Ly Huynh Thi Yen, Suk-Youl Park
and Jeong-Sun Kim*Department of Chemistry, Chonnam National
University, Gwangju 500-757, Republic of
KoreaCorrespondence e-mail:
jsunkim@chonnam.ac.krReceived 1 November 2013
Accepted 27 February 2014

Cloning, crystallization and preliminary X-ray diffraction analysis of an intact DNA methyltransferase of a type I restriction–modification enzyme from *Vibrio vulnificus*

Independently of the restriction (HsdR) subunit, the specificity (HsdS) and methylation (HsdM) subunits interact with each other, and function as a methyltransferase in type I restriction–modification systems. A single gene that combines the HsdS and HsdM subunits in *Vibrio vulnificus* YJ016 was expressed and purified. A crystal suitable for X-ray diffraction was obtained from 25% (w/v) polyethylene glycol monomethylether 5000, 0.1 M HEPES pH 8.0, 0.2 M ammonium sulfate at 291 K by hanging-drop vapour diffusion. Diffraction data were collected to a resolution of 2.31 Å using synchrotron radiation. The crystal belonged to the primitive monoclinic space group $P2_1$, with unit-cell parameters $a = 93.25$, $b = 133.04$, $c = 121.49$ Å, $\beta = 109.7^\circ$. With four molecules in the asymmetric unit, the crystal volume per unit protein weight was $2.61 \text{ Å}^3 \text{ Da}^{-1}$, corresponding to a solvent content of 53%.

1. Introduction

Bacterial restriction–modification (R–M) systems confer antiviral defence against foreign genetic materials by recognition of the methylation state of specific sequences (Krüger & Bickle, 1983). They are classified into three types, named I, II and III, based on their composition and cofactor requirements, the nature of their target sequence, and the position of the DNA cleavage site with respect to the recognition sequence (Murray, 2000). Type I systems are multi-functional enzyme complexes that act as a methyltransferase (MTase), an ATPase, a translocase and a nuclease. The complex includes polypeptides encoded by *hsdR*, *hsdM* and *hsdS*, where *hsd* denotes the ‘host specificity of DNA’. R_2M_2S heteropentamers cleave DNA thousands of base pairs away from the recognition sequence when two bipartite sequences are not methylated (Murray, 2000; Dryden *et al.*, 2001), while they methylate the N6 atom of the nonmethylated adenine base when one of the two specific DNA sequences is methylated (Sistla & Rao, 2004; Tock & Dryden, 2005). Independently of the HsdR (restriction) subunit, HsdM (methylation) and HsdS (specificity) subunits have been known to form an active heterotrimeric MTase complex.

The mechanism of type I R–M systems has been deciphered from the structural information on each subunit of the complex. The HsdS structure from *Methanococcus jannaschii* revealed two target-recognition domains (TRDs) attached on both ends of two anti-parallel α -helices of conserved regions (CRs) in the middle and at the end of the amino-acid sequence (Kim *et al.*, 2005), which shows an intramolecular pseudo-dyad symmetry. Two other HsdS structures from *Mycoplasma genitalium* (Calisto *et al.*, 2005) and *Thermoplasma volcanium* (Gao *et al.*, 2011) have shown the same overall organization of three domains with a slight difference in the inter-domain orientation. The fragment structure of a putative HsdR subunit from *V. vulnificus* YJ016 has shown three globular domains forming the N-terminal nuclease domain and two consecutive RecA-like domains (Uyen *et al.*, 2009). Another fragment structure of an *EcoR124I* HsdR subunit revealed a fourth α -helical domain at the C-terminus (Lapkouski *et al.*, 2009). The HsdM structure from *V. vulnificus* YJ016 is closely related to that of the N6 DNA MTase from *Thermus aquaticus* (M.TaqI; Park *et al.*, 2012).

An electron microscopy (EM) study of *M.EcoKI* MTase in complex with a DNA-mimicking anti-restriction protein Ocr

suggested inter-subunit interactions of HsdS and HsdM. In this model, the N-terminal domains of two HsdM subunits interact with each other *via* α -helices, and the C-terminal region of the HsdM subunit contacts the HsdS subunit with the intervening Ocr between two HsdMs (Kennaway *et al.*, 2008). Based on the *T. tengcongensis* HsdS structure that has a bent CR α -helix, an open DNA-binding model by MTase has been proposed (Gao *et al.*, 2011). More recently, a pentameric model for the *EcoR124I* type I R–M and DNA complex was constructed from the EM map, which suggested that the open and closed forms were equilibrated prior to DNA binding to the MTase core (Kennaway *et al.*, 2012). However, computational models have a limited ability to show interactions among subunits in detail.

Besides M₂S MTase, a single HsdM subunit bound to a truncated HsdS subunit that contains a single TRD and a single CR has been thought to form an active MTase, named the MS_{1/2} MTase. The central CR helix and the TRD of the HsdS were incompletely repeated at the C-terminus (MacWilliams & Bickle, 1996). For example, *EcoDXXI*, which lacks the C-terminal half of the HsdS polypeptide, formed a functional MTase that recognized the same

bipartite target sequences (Meister *et al.*, 1993), indicating that the half S_{1/2} subunit may self-assemble to form a full HsdS-like structure with an intermolecular dyad symmetry. Therefore, a structural study with an intact protein composed of the fused HsdM and HsdS_{1/2} subunits may elucidate the subunit interaction in the type I R–M MTases. We expressed one single-chain MS_{1/2} MTase from *V. vulnificus* YJ016 (vvMTase) and carried out a preliminary crystallographic study.

2. Materials and methods

2.1. Cloning, expression and purification of intact DNA MTase of type I R–M system

The *V. vulnificus* gene coding for an intact DNA MTase (vvMTase; gi:37680386; Met1–Ser604) was amplified by the polymerase chain reaction (PCR) using *V. vulnificus* genomic DNA. The PCR product of vvMTase was cloned into pProEX HTa expression vector (Invitrogen) using *Bam*HI and *Not*I restriction enzymes (Biolabs).

Figure 1 Sequence analysis of vvMTase. (a) Analysis using the Pfam protein families database (Pfam; <http://pfam.sanger.ac.uk>) shows the presence of the methyltransferase at the N-terminus and HsdS at the C-terminus. (b), (c) Sequence alignment of vvMTase with vvHsdM and vvHsdS using the *Constraint-based Multiple Alignment Tool (COBALT)*; <http://www.ncbi.nlm.nih.gov/tools/cobalt>. The N-terminal 463 residues and the C-terminal 175 residues of vvMTase were aligned with vvHsdM (Met1–Lys530) and the C-terminus of vvHsdS (Lys209–Glu389), respectively. The abbreviations used in this figure are vvMTase for MTase from *V. vulnificus* (gi:37680386), vvHsdM for HsdM from *V. vulnificus* (gi:37680450) and vvHsdS for HsdS from *V. vulnificus* (gi:37680389).

This construct expresses the cloned gene fused with an N-terminal six-His tag and the tobacco etch virus (TEV) cleavage sequence. The expression construct was transformed into *Escherichia coli* BL21 Gen-X and grown in Luria–Bertani (LB) media supplemented with 100 $\mu\text{g ml}^{-1}$ ampicillin at 310 K. When the optical density at 600 reached 0.6, recombinant protein was expressed with 1.0 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) for an additional 16 h at 291 K. The culture was harvested by centrifugation at 5000g at 277 K. The cell pellet was resuspended in an ice-cold buffer consisting of 20 mM Tris–HCl pH 7.5, 500 mM NaCl, 10 mM β -mercaptoethanol (buffer A) and disrupted by ultrasonication. Cell debris was removed by centrifugation at 11 000g for 30 min. The fusion protein was purified using a 5 ml HisTrap HP Chelating column (GE Healthcare, Uppsala, Sweden) and the bound protein was eluted using an imidazole gradient to 500 mM imidazole in buffer A. The eluted protein fractions were mixed with recombinant TEV protease to cleave the six-His tag at the N-terminus and dialyzed to remove salts. The protein containing five additional amino acids (YFQGA) at the N-terminus was further purified using a 5 ml HiTrapQ HP anion-exchange column (GE Healthcare, Uppsala, Sweden) to more than 90% purity, as estimated by SDS–PAGE (data not shown).

2.2. Crystallization

For crystallization, the purified vvMTase protein was concentrated to 6.5 mg ml^{-1} in 20 mM Tris–HCl pH 7.5, 300 mM NaCl, 10 mM β -mercaptoethanol, 5% (v/v) glycerol. The protein concentration was determined using an extinction coefficient at 280 nm of 0.788 $\text{mg ml}^{-1} \text{cm}^{-1}$, which was calculated from its amino-acid sequence. The initial crystallization of the target protein was performed by sparse-matrix screening (Jancarik & Kim, 1991) using Crystal Screen and Crystal Screen 2 (Hampton Research, Riverside, California, USA) at 291 K with the hanging-drop vapour-diffusion method. For each crystallization trial, a 2 μl drop was prepared by mixing 1 μl purified vvMTase protein solution with an equal volume of reservoir solution. The reservoir contained 60 μl of the precipitating solution. Initially small crystals were obtained using solution No. 26 of Crystal Screen 2 [30% (w/v) polyethylene glycol monomethylether 5000, 0.1 M MES pH 6.5, 0.2 M ammonium sulfate]. In order to obtain better crystals, this initial condition was optimized by changing the precipitant and protein concentrations, the buffer pH, the temperature and the vapour-diffusion method, for example, hanging-drop or sitting-drop vapour diffusion.

Figure 2

Representative crystals of vvMTase. The crystal grew at 291 K within 7 d to maximum dimensions of approximately $0.01 \times 0.05 \times 0.1$ mm.

Table 1

Data collection statistics for vvMTase.

Values in parentheses are for the highest resolution shell.

Synchrotron	5C SBII, PAL
Wavelength (Å)	0.98
Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 93.25$, $b = 133.04$, $c = 121.49$, $\alpha = \gamma = 90$, $\beta = 109.7$
Resolution range (Å)	50–2.31 (2.37–2.31)
Completeness (%)	91.7 (83.9)
$R_{\text{merge}}^{\dagger}$ (%)	13.4 (30.9)
Measured reflections	284629
Unique reflections	110965
Multiplicity	2.6 (2.0)
Temperature (K)	100
Matthews coefficient ($\text{\AA}^3 \text{Da}^{-1}$)	2.61
Solvent content (%)	53
No. of molecules in asymmetric unit	4
Mean $I/\sigma(I)$	9.67 (2.37)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity of an individual reflection and $\langle I(hkl) \rangle$ is the mean intensity of that reflection.

2.3. X-ray diffraction data collection

For the diffraction experiment, the crystals were briefly immersed into the precipitant solution containing 15% (v/v) glycerol as a cryo-protectant and immediately placed in a 100 K nitrogen-gas stream. Native X-ray diffraction data were collected at the 5C SBII Pohang Accelerator Laboratory (PAL, Republic of Korea) with 1° oscillation per image and a crystal-to-detector distance of 350 mm. The crystal was exposed for 10 s per image at a wavelength of 0.98 Å. A data set of 180 images was collected to a resolution of 2.31 Å from a single crystal. The data were indexed and scaled using HKL-2000 (Otwinowski & Minor, 1997). The data-collection statistics are summarized in Table 1.

3. Results and discussion

The intact $\text{MS}_{1/2}$ MTase from *V. vulnificus* (vvMTase) is comprised of 638 residues. Its sequence analysis indicates that approximately 450 and 180 residues at the N- and C-terminus correspond to the HsdM and the half HsdS subunit, respectively (Fig. 1). Secondary-structure prediction suggested that approximately 50 residues at the C-terminus would form the long CR α -helix of HsdS (data not shown). For the structural study, 34 amino acids at the extreme C-terminus were truncated. The truncated vvMTase protein was successfully expressed in *E. coli* and the purified protein was concentrated to 6.5 mg ml^{-1} in a buffer consisting of 20 mM Tris–HCl pH 7.5, 300 mM NaCl, 10 mM β -mercaptoethanol, 5% (v/v) glycerol. Crystals suitable for diffraction experiments (Fig. 2) were obtained within 7 d by the hanging-drop vapour-diffusion method at 291 K, in which 1 μl protein solution was mixed with 1 μl reservoir solution. The mixture was equilibrated against 200 μl reservoir solution consisting of 25% (w/v) polyethylene glycol monomethylether 5000, 0.1 M HEPES pH 8.0, 0.2 M ammonium sulfate. The dimensions of the crystal used for data collection were approximately $0.01 \times 0.05 \times 0.1$ mm and the crystal diffracted to a resolution of 2.10 Å (Fig. 3). Owing to radiation damage during data collection, the scaled data were reduced to 2.31 Å resolution. The crystal belonged to the primitive monoclinic space group $P2_1$, with unit-cell parameters $a = 93.25$, $b = 133.04$, $c = 121.49$ Å, $\alpha = \gamma = 90$, $\beta = 109.7^\circ$. Since a self-rotation function suggested four vvMTase molecules of 68 kDa in the asymmetric unit, the Matthews coefficient was $2.61 \text{ \AA}^3 \text{Da}^{-1}$, corresponding to a solvent content of 53% (Matthews, 1968). We have attempted molecular-replacement with *Phaser* (McCoy *et al.*, 2007) and

Figure 3

A representative X-ray diffraction image from a vvMTase crystal. The crystal was exposed for 10 s over a 1° oscillation range. The edge of the detector corresponds to a resolution of 2.10 Å.

MOLREP (Vagin & Teplyakov, 2010) from the *CCP4* program suite (Winn *et al.*, 2011) using the reported structures of the HsdS subunit (PDB code 1yf2; less than 30% identity; Kim *et al.*, 2005) and the HsdM subunit (PDB entry 3ufb; less than 25% identity; Park *et al.*, 2012) as search models, but have not yet been successful. Therefore, the crystal structure of vvMTase is now being solved by anomalous dispersion methods with selenium as the anomalous scatterer.

This study was financially supported by Chonnam National University, 2011 (Project administration No. 2011-0647).

References

- Calisto, B. M., Pich, O. Q., Piñol, J., Fita, I., Querol, E. & Carpena, X. (2005). *J. Mol. Biol.* **351**, 749–762.
- Dryden, D. T. F., Murray, N. E. & Rao, D. N. (2001). *Nucleic Acids Res.* **29**, 3728–3741.
- Gao, P., Tang, Q., An, X., Yan, X. & Liang, D. (2011). *PLoS One*, **6**, e17346.
- Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
- Kennaway, C. K., Kosinska, A. O., White, J. H., Tuszyńska, I., Cooper, L. P., Bujnicki, J. M., Trinick, J. & Dryden, D. T. F. (2008). *Nucleic Acids Res.* **37**, 762–770.
- Kennaway, C. K. *et al.* (2012). *Genes Dev.* **26**, 92–104.
- Kim, J.-S., DeGiovanni, A., Jancarik, J., Adams, P. D., Yokota, H., Kim, R. & Kim, S.-H. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 3248–3253.
- Krüger, D. H. & Bickle, T. A. (1983). *Microbiol. Rev.* **47**, 345–360.
- Lapkouski, M., Panjikar, S., Janscak, P., Smatanova, I. K., Carey, J., Ettrich, R. & Csefalvay, E. (2009). *Nature Struct. Mol. Biol.* **16**, 94–95.
- MacWilliams, M. P. & Bickle, T. A. (1996). *EMBO J.* **15**, 4775–4783.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Meister, J., MacWilliams, M., Hübner, P., Jütte, H., Skrzypek, E., Piekarowicz, A. & Bickle, T. A. (1993). *EMBO J.* **12**, 4585–4591.
- Murray, N. E. (2000). *Microbiol. Mol. Biol. Rev.* **64**, 412–434.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Park, S.-Y., Lee, H.-J., Song, J.-M., Sun, J., Hwang, H.-J., Nishi, K. & Kim, J.-S. (2012). *Acta Cryst.* **D68**, 1570–1577.
- Sistla, S. & Rao, D. N. (2004). *Crit. Rev. Biochem. Mol. Biol.* **39**, 1–19.
- Tock, M. R. & Dryden, D. T. F. (2005). *Curr. Opin. Microbiol.* **8**, 466–472.
- Uyen, N. T., Park, S.-Y., Choi, J.-W., Lee, H.-J., Nishi, K. & Kim, J.-S. (2009). *Nucleic Acids Res.* **37**, 6960–6969.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.