Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 334 (2005) 1233–1240
BBRC
www.elsevier.com/locate/ybbrc

Protein fragment complementation in M.HhaI DNA methyltransferase

Wonchae Choe [a,b,1], Srinivasan Chandrasegaran [b,*], Marc Ostermeier [a,*]

[a] Department of Chemical and Biomolecular Engineering, Johns Hopkins University, 3400 N. Charles St. Baltimore, MD 21218, USA
[b] Department of Environmental Health Sciences, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, MD 21205, USA

Received 24 June 2005
Available online 18 July 2005

Abstract

The 5mC DNA methyltransferase M.HhaI can be split into two individually inactive N- and C-terminal fragments that together can form an active enzyme in vivo capable of efficiently methylating DNA. This active fragment pair was identified by creating libraries of M.HhaI gene fragment pairs and then selecting for the pairs that code for an active 5mC methyltransferase. The site of bisection for successful protein fragment complementation in M.HhaI was in the variable region near the target recognition domain between motif VIII and TRD. This same region is the location of bifurcation in the naturally split 5mC methyltransferase M.AquI, the location for circular permutation in M.BssHII, and the location for previously engineered split versions of M.BspRI.

Keywords: M.HhaI; DNA methyltransferase; Protein fragment complementation; Leucine zipper; Incremental truncation; Protein engineering; M.AquI; M.BssHII; M.BspRI; DNA methylation

DNA methyltransferases (MTases) are DNA modifying enzymes that catalyze the transfer of a methyl group from *S*-adenosylmethionine to an adenine or cytosine base within a specific DNA sequence. DNA MTases are divided into three main groups based on the site of methylation. MTases that methylate the 5 position of cytosine are referred to as 5mC MTases. The modified base 5-methylcytosine is the only methylated base found in vertebrate and plant DNA. DNA methylation, especially in eukaryotes, has a profound effect on gene regulation. Epigenetic gene silencing by methylation is achieved by a non-mutational gene inactivation that occurs by the transfer of methyl groups to the cytosine residues of promoter sequences. This modification is faithfully propagated from mother to daughter cells. Because aberrations in cytosine-5-methylation are involved in many human genetic diseases, 5mC MTases are important targets for drug development to treat these diseases.

The crystal structure of *HhaI* cytosine-5-methyltransferase (M.HhaI: 5′GCGC3′) from *Haemophilus haemolyticus* was the first to be solved [1]. The structure of M.HhaI with its cognate sequence DNA showed that the enzyme flips the target cytosine residue that is to be methylated out of the DNA helix without any major distortion to the DNA [2]. M.HhaI has a large N-terminal domain and a smaller C-terminal domain connected by a 'long linker' region (approximately residues 175–191). The large domain is discontinuous, as the C-terminal end of the protein is a α-helix that is inserted into the large domain. Sequence comparison of 5mC MTases revealed 10 highly conserved protein motifs [3]. Motif I, also known as the FXGXG motif, is found in all SAM-dependent MTases and it appears to be involved in SAM binding; motif IV also known as the PC motif, which has a conserved cysteine residue, has been shown to be involved in enzyme catalysis; motifs II–V are

* Corresponding authors. Fax: +1 410 516 5510.
E-mail addresses: schandra@jhsph.edu (S. Chandrasegaran), oster@jhu.edu (M. Ostermeier).
[1] Present address: Department of Molecular Biology and Biochemistry, School of Medicine, Kyung Hee University, 1 Hoeki-Dong, Dongdaemoon-Ku, Seoul, 130-701, Korea.

0006-291X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2005.07.017

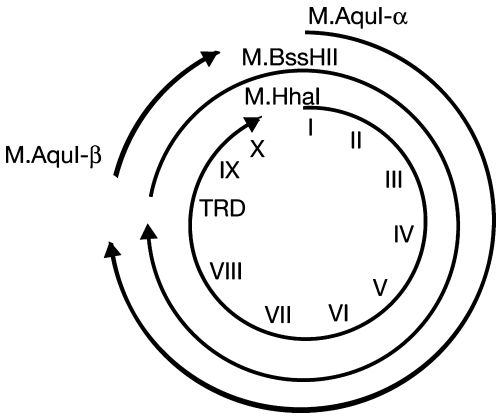

Fig. 1. Sequence comparison of M.HhaI with M.BssHII and M.AquI. The order of conserved motifs is circularly permuted in M.BssHII [6] and is split between two peptides in M.AquI [7].

likely responsible for the key interactions between the co-factor and the protein that are mediated through a set of hydrophobic amino acids [4,5]. The variable region between motifs VIII and IX (amino acids 171–271 in M.HhaI), which include a small sub region called the target recognition domain (TRD), shows the greatest heterogeneity in size and sequence among all 5mC MTases. Interestingly, 5mC MTases can have different assemblies of these motifs (Figs. 1 and 2).

Relative to most 5mC MTases, including M.HhaI, M.BssHII is circularly permuted within the variable region and the TRD, IX, and X motifs precedes motifs I–VIII [6]. The evolutionary origin of this circular permutation of motifs I–X within the 5mC methyltransferase family is not clear [10]. One possibility is that evolution could have proceeded by gene duplication, modifications of this gene pair (e.g., introduction of new stop and start codons) to create interacting fragment pairs, and finally fusion of these gene fragments in a new order. Such a mechanism is supported by the existence of a fragmented 5mC methyltransferase, M.AquI. M.AquI is comprised of two polypeptide chains coded for by partially overlapping reading-frames but is structurally organized quite similarly to other 5mC methyltransferases [7,11]. The 10 highly conserved motifs are present in the correct order on two distinct polypeptides: the α-polypeptide contains motifs I–VIII, and the β-subunit contains the TRD together with motifs IX and X. It has been shown that M.BspRI can be functionally bifurcated at several locations including within the variable region [9] (Fig. 2). The plasticity of the variable region to modifications is also evident in multi-specific mC5 MTases (those which methylate more than one specific DNA target) that have been shown to be tolerant to deletions, rearrangements and exchanges between different enzymes within this region [12]. In addition, some hybrids between M. HpaII and M.MspI in which the fusion was at the beginning or end of the variable region showed partial methylation activity [13].

Previous combinatorial searches for split fragments of *Escherichia coli* glycinamide ribonucleotide phosphotransferase [14] and aminoglycoside phosphotransferase (3′)-IIa [15] resulted in the identification of complementing fragments that were split both between and within sub-domains and at unexpected locations such as within secondary structural elements and within the active site. We were interested whether a combinatorial search for complementing protein fragments of M.HhaI would also result in fragment pairs split at similar locations to those naturally found in M.AquI and the engineered M.BspRI fragment pairs, or whether new bisection locations would be found. Furthermore, we also have a future interest in conferring additional properties to M.HhaI without affecting it primary function of site-specific methylation of DNA. Identification of bisection sites of complementing fragment partners is a first step towards this goal. Here, we report the identification of a bisection site within the M.HhaI molecule where new functionalities could potentially be introduced in the future without affecting its enzymatic activity.

Materials and methods

Bacterial strains and media. Escherichia coli K-12 strains ER1727 [*F′ proA + B + lacIq* Δ(*lacZ*)*M15*/*t-31 his-1 rpsL104* (*StrR*) Δ(*lacZ*)*r1 glnV44 xyl-7 mtl-2* Δ(*mcrBC-hsdRMS-mrr*)*2::Tn10mcrA1272::Tn10*(*-TetR*) *metB1 fhuA2*] and ER2267 [F′ *proA + B + lacIq* Δ(*lacZ*)*M15 zzf::mini-Tn*10 (KanR)/Δ(*argF-lacZ*)*U169 glnV44 e14-*(*McrA-*) *rfbD1? recA1 relA1? endA1 spoT1? thi-1* Δ(*mcrC-mrr*)*114::IS10*] were used as hosts in DNA subcloning experiments and library construction and were obtained from New England Biolabs (Beverly, MA). Library selection was carried using the *E. coli* strain GeneHogs [*F⁻ mcrA* Δ*(mrr-hsdRMS-mrcBC) f80lacZ*Δ*M15* Δ*lacX74 deoR recA1 endA1 araD139* Δ*(ara-leu)7697 galU galK rpsL(Str*^{R}*) nupG*] which was obtained from Invitrogen (Carlsbad, CA). Cells were grown in LB medium (1% w/v tryptone, 1% w/v NaCl, and 0.5% w/v yeast extract) at 37 °C. Chloramphenicol (75 μg/ml) and/or ampicillin (100 μg/ml) were added to media to maintain the plasmids, as needed.

Plasmids. Restriction enzyme digestions and ligations were performed as recommended by the manufacturer (New England Biolabs, Beverly, MA). DNA manipulations including subcloning and agarose gel electrophoresis were performed essentially as described elsewhere [16]. Plasmid DNA was purified by Qiagen plasmid midi prep kit. The following double-stranded oligonucleotide was digested with *Kpn*I and *Dra*III (sites underlined) and inserted into *Kpn*I/*Dra*III digested pDIM-N2 [17]: 5′-GCCAGC<u>GGTACC</u>ACCGAT***CCCGGG*** AAGCTTGGGGCGGAACAACTGAGGCGCGCCGCTTGGGGA AGAAAC**GAATTC**ATGATAA**ACTAGT**<u>CACTACGTG</u>GGAC-3′. This DNA sequence contains *Eco*RI and *Spe*I sites (bold) to allow for cloning of fragments from pDIM-C8. It also contains a *Sma*I site (bold and italics). The TetR operon was amplified from pACYC184 and inserted between the *Kpn*I and *Sma*I sites to create plasmid pDIM-N7. The gene encoding *M.HhaI* (Genbank Accession No. J02677, bases 437–1420) was amplified by PCR using chromosomal DNA purified from *Haemophilus heamolyticus* bacteria as template and primers complementary to the 5′ end (5′-GCGGGCCCATATGATTGAAAT AAAAGATAAACAG-3′) and 3′end (5′-GCGGGCCTCTAGATTA ATATGGTTTGAAATTTAATGA-3′) of the M.HhaI gene. The M.HhaI PCR product was digested with *Nde*I and *Xba*I and ligated into these same sites in pDIM-N7 under the control of lac promoter to