Structure of the Q237W mutant of HhaI DNA methyltransferase: an insight into protein-protein interactions

Aiping Dong, Emory University
Lan Zhou, Emory University
Xing Zhang, Emory University
Shawn Stickel, New England Biolabs
Richard J. Roberts, New England Biolabs
Xiaodong Cheng, Emory University

Journal Title: Biological Chemistry
Volume: Volume 385, Number 5
Publisher: De Gruyter | 2004-01-14
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1515/BC.2004.041
Permanent URL: https://pid.emory.edu/ark:/25593/mz4h1

Final published version: http://dx.doi.org/10.1515/BC.2004.041

Copyright information:
© by Walter de Gruyter

Accessed January 7, 2024 9:59 PM EST
Structure of the Q237W mutant of HhaI DNA methyltransferase: an insight into protein-protein interactions

Aiping Dong1,a, Lan Zhou1b, Xing Zhang1, Shawn Stickel2, Richard J. Roberts2 and Xiaodong Cheng1,*

1 Department of Biochemistry, Emory University School of Medicine, 1510 Clifton Road, Atlanta, GA 30322, USA
2 New England Biolabs, 32 Tozer Road, Beverly, MA 01915, USA
*Corresponding author
e-mail: xcheng@emory.edu

Abstract

We have determined the structure of a mutant (Q237W) of HhaI DNA methyltransferase, complexed with the methyl-donor product AdoHcy. The Q237W mutant proteins were crystallized in the monoclinic space group C2 with two molecules in the crystallographic asymmetric unit. Protein-protein interface calculations in the crystal lattices suggest that the dimer interface has the specific characteristics for homodimer protein-protein interactions, while the two active sites are spatially independent on the outer surface of the dimer. The solution behavior suggests the formation of HhaI dimers as well. The same HhaI dimer interface is also observed in the previously characterized binary (M. HhaI-AdoMet) and ternary (M. HhaI-DNA-AdoHcy) complex structures, crystallized in different space groups. The dimer is characterized either by a non-crystallographic two-fold symmetry or a crystallographic symmetry. The dimer interface involves three segments: the amino-terminal residues 2–8, the carboxy-terminal residues 313–327, and the linker (amino acids 179–184) between the two functional domains – the catalytic methylation domain and the DNA target recognition domain. Both the amino- and carboxy-terminal segments are part of the methylation domain. We also examined protein–protein interactions of other structurally characterized DNA MTases, which are often found as a 2-fold related ‘dimer’ with the largest dimer interface area for the group-β MTases. A possible evolutionary link between the Type I and Type II restriction-modification systems is discussed.

Keywords: DNA methyltransferase; evolutionary link; protein-protein interactions; Type I and II restriction-modification systems.

Introduction

The DNA cytosine methyltransferase (MTase) M. HhaI recognizes the sequence 5′-CCGCG-3′ and methylates the internal cytosine by flipping the target cytosine completely out of the DNA duplex (Klimašauskas et al., 1994). The remainder of the duplex is left as nearly intact B-form and the DNA gap created by base flipping is filled by insertion of the side chain of Gln237. No covalent bonds are broken to carry out the base flipping (Roberts and Cheng, 1998) – only the base-pairing hydrogen bonds are broken and the π-stacking interactions between adjacent base pairs are lost. Insertion of Gln237 into the π-gap restores the hydrogen-bonding network to the orphan guanine but interrupts the aromatic manifold by which charge migration through the DNA can occur. However, a mutant of M. HhaI in which Gln237 is replaced with tryptophan (Q237W), an aromatic heterocyclic amino acid, appeared to restore charge migration through the DNA duplex, implying that the tryptophan insertion into the DNA gap maintains the π-stacked array (Rajski et al., 1999). In an effort to study the nature of tryptophan insertion into the DNA gap, we attempted to crystallize the ternary complex containing the Q237W mutant, the methyl-donor product AdoHcy, and a short oligonucleotide containing the HhaI recognition sequence. Although no DNA is incorporated in the crystal, the mutant enzyme was crystallized in a new space group with two molecules per asymmetric unit. This observation prompted us to examine the protein–protein interactions in different crystal lattices as well as in solution.

Early solution studies suggested that some DNA MTases exist predominantly as dimers: M. Rsrl shows partial dimerization in solution (Kaszubska et al., 1989), two DpnI MTases appear to exist as dimers (de la Campa et al., 1987), and M. Mspl dimerizes in solution at high concentration (>3 mg/ml; Dubey et al., 1992), while some MTases remain monomeric in solution (M. BamHI, Nardone et al., 1984; M. EcoRI, Rubin and Modrich, 1977). In experiments involving crosslinking of the BamHI MTase subunits, it was shown that the enzyme in a free state exists as a dimer (Malgyn et al., 2001). A similar behavior was observed for adenine M. CcrM from C. crescentus dimerized at physiological concentrations, and the dimerization does not affect DNA methylation (Shier et al., 2001).

In the absence of DNA substrate, M. HhaI, M. TaqI, M. PvuII, and M. MboII were crystallized with two molecules per asymmetric unit (Cheng et al., 1993; Labahn et al., 1994; Gong et al., 1997; Osipiuk et al., 2003), while M. DpnII, M. Rsrl, and M. EcoT4Dam with one molecule (Tran et al., 1998; Scavetta et al., 2000; Yang et al., 2003). In the presence of DNA substrate, M. HhaI-DNA existed as one protomer and a single copy of double-stranded DNA (Klimašauskas et al., 1994), while M. HaeIII-DNA and
The M.HhaI dimer in the crystal lattices and solution. (A) Two views of the dimer of the M.HhaI mutant Q237W. The insert shows the side chains from one protomer in yellow and from the other in black. (B) HhaI MTase dimer in the presence of DNA. A schematic drawing shows the MTase domain in green, the TRD domain is indicated in red, and the dimer interface in yellow. (C) Elution profile of M.HhaI on a Superdex 2000HR column. (D) Cross-linking of M.HhaI with disuccinimidyl suberate (DSS). The concentration of protein and cross-linker and the size of the molecular mass marker are indicated. The amount of protein (5 μg) loaded onto the gel was the same in all five slots.

M.TaqI-DNA complexes contained two such complexes in the asymmetric unit (Reinisch et al., 1995; Goedecke et al., 2001). For non-DNA AdoMet-dependent MTases, many enzymes are monomeric, while some exhibit features of dimerization and/or oligomerization. For example, small molecule glycine N-MTase is a tetrameric enzyme (Takata et al., 2003), guanidinoacetate MTase forms a dimer (Komoto et al., 2002), RNA MTases of the SPOUT family are dimers with their active sites constructed as a topological knot (Michel et al., 2002; Nureki et al., 2002), and protein arginine MTases (PRMTs) form ring-like dimers that can be further oligomerized in some cases (Weiss et al., 2000; Zhang et al., 2000; Zhang and Cheng, 2003). Here we examined the protein-protein interactions of M.HhaI in three different crystal lattices. The dimerization behavior of M.HhaI in solution and its potential link to other types of restriction-modification systems are also discussed.

Results

Structure of the Q237W mutant

Although a short DNA duplex containing a HhaI recognition site was used for crystallization, the cognate oligonucleotide was missing from the new crystal structure. There are two possible reasons. First, although the mutant methylates DNA in vivo and in vitro, the specific activity is only about 2% of the wild-type enzyme (Mi et al., 1995). Second, Trp237 radical formation would be thermodynamically feasible but appears not to yield significant amounts of kinetically trapped products (Rajski et al., 1999). However, the bound AdoHcy indicates that cofactor binding is not affected by the mutation.

Q237W mutant proteins were crystallized in the C-centered monoclinic space group C2 with two molecules in the crystallographic asymmetric unit. The two protomers present in the asymmetric unit are very similar: least squares superposition of the two molecules gave a root-mean-square-deviation of 0.647 Å for 307 pairs of Cα atoms. The overall structure of Q237W is shown in Figure 1A. The structure contains two parts: a catalytic methylation domain (green) and a DNA binding domain (red). The methylation domain is formed by a mixed seven-stranded β-sheet (green), a characteristic feature of the class-I AdoMet-dependent MTases (Schubert et al., 2003). While a contiguous region forms the DNA binding domain (red), the catalytic domain consists of two non-contiguous regions: the N-terminal region (green) and a carboxy-terminal helix αZ (yellow) located on one side of the β-sheet.
While the mutant Q237W-AdoHcy was crystallized in the C-centered monoclinic space group C2, the M.HhaI-AdoMet complex was previously crystallized in the monoclinic space group P2₁ (Cheng et al., 1993; O’Gara et al., 1999). The crystallization conditions (polyethylene glycol 4000 vs. ammonium sulfate) and crystal properties such as space groups (P2₁ vs. C2) and unit cell dimensions (a = 55.3 Å, b = 74.68 Å, c = 90.96 Å, and β = 103° vs. a = 155.0 Å, b = 47.0 Å, c = 122.7 Å, and β = 121°) are completely different in these two cases. One common feature is that both crystals contain two molecules per asymmetric unit and the loop containing the residue 237 is not involved in any crystal packing contacts. Furthermore, in both monoclinic lattices, a pair of the protomers is related through the carboxy-terminal helix αZ (yellow) by a two-fold non-crystallographic rotation. A more intriguing finding is that the same dimer interface also exists in the M.HhaI-DNA-AdoHcy complexes (Klimašauskas et al., 1994; Vilkaitis et al., 2000); in the ternary complexes, the dimer is characterized by a 2-fold crystallographic symmetry in the rhombohedral space group R32 (Figure 1B). The two DNA substrates, and hence the active sites, are located on the outer surface of the dimer. These observations prompted us to examine the protein-protein interactions in different crystal lattices.

**HhaI dimer formation in crystal lattices**

We calculated all possible protein-protein interfaces in the lattices of P2₁, C2, and R32 based on the change in solvent accessible surface area (ΔASA) when going from a monomeric to a dimeric state (http://www.biochem.ucl.ac.uk/bsm/PP/server; Jones and Thornton, 1996). The largest change in ΔASA was found for a pair of the protomers related through the carboxy-terminal helix αZ (yellow) by a two-fold non-crystallographic rotation. Several indications (Table 1) that this interface has the characteristics specific for dimer protein-protein interactions (Jones and Thornton, 1995): (i) the 652 Å² of ΔASA is within the limits, though in the lower end, of the 32 homodimers examined (Jones and Thornton, 1996); (ii) the interface is mainly hydrophobic.
The 2-fold related DNA MTase structures.
One protomer is shown in yellow and the other in green. The dimer is related by a 2-fold non-crystallographic symmetry (NCS) or crystallographic symmetry (NC). The dimer interface area (∆ASA) was calculated using AREAIMOL of CCP4.

formed by Tyr313, Tyr316, Phe324, Pro326, and Tyr327 (Figure 1A) and 62.5% of the atoms in the interface are non-polar; (iii) the interface also involves two hydrogen bonds – the side chain of Asn323 to the side chain of Gln8 and the side chain of Asn317 to the main chain of Tyr327; (iv) three segments (amino acids 2–8, 179–184, 313–327) of the polypeptide chain are involved in the interface. The amino-terminal segment and the carboxy-terminal segment are part of the methylation domain (colored green in Figure 1A and B), while the residues 179–184 are part of the linker between the methylation and DNA-binding domains.

HhaI dimer formation in solution
There are several indications that this structural dimer may also exist in solution: (i) gel filtration results suggest that M.HhaI has an apparent molecular mass of a dimer. On a Superdex 200 column (Pharmacia), M.HhaI eluted as a 60 kDa globular protein (Figure 1C), larger than the calculated molecular mass (37 kDa). (ii) Dynamic light scattering measurement at a range of protein concentrations suggested that the hydrodynamic radius of M.HhaI is 4 nm and would correspond to an 85 kDa globular protein (Table 2). This dimension (2×4=8 nm) is between the shorter (4.7 nm) and the longer (12 nm) dimensions of the dimer (Figure 1A). (iii) A portion of M.HhaI formed a covalent dimer after cross-linking by glutaraldehyde or disuccinimidyl suberate (Figure 1D). These data suggest that the dimers seen in the crystal structures of M.HhaI could also exist in solution.

Comparison with other structurally characterized DNA MTases
At this time, the structural characterization of nine DNA MTases has been reported (Table 3). These include three generating 5mC (M.HhaI, M.HaeIII, and human Dnmt2), one generating N4mC (M.PvuII), and five generating N6mA (M.TaqI, M.DpnII, M.RsrI, M.MboIIA, and M.EcoT4Dam). We examined the protein-protein interactions in the crystal lattices to search for a pair of protomers related by a 2-fold rotation, either through non-crystallographic symmetry (NCS) if two proteins or protein-DNA complexes existed in the asymmetric unit or through crystallographic symmetry (CS) if only one protein or protein-DNA complex was present in each asymmetric unit.

With the exception of M.DpnII, all other structurally characterized MTases can be found as a 2-fold related
‘dimer’ (Figure 2). The dimer interface does not block the active site or the DNA binding cleft. The ΔASA values are distributive: (i) the largest ΔASA values are from M.MboIIA (1966 Å²), M.RsrI (1706 Å²), and M.Pvull (2317 Å²) (Figure 2D–F). These values are at the high end of the homodimers examined by Jones and Thornton (1996). All three are β-group DNA MTases (Malone et al., 1995) and show, interestingly, a protomer interface between two molecules (Osipiuk et al., 2003). Most likely, these β-group DNA MTases are dimeric in solution (Kaszubska et al., 1989; Osipiuk et al., 2003). (ii) For the 5mC MTases examined (Figure 2A–C), the ΔASA values, being at the low end but within the limit of Jones and Thornton (1996), are approximately 3 times less than that of β-group enzymes. As demonstrated by M.MspI (Dubey et al., 1992), these enzymes could dimerize in solution in a concentration-dependent fashion. (iii) The smallest ΔASA value is from M.TaqI in the absence of DNA (164 Å²); this value is outside of the lower limit of Jones and Thornton (1996). This may explain that the M.TaqI interface shown in Figure 2G is not observed for the M.TaqI-DNA complex, even though two such complexes existed in the asymmetric unit. (iv) Similarly, the M.EcoT4Dam interface (Figure 2H) observed in the presence of DNA (404 Å²) does neither exist in the substrate-free T4Dam nor in a GATC-related orthologous M.DpnI structure, although both α-group MTases were previously suggested to dimerize in solution (de la Campa et al., 1987; Evdokimov et al., 2002).

## Discussion

The type II restriction-modification systems consist of two independent components that, respectively, restrict and modify DNA (Roberts and Halford, 1993). Restriction involves a homodimeric endonuclease (ENase) cleaving DNA by hydrolyzing the phosphodiester backbone on both strands, while modification involves a MTase adding a methyl group to a DNA base on one strand at a time. Both the restriction and the modification activities recognize the same DNA sequence. The cleavage occurs at fixed positions that are symmetrically-disposed relative to the symmetrical recognition sequence, and within or adjacent to that sequence. The methylation occurs at a position within the palindromic recognition sequence that blocks the paired restriction activity.

Most of the structurally characterized Type II ENase-DNA complexes comprise a protein dimer and a single copy of a double-stranded oligonucleotide in the crystallographic asymmetric unit, the smallest repeating unit within a crystal. However, for the R.BglII-DNA complex, there was only one protein subunit and one DNA strand in the asymmetric unit, and thus the biologically active homodimer was generated by a two-fold axis of crystallographic symmetry (Newman et al., 1998).

Although the functional significance, if any, of the Type II MTase dimer is unclear, one possibility is that the dimeric form of Hhal MTase allows formation of a large enzyme-substrate network with high molecular weight DNA, as demonstrated for Serratia nuclease (Franke et al., 1999). Although the two subunits of Serratia nuclease function independently of each other, they bind simultaneously to one macromolecular DNA, forming a large network between the enzyme and the substrate. Homodimeric DNA MTases may serve the same function. At present, we do not have any evidence against the notion that two monomers act independently from each other in the Hhal homodimer.

Type II MTase dimers might also provide an evolutionarily link to other types of restriction-modification systems. Type II MTases are bilobal molecules with a catalytic methylation domain (shown in green in Figure 1A and B), responsible for AdoMet binding and the methyl transfer reaction, and a target recognition domain (TRD) responsible for DNA binding. In some cases, MTases are expressed as two polypeptides that associate to form the active enzymes (Kareman and de Waard, 1990; Lee et al., 1996). In addition, the Type IIS systems consist of two MTases (sometimes expressed as a single fusion protein) that individually modify the two asymmetric strands of the recognition sequences.

### Table 4 Summary of crystallographic data.

<table>
<thead>
<tr>
<th>Space group</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit cell dimensions</td>
<td>a=155.0 Å, b=47.0 Å, c=122.7 Å</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>24.54–2.66, 2.75–2.66</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>91.6, 89.1</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.082, 0.231</td>
</tr>
<tr>
<td>⟨/⟩</td>
<td>16.8, 3.0</td>
</tr>
<tr>
<td>Observed reflections</td>
<td>69 737, -</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>19 780, 1916</td>
</tr>
</tbody>
</table>

### Refinement

<table>
<thead>
<tr>
<th>Number of atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein</td>
</tr>
<tr>
<td>AdoHcy</td>
</tr>
<tr>
<td>SO₄</td>
</tr>
<tr>
<td>Water</td>
</tr>
<tr>
<td>R_factor</td>
</tr>
<tr>
<td>R_mol</td>
</tr>
<tr>
<td>Rms deviation from ideal</td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
</tr>
<tr>
<td>Bond angles (°)</td>
</tr>
<tr>
<td>Dihedrals (°)</td>
</tr>
<tr>
<td>Improper (°)</td>
</tr>
<tr>
<td>Estimated coordinate error</td>
</tr>
<tr>
<td>uzzati plot (Å)</td>
</tr>
<tr>
<td>Mean B value (Å²)</td>
</tr>
</tbody>
</table>
a clear link between Type I and Type II restriction-modification systems by utilizing gene fusions and the shuffling of structural and functional domains (Dryden, 1999). The M.
Hhal dimer, mediated mainly by the catalytic domain (M subunit) and the linker, but not the DNA binding domain (S subunit), may simply reflect such an evolutionary relationship.

Materials and methods

The mutant protein Q237W was engineered, expressed, and purified as described previously (Kumar et al., 1992; Mi et al., 1995). For recording the elution profile of M.
Hhal on a Superdex 2000HR column (Pharmacia, Piscataway, USA), the protein was loaded onto the column at a concentration of 5 mg/ml. The column buffer was 20 mM HEPES pH 7.0, 50 mM NaCl, 0.1 mM EDTA, 5% glycerol and 0.1% 2-mercaptoethanol. Total protein (5 mg) was loaded on a 10% SDS polyacrylamide gel. The dynamic light scattering experiments were performed at 16 °C using a DynaPro Molecular Sizing Instrument (Protein Solutions, High Wycombe, UK). Because polydispersity is greater than 25%, a multiple-exponential algorithm (regularization analysis) was employed.

The cofactor-free Q237W protein in 20 mM Tris-HCl, 100 mM NaCl, 5% glycerol, 1 mM EDTA, 1 mM DTT was concentrated to 25 mg/ml (0.68 mM) and incubated with AdoHcy and DNA (5'-TATC-3') in an equal molar ratio. Crystals were grown under the conditions of 2.2 M ammonium sulfate, 50 mM HEPES pH 7.0, 0.2 mM NaCl, 0.1 mM EDTA, 5% glycerol and 0.1% 2-mercaptoethanol. Total protein (5 mg) was loaded onto the column at a concentration of 5 mg/ml. The column buffer was 20 mM HEPES pH 7.0, 50 mM NaCl, 0.1 mM EDTA, 5% glycerol and 0.1% 2-mercaptoethanol. Total protein (5 mg) was loaded on a 10% SDS polyacrylamide gel. The dynamic light scattering experiments were performed at 16 °C using a DynaPro Molecular Sizing Instrument (Protein Solutions, High Wycombe, UK). Because polydispersity is greater than 25%, a multiple-exponential algorithm (regularization analysis) was employed.

The structure was solved by the molecular replacement program AmoRe (Navaza, 2001) using the refined protomer T250G structure (Vilkaitis et al., 2000) as the search model by deleting the side chain of residue 237; two solutions were found indicating each asymmetric unit contains two molecules. The structure was confirmed by the appearance of the electron densities corresponding to the side chains of Thr250 and Trp237. Numerous cycles of model refinement using XPLOR (Brunger et al., 1998) and manual rebuilding using O (Jones and Kjeldgaard, 1997) were conducted. A series of simulated annealing omit maps were used to correct the model. The non-crystallographic restraints were imposed on the two protein molecules throughout the refinement and were only released from protein side chains at the last cycle to account for a different packing environment with each molecule. The final model consists of 307 out of 328 amino acids in protein chain A, 314 amino acids in protein chain B, and two AdoHcy molecules. The missing residues, amino acid residues 80–99 in protomer A and residues 84–96 in protomer B that could not be modeled due to the poor quality of the electron density, are part of the active-site loop that undergoes large conformational changes upon DNA binding (Cheng et al., 1993; Klimašauskas et al., 1994).

Acknowledgments

We thank Susan Sunay (Emory University, Atlanta, USA) for dynamic light scattering measurement and gel filtration experiments and G.G. Wilson (New England Biolabs, Beverly, USA) for sharing his unpublished data. This work was supported in part by National Institutes of Health grants GM46127 to R.J.R. and GM49245 to X.C.

References


Received January 9, 2004; accepted February 9, 2004.