In Vivo Site-Specific DNA Methylation with a Designed Sequence-Enabled DNA Methylase

Wataru Nomura and Carlos F. Barbas, III*

The Skaggs Institute for Chemical Biology and the Departments of Chemistry and Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received January 25, 2007; E-mail: carlos@scripps.edu

As an alternative to the continual expression of transcriptional repressors to turn off genes after they have served their purpose, nature has developed epigenetic strategies that result in the covalent modification of DNA itself to induce heritable gene silencing. Mounting evidence supports the notion that once a genomic region has been targeted for silencing by acquisition of one or more covalent epigenetic marks, mark can be propagated and may influence acquisition of others.[1] If epigenetic modifications can be made specifically by the addition of targeted exogenous agents, new approaches to transcriptional therapy should result.

Methyltransferases recognize specific DNA sequences and transfer a methyl group from the cofactor S-adenosyl-L-methionine (AdoMet) to an amine on either cytosine or adenine.[2] The best characterized methyltransferase is HhaI DNA m5c-methyltransferase (M.HhaI) that recognizes the sequence 5′-GCGC-3′ and converts the internal cytosine (in bold) to 5-methylcytosine using a base-flipping mechanism.[3] In eukaryotes, CpG methylation serves as a signal for the recruitment of proteins that ultimately act to silence transcription.[4] Recently, fusions of zinc finger proteins (ZFPs) to full-length methyltransferases such as M.HpaII and fusions with the enzymatic domains derived from the murine enzymes Dnmt3a and 3b have been studied.[5] These fusion proteins methylate native target sites as well as sites adjacent to the binding sites for the attached ZFPs.[6] Although these proteins are biased in their activity, they do not direct the truly programmable site-specific methylation we aim to achieve. To address this challenge we sought to design methyltransferases that would act only at a targeted site by adopting a sequence-enabled reassembly strategy.[7] Our aim was to reassemble a fragmented methylase on DNA as directed by zinc finger binding, thereby restoring its activity at a specific site and providing for site-specific cytosine methylation (Figure 1).

We hypothesized that a functional and site-specific enzyme could be self-assembled on a particular DNA sequence using ZFPs appropriately fused to a recently described split M.HhaI enzyme.[8] If correctly designed, such a fragmented protein would be active only at the site of assembly and not elsewhere. To explore this hypothesis, each domain of split M.HhaI (N- and C-terminal domain fragments) was fused to previously characterized three-finger ZFPs, HS1 and HS2.[9] These ZFPs, HS1 and HS2, bind the DNA sequences 5′-GGGGCCGGA-3′ and 5′-GCCGCAGTG-3′, respectively. The N-terminal fragment was fused to HS2 to create MeNDHis and the C-terminal fragment was fused to HS1 to create MeCD proteins. A His$_6$ tag was fused to MeNDHis and an HA epitope tag was fused to MeCD to facilitate protein detection. A fusion of intact M.HhaI with HS2 (HS2Me) was also constructed (Figure 1A).

Genes for the fusion proteins were cloned downstream of a *pBAD* promoter with Shine−Dalgarno sequences for bacterial expression. A 26 bp DNA methylation target sequence, GCGC-ZFS, consisting of a GGCGCC site flanked by ZFP binding-sites was present on
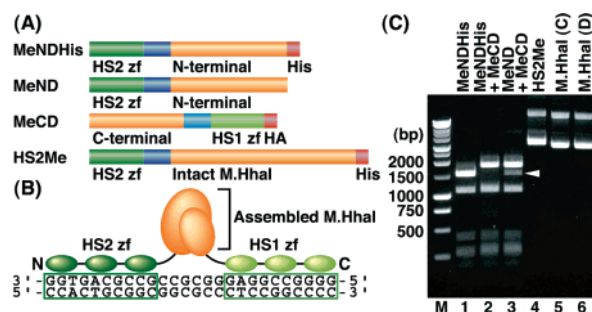

Figure 1. (A) Schematic drawing of Split-M.HhaI and ZFP fusions. Blue segment shows linker sequences. (B) Cartoon illustrating ZFP directed assembly of a fragmented DNA methylase on the target DNA sequences. The DNA sequence bound by ZFPs, HS2, and HS1, are boxed. (C) HhaI digestion assays for methylation detection. Plasmid DNA expressing the various enzyme fragments was isolated and digested with HhaI: lane M, marker; lane 1, MeNDHis; lane 2, MeNDHis + MeCD; lane 3, MeND + MeCD; lane 4, HS2Me; lanes 5 and 6, M.HhaI only (C, no digestion control; D, HhaI digestion). White arrow in lane 3 indicates a band created by inhibition of HhaI digestion owing to ZFP-targeted methylation.

the plasmid (Figure 1B). To stringently test the specificity of the assembled enzyme, 18 native M.HhaI GCGC sites without flanking ZFP binding sites were also encoded on the plasmid. Expression of the fusion proteins in *E. coli* was confirmed by Coomassie blue staining of cell extracts and western blotting using antibodies against His$_6$ and HA tags. DNA binding of fusion proteins was detected by ELISA as described previously.[9] DNA binding affinities of ZFPs HS1 and HS2 are 35 and 25 nM, respectively.[10]

As our first methylation test, HhaI restriction enzyme cleavage was performed on plasmid DNA recovered from *E. coli* that expressed our designed enzymes (Figure 1C). DNA cleavage by the HhaI endonuclease is inhibited when the internal cytosine in the GCGC site is methylated. In our assay, DNA fragmentation patterns of plasmid following HhaI restriction enzyme treatment will differ in accord with the methylation status of the various HhaI sites.[11a] HhaI endonuclease cleavage of plasmid coding MeNDHis and MeCD produces 20 fragments sized 312, 353, 390, 1108, and 1710 bp together with 15 fragments smaller than 300 bp. Site-specific methylation at GCGC-ZFS will inhibit endonuclease cleavage between the 353 and 1108 bp fragments, resulting in production of a novel 1461 bp fragment. Results of this study revealed that the MeNDHis fragment did not methylate any of the 19 HhaI sites when expressed alone (lane 1). To test ZFP directed self-assembly, methylation of the plasmid derived from cells expressing the combination of N- and C-terminal enzyme fragments, MeNDHis and MeCD, was evaluated. Here too, no band indicative of methylation was detected (lane 2). On the basis of the possibility that a His$_6$ tag could inhibit the reassembly of split domains, we further tested expression of MeND (identical to MeNDHis but lacking the C-terminal tag) and MeCD. Reassembly of this
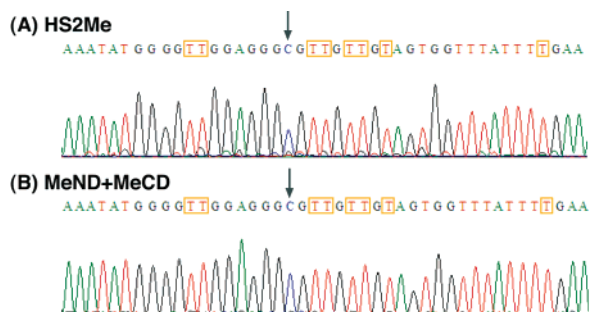
Figure 2. Site-specific methylation at the GCGC-ZFS detected by bisulfite sequencing. Arrows indicate methylcytosines. Thymines converted from cytosines after bisulfite reaction are boxed. Additional sequencing data for MeND+MeCD is available in Supporting Information Figure S7.

combination resulted in specific methylation of the GCGC-ZFS sequence (lane 3). Intact M.HhaI fusion protein (HS2Me) produced a different digestion pattern indicative of methylation at most of the 19 GCGC sites (lane 4). M.HhaI expression alone completely inhibited HhaI cleavage (lanes 5 and 6). Results concerning M.HhaI and HS2Me results were very similar supporting the idea that the native methylation activity of HS2Me is still dominant.[11b]

To confirm that the results obtained from HhaI restriction enzyme analysis were due to site-specific methylation of DNA, bisulfite sequencing was performed. Plasmid DNA recovered from cells expressing HS2Me or both MeND and MeCD showed methylation at GCGC-ZFS (Figure 2). Only HS2Me expression resulted in methylation at GCGC sequences lacking ZFP binding sites (non-targeted native M.HhaI sites); methylation was also detected at a site 1500 bp distal to the GCGC-ZFS site.[11b] This result suggests that the DNA binding activity of M.HhaI remains when the intact enzyme is fused to a heterologous DNA binding domain, a result consistent with earlier studies indicating how challenging site-specific methylation can be using a simple enzyme-fusion approach.[5,6] MeNDHis alone did not show any methylation activity.[11c] These sequencing results revealed that only plasmids derived from cells expressing the combination of N- and C-terminal enzyme fragments, MeND and MeCD, methylate DNA specifically at the designed site, without any background methylation at native M.HhaI sites. As additional evidence supporting the lack of nonspecific methylation in our approach, no methylation was observed in plasmids with mutant GCGC-ZFS sites.[11d]

High background methylation has been a difficult problem in approaches based on fusions of methyltransferases with heterologous DNA binding domains.[5,6] This difficulty stems from the fact that DNA recognition and catalytic activity are coupled in bacterial methyltransferases.[3b] In our sequence-enabled split M.HhaI approach, the site of the bisection of the native enzyme is located in the variable region near the target recognition domain between motif VIII and TRD. This site is adjacent to sequences found in the active site of the enzyme and residues utilized for DNA recognition. Only when the split enzyme is appropriately assembled is it active. Significantly, the absence of background methylation mediated by our methylase at native GCGC sites—despite high-level expression in *E. coli*—indicates that the native DNA binding and methylation activity of split-M.HhaI is absent or very substantially reduced by fusion with ZFPs and that ZFP binding to the designed site directs reconstitution of the fragmented enzyme.

Reassembly of split protein domains is a powerful strategy for creating conditionally active proteins that has been applied in other systems.[12] When this approach incorporates fusions with designed ZFPs, functional assembly of the active enzyme is conditionally dependent on sequence-specific DNA binding of ZFPs.[7] Previous studies have demonstrated activity in vitro with fusions created through sequence-enabled reassembly; however, specific activity in a living cell had not been previously demonstrated. Studies in living cells are challenging since the signal-to-noise ratio must be very high as manipulation of the substrate DNA and enzyme concentrations are difficult in vivo.

Application of our methylases in living systems should provide powerful tools to control gene expression. Indeed, methylation of a single CpG site within eukaryotic promoters, such as rat *p53*, can be sufficient to silence transcription.[13] Single-site CpG methylation directed by our sequence-enabled methylase should provide for silencing of transcription at virtually any desired site as ZFPs can now be readily prepared to bind most DNA sequences.[14] Beyond transcriptional regulation, we envision that our methylases will have applications in DNA tagging approaches and nanotechnology.[15]

In summary, we have developed sequence-enabled DNA methylases that direct programmable site-specific DNA methylation in vivo. These results encourage us to advance this split domain reassembly strategy in order to develop programmable zinc finger methylases that act at any specific CpG site in the mammalian genome to orchestrate heritable gene silencing.

Supporting Information Available: Materials and methods and additional supporting experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Bird, A. *Genes Dev.* **2002**, *16*, 6−21. (b) Richards, E. J.; Elgin, S. C. *Cell* **2002**, *108*, 489−500.
(2) (a) Hermann, A.; Gowher, H.; Jeltsch, A. *Cell. Mol. Life Sci.* **2004**, *61*, 2571−87. (b) Jeltsch, A. *Curr. Top. Microb. Immunol.* **2006**, *301*, 203−25.
(3) (a) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994**, *76*, 357−69. (b) Youngblood, B.; Buller, F.; Reich, N. O. *Biochemistry* **2006**, *45*, 15563−72.
(4) Wade, P. A. *BioEssays* **2001**, *23*, 1131−1137.
(5) (a) Xu, G. L.; Bestor, T. H. *Nat. Genet.* **1997**, *17*, 376−8. (b) McNamara, A. R.; Hurd, P. J.; Smith, A. E.; Ford, K. G. *Nucleic Acids Res.* **2002**, *30*, 3818−30. (c) Carvin, C. D.; Parr, R. D.; Kladde, M. P. *Nucleic Acids Res.* **2003**, *31*, 6493−501. (d) Smith, A. E.; Ford, K. G. *Nucleic Acids Res.* **2007**, *35*, 740−54. (e) Li, F.; Papworth, M.; Minczuk, M.; Rohde, C.; Zhang, Y.; Ragozin, S.; Jeltsch, A. *Nucleic Acids Res.* **2007**, *35*, 100−12.
(6) Buryanov, Y.; Shevchuk, T. *Anal. Biochem.* **2005**, *338*, 1−11.
(7) (a) Stains, C. I.; Porter, J. R.; Ooi, A. T.; Segal, D. J.; Ghosh, I. *J. Am. Chem. Soc.* **2005**, *127*, 10782−3. (b) Ooi, A. T.; Stains, C. I.; Ghosh, I.; Segal, D. J. *Biochemistry* **2006**, *45*, 3620−5. (c) Stains, C. I.; Furman, J. L.; Segal, D. J.; Ghosh, I. *J. Am. Chem. Soc.* **2006**, *128*, 9761−5.
(8) Choe, W.; Chandrasegaran, S.; Ostermeier, M. *Biochem. Biophys. Res. Commun.* **2005**, *334*, 1233−40.
(9) Eberhardy, S. R.; Goncalves, J.; Coelho, S.; Segal, D. J.; Berkhout, B.; Barbas, C. F., III. *J. Virol.* **2006**, *80*, 2873−83.
(10) Beerli, R. R.; Segal, D. J.; Dreier, B.; Barbas, C. F., III. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14628−33.
(11) See Supporting Information figures S3 (a), S4 (b), S5 (c), and S6 (d).
(12) (a) Magliery, T. J.; Wilson, C. G.; Pan, W.; Mishler, D.; Ghosh, I.; Hamilton, A. D.; Regan, L. *J. Am. Chem. Soc.* **2005**, *127*, 146−57. (b) Galarneau, A.; Primeau, M.; Trudeau, L. E.; Michnick, S. W. *Nat. Biotechnol.* **2002**, *20*, 619−22. (c) Spotts, J. M.; Dolmetsch, R. E.; Greenberg, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15142−7.
(13) Pogribny, I. P.; Pogribna, M.; Christman, J. K.; James, S. J. *Cancer Res.* **2000**, *60*, 588−94.
(14) Mandell, J. G.; Barbas, C. F., III. *Nucleic Acids Res.* **2006**, *34*, W516−23. Blancafort, P.; Magnenat, L.; Barbas, C. F., III. *Nat. Biotechnol.* **2003**, *21*, 269−274.
(15) Lukinavicius, G.; Lapine, V.; Stasevskij, Z.; Dalhoff, C.; Weinhold, E.; Klimasauskas, S. *J. Am. Chem. Soc.* **2007**, *129*, 2758−9

JA0705588