DNA Methylation through a Locally Unpaired Intermediate

Daniel A. Erlanson, Lin Chen, and Gregory L. Verdine*

Department of Chemistry
Harvard University
Cambridge, Massachusetts 02138

Received August 24, 1993

Methylation of DNA serves essential roles in mammalian development and in bacterial resistance to viral pathogens.[1] In this process, a DNA (cytosine-5)-methyltransferase (DCMtase) mediates delivery of a methyl group from *S*-adenosyl-L-methionine to the 5-position of cytosine residues in DNA. DCMtases operate by conjugate addition of a cysteine thiolate to the 6-carbon (C6) of the substrate cytosine followed by transfer of a methyl group to C5. β-Elimination regenerates the free enzyme[2,3] (Figure 1). We have noted that the stereoelectronic attack trajectories for thiolate addition and methyl transfer cannot be accommodated in a canonical B-form duplex, suggesting that DCMtases cause transient helical disruption during the catalytic event.[3] Here we report evidence in favor of DCMtase-induced distortion of DNA and propose a structural model for the enzyme–DNA intermediate.

The DCMtase employed in these studies, M.*Hae*III, recognizes the dyad symmetric site,

$$5'\text{-d(G}\underline{\mathbf{G}}\mathit{C}\mathit{C}\text{)-}3'$$

$$3'\text{-d(}\mathit{C}\mathit{C}\underline{\mathbf{G}}\text{G)-}5'$$

and transfers a methyl group to the inner cytosine on each strand (italics). We reasoned that by cross-linking the complementary dG residues (bold, underlined), global strand separation of the sort widely employed by catalytic DNA-binding proteins could be severely restricted (Figure 2). We have previously reported controlled manipulation of duplex DNA structure and dynamics by site-specific incorporation of an alkane disulfide cross-link.[4,5] Disulfide cross-linking allows introduction of localized torsional stress into DNA,[5] and thus serves as a probe for protein-induced stress.

The cross-linked DNA was synthesized by the convertible nucleoside approach[6] as outlined in Figure 3. O^6-[2-(*p*-Nitrophenyl)ethyl]-2-fluoro-2′-deoxyinosine (NPE-FdI)[7] was incorporated into the appropriate position of an 18-mer by solid-phase synthesis.[8] The resin-bound oligonucleotide was treated with concentrated aqueous solutions of disulfide-containing diamines (n = 2, 3) or 2-(methylthio)ethylamine (MTE amine), resulting in replacement of the 2-fluoro substituent with the respective *N*-alkyl disulfide or *N*-MTE group.[9] The oligonucleotides were then treated with 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in formamide to remove the NPE protecting group.[10] The self-complementary 18-mers were reduced with dithiothreitol and then air-oxidized to furnish the cross-linked duplexes $\mathbf{C_2X}$ and $\mathbf{C_3X}$ in high yield[8] (Figure 3B).

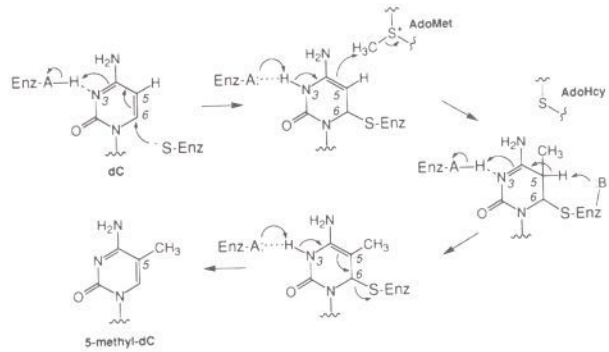

Figure 1. Proposed mechanism of DNA methylation catalyzed by DCMtases.

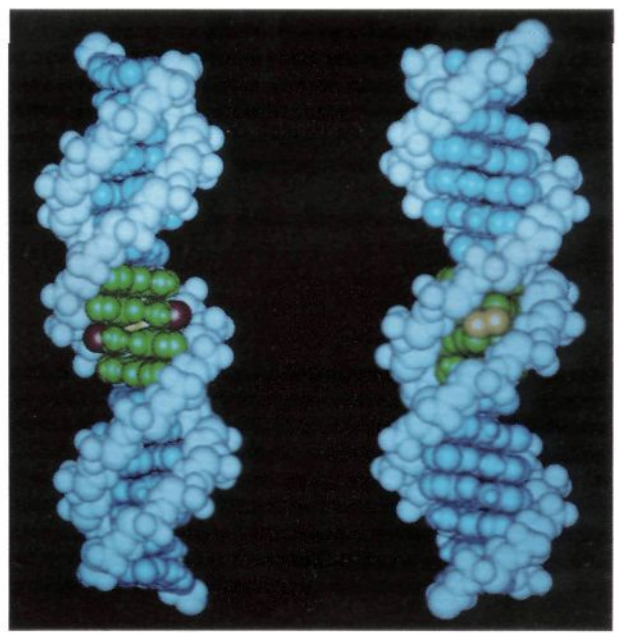

Figure 2. Model of the disulfide-cross-linked 18-mer, $\mathbf{C_2X}$. Left: View into the major groove of the M.*Hae*III recognition site (green). C5 and C6 of the substrate cytosine residues are colored red; in this view, C5 is in front of C6 and partially eclipses it. Right: View rotated by 180° about the helical axis, now looking into the minor groove. The sulfur atoms of the disulfide cross-link are colored yellow.

The presence of the disulfide cross-link in $\mathbf{C_2X}$ and $\mathbf{C_3X}$ greatly elevated the half-transition temperature (T_m) for thermal duplex denaturation, consistent with the known ability of such cross-links to stabilize DNA entropically.[4a,b,d] The T_m of $\mathbf{C_2X}$ was lower than that of $\mathbf{C_3X}$, however, indicating that the shorter cross-link induces more torsional stress in the duplex 18-mer. Comparison of the T_m values for the MTE-tethered and unmodified controls revealed that the tether alone has a negligible effect on duplex stability.[11]

To assess the effect of minor groove DNA cross-linking on recognition by M.*Hae*III in the major groove, we carried out

* To whom correspondence should be addressed.

(1) Li, E.; Bestor, T.; Jaenisch, R. *Cell* **1992**, *69*, 915–926. Bestor, T.; Coxon, A. *Curr. Biol.* **1993**, *3*, 384–386.

(2) Santi, D. V.; Garrett, C. E.; Barr, P. J. *Cell* **1983**, *33*, 9–10. Wu, J. C.; Santi, D. V. *J. Biol. Chem.* **1987**, *262*, 4778–4786.

(3) Chen, L.; MacMillan, A. M.; Chang, W.; Lane, W. S.; Verdine, G. L. *Biochemistry* **1991**, *30*, 11018–11025.

(4) (a) Ferentz, A. E.; Verdine, G. L. *J. Am. Chem. Soc.* **1991**, *113*, 4000–4002. (b) Ferentz, A. E.; Keating, T. A.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 9006–9014. (c) Ferentz, A. E.; Workiewicz-Kuzera, J.; Karplus, M.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 7569–7583. (d) Wolfe, S. A.; Ferentz, A. E.; Grantcharova, V.; Verdine, G. L. Submitted for publication.

(5) Wolfe, S. A.; Verdine, G. L. *J. Am. Chem. Soc.*, following paper in this issue.

(6) MacMillan, A. M.; Verdine, G. L. *Tetrahedron* **1991**, *47*, 2603–2616. MacMillan, A. M.; Verdine, G. L. *J. Org. Chem.* **1990**, *55*, 5931–5933.

(7) Lee, H.; Hinz, M.; Stezowski, J. J.; Harvey, R. G. *Tetrahedron Lett.* **1990**, *31*, 6773–6776.

(8) Further details are available in the supplementary material or can be obtained directly from the authors by FAX: (617) 495-8755.

(9) See also: Harris, C. M.; Zhou, L.; Strand, E. A.; Harris, T. M. *J. Am. Chem. Soc.* **1991**, *113*, 4328–4329.

(10) Unexpectedly, the disulfide tether of the N^2-alkyl-dG residue was found to be N-formylated on its dispensible half.

(11) Melting temperatures in 1 M NaCl (μM strand concentrations): unmodified control, 5′-d(ACGCATAGGCCTATGCGT)-3′, 71.3 °C; MTE-tethered control, 73.2 °C; $\mathbf{C_2X}$, 86.3 °C; and $\mathbf{C_3X}$, 90.1 °C; error limits were ±0.5 °C, for all except $\mathbf{C_3X}$ (±1.5 °C).[8]

0002-7863/93/1515-12583$04.00/0

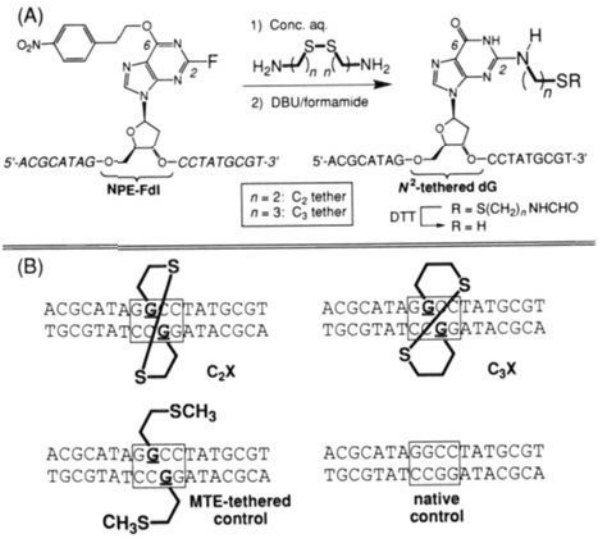

Figure 3. Synthesis of disulfide-cross-linked 18-mers. (A) Site-specific modification of dG residues. Italics denote the presence of protecting groups on the DNA. (B) Oligonucleotides studied in this work.

Table I. Relative Equilibrium Dissociation Constants (K_ds) for Interaction of M.*Hae*III[3] with Oligonucleotides[a]

oligonucleotide	K_d
unmodified 18-mer	43 ± 29
MTE-tethered control	85 ± 44
C_2X	1
C_3X	55 ± 20

[a] Values are normalized to the K_d of **C_2X**; a relative K_d of 1 corresponds to ~13 pM.

equilibrium binding assays under noncatalytic conditions.[8] Binding of M.*Hae*III to the cross-linked duplexes was compared with binding to unmodified and MTE-tethered 18-mer controls (Figure 3B). Wild-type M.*Hae*III bound the MTE-tethered control and **C_3X** with affinity similar to that of the unmodified oligonucleotide (Table I). In contrast, **C_2X** bound wild-type M.*Hae*III ~43-fold more tightly than unmodified DNA. These effects do not depend upon the ability of the enzyme to form a covalent complex with the DNA, since similar results were obtained with a mutant M.*Hae*III in which the active site Cys was replaced by Ala (data not shown).[12] In preliminary methyl-transfer experiments[8] conducted under conditions of substrate saturation, both **C_2X** and **C_3X** were found to be substrates for M.*Hae*III, being methylated at roughly 10% of the rate of the unmodified 18-mer control. Moreover, the cross-linked duplexes were methylated no more slowly than the MTE-tethered control, suggesting that the decrease in rate was not due to cross-linking *per se*.[13]

These experiments demonstrate that a minor groove disulfide cross-link can allosterically *activate* binding by a protein in the major groove.[14] Allosteric activation by the cross-link in **C_2X** most likely arises from its ability to induce torsional stress in DNA, which serves to prepay some of the thermodynamic price for helical distortion ordinarily brought about by the protein. The precise nature of this distortion is suggested by the mechanism

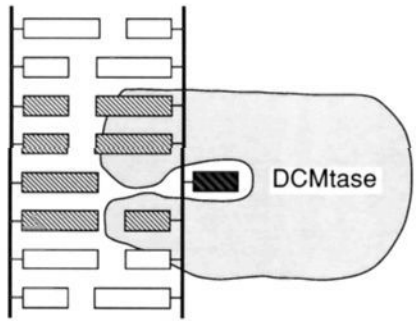

Figure 4. Model for helical distortion during DCMtase catalysis of DNA methylation. The bases are represented by boxes: light hatching, M.*Hae*III recognition site; dark hatching, substrate cytosine. The substrate cytosine is envisioned to undergo enzyme-induced extrusion to an extrahelical orientation.

presented in Figure 1. We have proposed earlier that protons are transferred to and from N3 during enzymatic processing.[12,15] However, because N3 is ordinarily involved in Watson–Crick base-pairing, this position can be made available to an enzymic acid only if the substrate G·C base pair is separated. Since the enzyme is still able to methylate cross-linked substrates, for which global strand separation is prevented, the strand separation event must be highly localized. To account for these findings, we propose that DCMtases cause transient extrusion of the substrate cytosine to an extrahelical position (Figure 4). This model not only allows for protonation at N3 but also affords the enzyme stereoelectronic access to C5 and C6. Such extrahelical bases have occasionally been observed in high-resolution structures of DNA containing various lesions.[16] Indeed, preliminary NMR studies on a DCMtase peptide–DNA complex derived from a trapped covalent intermediate (Figure 1, right-hand structure, H5 = F) have revealed that the substrate cytosine is in fact not Watson–Crick base-paired to its complementary guanine.[17,18]

Supplementary Material Available: General methods of spectroscopy, kinetic and thermodynamic analyses, gel electrophoresis, and HPLC purification of oligonucleotides, details of phosphoramidite, modified nucleoside, and oligonucleotide syntheses, and details of oligonucleotide characterization (23 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

Note Added in Proof: X-ray crystallographic studies have now confirmed the model proposed here: Klimaskausas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell*, in press.

(12) Chen, L.; MacMillan, A. M.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 5318–5319.

(13) Catalytic DNA-binding proteins are known to exert sequence selection in the transition state: Lesser, D. R.; Kurpiewski, M. R.; Jen-Jacobson, L. *Science* **1990**, *250*, 776–786.

(14) See also: Nosikov, V. V.; Braga, E. A.; Karlishev, A. V.; Zhuze, A. L.; Polyanovsky, O. L. *Nucleic Acids Res.* **1976**, *3*, 2293–2301. Dorn, A.; Affolter, M.; Müller, M.; Gehring, W. J.; Leupin, W. *EMBO J.* **1992**, *11*, 279–286. Fesen, M.; Pommer, Y. *J. Biol. Chem.* **1989**, *264*, 11354–11359. Oakley, M. G.; Mrksich, M.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10969–10975.

(15) Although protonation at O^2 could take place in canonical B-form DNA, the evidence presented here suggests the presence of helix disruption. Assuming that the substrate G·C base pair is broken, we favor protonation at N3 because of its higher pK_a (outside the DNA helix).

(16) Kao, Y. K.; Goljer, I.; Phan, T. A.; Bolton, P. H. *J. Biol. Chem.* **1993**, *268*, 17787–17793. Morden, K. M.; Maskos, K. *Biopolymers* **1993**, *33*, 27–36.

(17) Wolfe, S. A.; Chen, L.; Verdine, G. L. Unpublished results.

(18) We thank Ann Ferentz, Scot Wolfe, Andrew MacMillan, and Dusan Stanojevic for helpful discussions and Larry McLaughlin for use of his T_m instrument. This work was supported by the National Institutes of Health (GM 44853-03).