8$, $b = 77.4$, $c = 117.4$ Å, $\beta = 93.5^\circ$, respectively. In the case of the DNA-free form, the crystallographic asymmetric unit was estimated to contain one molecule of M.*Eco*O109I, with a corresponding crystal volume per protein weight (V_M; Matthews, 1968) of 3.16 Å^3 Da^{-1} and a solvent content of 61.1%. In contrast, the asymmetric unit of the DNA-bound form was estimated to contain two molecules of the M.*Eco*O109I–dsDNA complex with a corresponding V_M value (Matthews, 1968) of 2.33 Å^3 Da^{-1} and a solvent content of 47.2%. Crystallographic data are summarized in Table 1. Crystals of the DNA-free and DNA-bound forms of SeMet-M.*Eco*O109I^{C84S} were also obtained under similar conditions to the wild-type crystals. Structure determinations of the DNA-free and the DNA-bound forms using anomalous dispersion are now in progress.

We acknowledge the kind support of the beamline staff at the Photon Factory and SPring-8 for data collection. This work was supported by KAKENHI, National Project on Protein Structural and Functional Analyses (Protein 3000 Project) and Target Protein Research Program grants to MS, TS and HH from MEXT.

References

Goedecke, K., Pignot, M., Goody, R. S., Scheidig, A. J. & Weinhold, E. (2001). *Nature Struct. Biol.* **8**, 121–125.
Hashimoto, H., Shimizu, T., Imasaki, T., Kato, M., Shichijo, N., Kita, K. & Sato, M. (2005). *J. Biol. Chem.* **280**, 5605–5610.
Horton, J. R., Liebert, K., Bekes, M., Jeltsch, A. & Cheng, X. (2006). *J. Mol. Biol.* **358**, 559–570.
Kita, K., Tsuda, J., Kato, T., Okamoto, K., Yanase, H. & Tanaka, M. (1999). *J. Bacteriol.* **181**, 6822–6827.
Kita, K., Tsuda, J. & Nishigaki, R. (2001). *Biosci. Biotechnol. Biochem.* **65**, 2512–2518.
Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. (1994). *Cell*, **76**, 357–369.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Reinisch, K. M., Chen, L., Verdine, G. L. & Lipscomb, W. N. (1995). *Cell*, **82**, 143–153.
Roberts, R. J. *et al.* (2003). *Nucleic Acids Res.* **31**, 1805–1812.