

Effect of base analog substitutions in the specific GATC site on binding and methylation of oligonucleotide duplexes by the bacteriophage T4 Dam DNA-[N⁶-adenine] methyltransferase

Ernst G. Malygin, Victor V. Zinoviev, Nicolay A. Petrov, Alexey A. Evdokimov, Linda Jen-Jacobson¹, Valeri G. Kossykh^{2,+} and Stanley Hattman^{2,*}

Institute of Molecular Biology, State Research Center of Virology and Biotechnology 'Vector', Novosibirsk 633159, Russia, ¹Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA and ²Department of Biology, University of Rochester, Rochester, NY 14627-0211, USA

Received September 23, 1998; Revised and Accepted November 28, 1998

ABSTRACT

The interaction of the phage T4 Dam DNA-[N⁶-adenine] methyltransferase with 24mer synthetic oligonucleotide duplexes having different purine base substitutions in the palindromic recognition sequence, GATC, was investigated by means of gel shift and methyl transfer assays. The substitutions were introduced in either the upper or lower strand: guanine by 7-deazaguanine (G→D) or 2-aminopurine (G→N) and target adenine by purine (A→P) or 2-aminopurine (A→N). The effects of each base modification on binding/methylation were approximately equivalent for both strands. G→D and G→N substitutions resulted in a sharp decrease in binary complex formation. This suggests that T4 Dam makes hydrogen bonds with either the N7- or O6-keto groups (or both) in forming the complex. In contrast, A→P and A→N substitutions were much more tolerant for complex formation. This confirms our earlier observations that the presence of intact 5'-G:C base pairs at both ends of the methylation site is critical, but that base substitutions within the central A:T base pairs show less inhibition of complex formation. Addition of T4 Dam to a complete substrate mixture resulted in a burst of [³H]methylated product. In all cases the substrate dependencies of bursts and methylation rates were proportional to each other. For the perfect 24mer $k_{\text{cat}} = 0.014/\text{s}$ and $K_{\text{m}} = 7.7 \text{ nM}$ was obtained. In contrast to binary complex formation the two guanine substitutions exerted relatively minor effects on catalytic turnover (the k_{cat} was reduced at most 2.5-fold), while the two adenine substitutions showed stronger effects (5- to 15-fold reduction in k_{cat}). The effects of base analog substitutions on $K_{\text{m}}(\text{DNA})$ were more variable: A→P (decreased); A→N and G→D (unchanged); G→N (increased).

INTRODUCTION

Three kinds of DNA methyltransferases (MTases) are known to exist in prokaryotes: C5-cytosine (Cyt), N⁴-Cyt and N⁶-adenine (Ade or A) MTases (1). Type II DNA MTases generally recognize short palindromic sequences and catalyze methyl transfer from donor *S*-adenosyl-L-methionine (AdoMet) to the N6-amino group of an Ade or the C5 atom or N4-amino nitrogen of a Cyt in the target sequence. In restriction–modification systems, two proteins, the cognate MTase and restriction endonuclease (ENase), are able to recognize a common specific nucleotide sequence, but catalyze very different reactions (2). Among the MTases, the most progress in the study of their mechanism of action has been attained for the C5-Cyt MTases. Not only has the chemical mechanism of catalysis been elucidated (3), but three-dimensional structures of MTase complexes with their substrates have been solved by X-ray crystallography (3–7). A most surprising and exciting result is that the Cyt residue to be methylated is flipped out of the DNA helix (6). Among the N⁶-Ade and N⁴-Cyt MTases, structures have been reported only for the *Taq*I and *Pvu*II MTases, respectively (8,9), but co-crystallization with DNA was not successful. Thus, the possible flipping of these target bases has not been shown directly, but alternative methodology (10–13) or modeling (14) suggests that it occurs with the N⁶-Ade MTases. In the absence of definitive structures from X-ray crystallography, genetic/biochemical studies (15,16,20–24) of these MTases are important to carry out. In fact, they are warranted even when crystal structures are known, as crystallographic images just indicate proximity relationships, they do not give direct insight into the individual contributions of each DNA nucleoside or backbone contact with the polypeptide side chain towards either target recognition or the catalytic process. One of the approaches used is to employ defined synthetic oligodeoxynucleotide duplexes containing a base analog that alters or removes a single contact (18–20); in some cases, the free energy change can be attributed to loss of a single

*To whom correspondence should be addressed. Tel: +1 716 275 8046; Fax: +1 716 275 2070; Email: moddna@uhura.cc.rochester.edu

⁺Present address: Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, TX 77555, USA

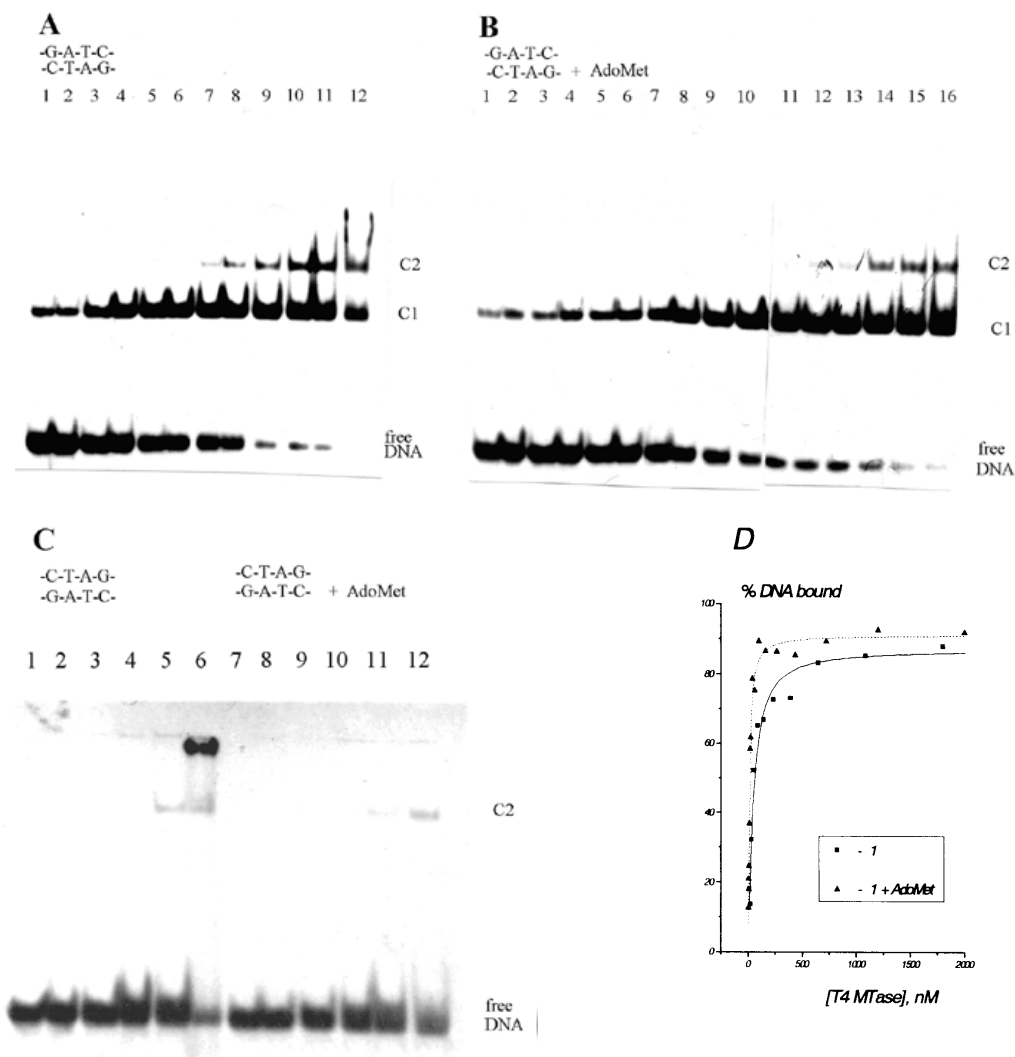


Figure 1. Binding of T4 Dam to 24mer specific duplex 1 containing the canonical palindromic methylation sequence, GATC. In this and subsequent figures, the binding mixture contained oligonucleotide duplex at 5 nM. Binding in the absence (A) or presence (B) of AdoMet (5 μ M). (A) Lanes 1–12 had concentrations of 11, 18, 30, 50, 84, 140, 233, 389, 648, 1080 and 3000 nM, respectively. (B) Lanes 1–16 had concentrations of 2, 3, 4, 6, 8, 12, 20, 34, 56, 93, 156, 260, 432, 720, 1200 and 2000 nM T4 Dam, respectively. (C) Binding of T4 Dam to 24mer non-specific duplex 0 containing the palindrome, CTAG, in place of GATC. Lanes 1 and 7, 2 and 8, 3 and 9, 4 and 10, 5 and 11 and 6 and 12 had concentrations of 194, 324, 540, 900, 1500 and 2500 nM T4 Dam, respectively. (D) Binding of T4 Dam with native specific 24mer duplex 1 in the absence and presence of AdoMet, as a function of T4 Dam concentration. Binding curves were fitted to the Hill equation in the form $P = P_{\max}/[1 + (K_d/[E])^n]$, as proposed by Kourganov (30). The K_d value here is equal to [E] at half of P_{\max} binding. The fitting was achieved using a computer program for non-linear regression analysis (developed by A. N. Naumochkin, Novosibirsk) with minimization of X^2 determining the fit.

contact. In this regard, studies using duplexes with 2-aminopurine substituted in the DNA methylation site have provided strong support for the notion that the target Ade residue also flips out of the helix (10–12).

Studying complex formation between a DNA MTase and its substrates, as well as the kinetics of reaction product formation, also may be of value towards elucidating the contribution of individual functional groups in DNA recognition by the enzyme. The Dam DNA-[N⁶-Ade] MTase encoded by phage T4 catalyzes methyl group transfer from AdoMet to the N6 position of Ade in the palindromic sequence, GATC. Earlier, we investigated the capability of T4 Dam to form complexes with synthetic duplex oligonucleotide substrates by means of gel filtration, ultracentrifugation and gel shift electrophoresis assays (23,25). For the gel shift assay, we applied the 'dissected duplex' principle (26,27), using variant defective duplexes lacking a phosphate or nucleo-

side(s) within the GATC target site. Studies with such defective duplexes showed that the formation of a stable complex with T4 Dam did not require that both strands be contiguous nor completely complementary. In fact, having only one half of the recognition site intact was sufficient for stable complex formation provided that both 5'-G:C base pairs were present at both ends of GATC. In addition, steady-state kinetic parameters of T4 Dam methylation were studied (24,28) using both unmethylated, unglucosylated T4 *gt⁻ dam⁻* DNA, which is a natural substrate for the enzyme, and synthetic oligonucleotide duplexes containing some defect in the target site, e.g. absence of an internucleotide phosphate or presence of an abasic site or having only a partially double-stranded recognition site. We showed that although having only half of the recognition site intact was sufficient for stable complex formation, the T4 Dam catalytic turnover process had a strict requirement for uninterrupted GAT sequences on both strands.

In this paper we describe studies on the effect of base modifications in GATC on complex formation and the kinetics of T4 Dam methylation of synthetic 24mer duplexes.

MATERIALS AND METHODS

Enzymes and chemicals

All oligonucleotide duplexes studied had the same general structure:



except for individual Ade or guanine (Gua or G) substitutions (Table 1) in the specific GATC target site (indicated in bold), while the flanking sequences were the same in all the duplexes. The substitutions were introduced in either the upper or lower strand: G→2-aminopurine (G→N), G→7-deazaguanine (G→D), A→2-aminopurine (A→N) or A→purine (A→P). Oligonucleotide concentrations were determined spectrophotometrically from the known molar extinction coefficients of individual oligonucleotides and their known sequences. Oligonucleotide duplexes were radiolabeled using [γ - 32 P]ATP and T4 polynucleotide kinase and then purified by preparative gel electrophoresis. T4 polynucleotide kinase was from SibEnzyme (Novosibirsk), [γ - 32 P]ATP from Biosan (Novosibirsk) and unlabeled AdoMet (Sigma) was purified further by chromatography on a C₁₈ reversed-phase column as described previously (15). [3 H-CH₃]AdoMet (15 Ci/mmol, 1 mCi/ml) was from Amersham. T4 Dam MTase was purified to homogeneity as previously described (16,28). Protein concentrations were determined by the Bradford method (29).

Binding experiments

Binding reactions and gel shift electrophoresis assays were carried out as described previously (23). Since the reaction tubes were kept on ice until the time of gel electrophoresis, methylation of duplexes in the presence of AdoMet was essentially eliminated during the incubation period.

DNA MTase assay

Methyl transfer assays were carried out as described (28) and were similar to those previously reported (24). Reaction mixtures (150 μ l) contained 100 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM DTT, 5% glycerol and 0.2 mg/ml BSA. [3 H-CH₃]AdoMet concentration was 2 μ M, if not noted otherwise. For determination of steady-state kinetic parameters, 2.25 or 4.5 nM of T4 Dam and varying concentrations of substrate were used. Reactions were initiated by addition of an aliquot of prewarmed T4 Dam solution to a preincubated mixture of [3 H-CH₃]AdoMet and substrate DNA at 25°C. Aliquots (20 μ l) were withdrawn at intervals and spotted on DE81 anion-exchange filter papers (2.0 cm; Whatman). The molar concentrations of [3 H]CH₃ groups incorporated into DNA were quantitated as described (22). Under the maximum methylation conditions, the calculated concentrations of [3 H]CH₃ groups incorporated into DNA coincided with the reaction mixture concentrations of methylatable Ade residues for various substrates, confirming the validity of the method. Steady-state parameters were obtained by fitting the experimental data using a computer program for non-linear regression analysis (developed by A. N. Naumochkin, Novosibirsk).

Table 1. Apparent K_d values for T4 Dam binding to duplexes with a native or modified recognition site^a

Duplex no.	Recognition site ^b	Group deleted	In absence of AdoMet			In presence of AdoMet		
			K_d (nM)	n	$\Delta\Delta G^\circ_{\text{bind}}$ (kcal/mol) ^c	K_d (nM)	n	$\Delta\Delta G^\circ_{\text{bind}}$ (kcal/mol) ^c
0	-C-T-A-G- -G-A-T-C-		>1500			>2500		
1	-G-A-T-C- -C-T-A-G-		54 (7)	1.0 (0.1)		18 (3)	1.0 (0.1)	
2	-G-P-T-C- -C-T-A-G-	N6 amino	105 (4)	3.0 (0.3)	0.4	47 (2)	5.1 (1)	0.5
3	-G-A-T-C- -C-T-P-G-	N6 amino	130 (2)	2.9 (0.1)	0.5	46 (2)	3.4 (0.4)	0.5
4	-G-N-T-C- -C-T-A-G-	N6 amino	256 (8)	2.9 (0.3)	0.9	95 (3)	3.9 (0.5)	0.9
5	-G-A-T-C- -C-T-N-G-	N6 amino	308 (10)	2.6 (0.2)	1.0	114 (3)	5.3 (0.8)	1.1
6	-N-A-T-C- -C-T-A-G-	O6 keto	1700 (85)	1.2 (0.1)	2.0	1480 (80)	0.9 (0.05)	2.6
7	-G-A-T-C- -C-T-A-N-	O6 keto	1400 (37)	1.4 (0.1)	1.9	1750 (95)	1.3 (0.1)	2.6
8	-D-A-T-C- -C-T-A-G-	N7 imidazole	1450 (150)	1.2 (0.2)	1.9	935 (90)	1.1 (0.1)	2.3
9	-G-A-T-C- -C-T-A-D-	N7 imidazole	1570 (90)	0.9 (0.1)	1.9	990 (110)	0.8 (0.1)	2.3

^aThe K_d values were obtained from binding curve fitting for the Hill equation (30), as described in the text; n is the calculated Hill coefficient. Standard deviation values are in parentheses.

^bP, purine; D, 7-deazaguanine; N, 2-aminopurine.

^c $\Delta\Delta G^\circ_{\text{bind}}$ was calculated for $T = 20^\circ\text{C}$, the temperature of gel electrophoresis.

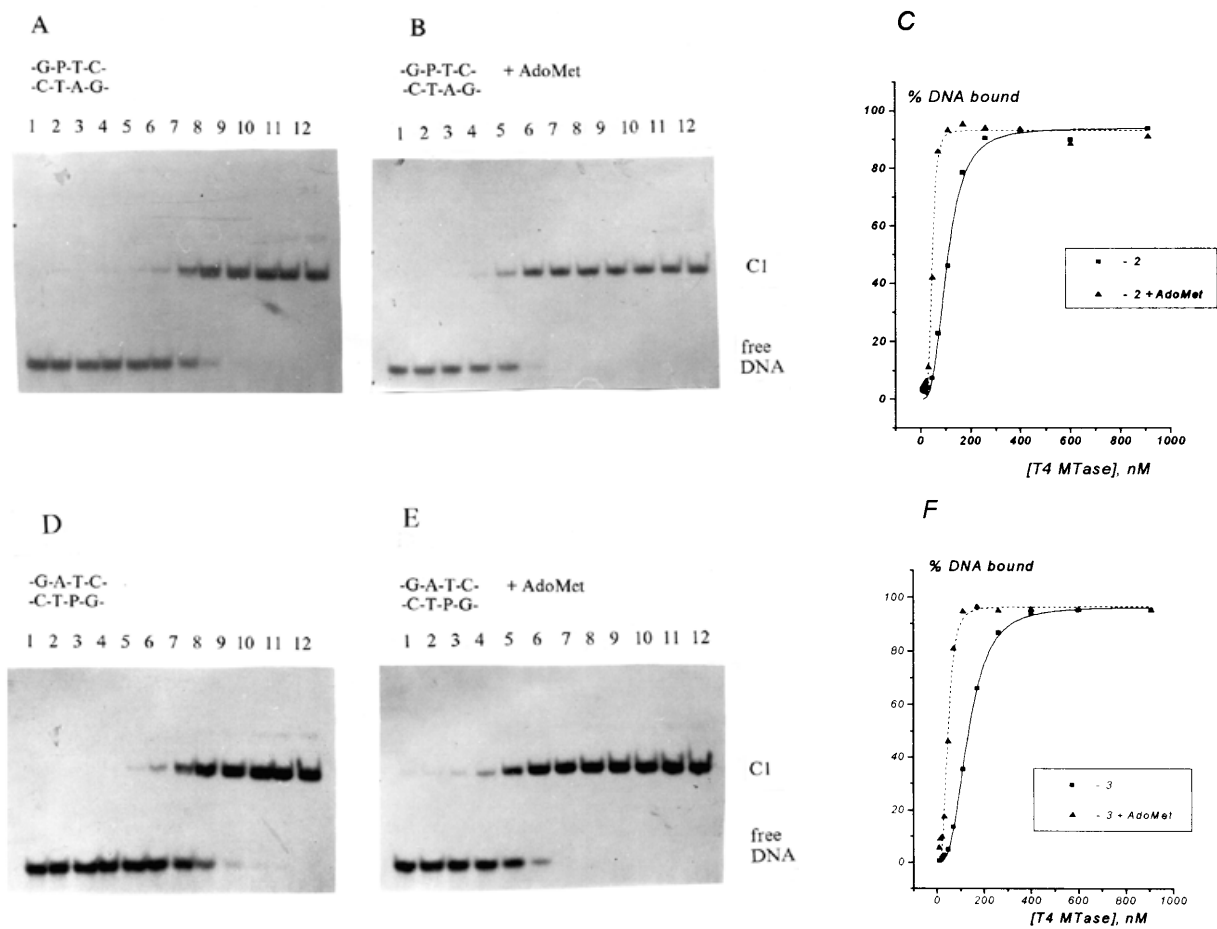


Figure 2. Binding of T4 Dam with duplexes 2 and 3 containing the Ade residue substituted by purine (P) in the absence (A and D) and in the presence (B and E) of AdoMet (5 μ M). Lanes 1–12 had concentrations of 9, 13, 20, 30, 46, 70, 110, 170, 260, 400, 600 and 910 nM T4 Dam, respectively. Binding curves (C and F) in this and subsequent figures were fitted as described in the legend to Figure 1D.

RESULTS

Binding of T4 Dam to a 24mer duplex with an unmodified recognition site

Figure 1A, B and D shows the results of the gel shift assay for complexes formed between T4 Dam and the 24mer duplex 1, having an unmodified GATC/GATC site; it should be noted here that at 4°C, the temperature of the binding reaction mixture, methylation of the duplexes does not occur (22). Duplex 1 complexed well with T4 Dam in the absence or presence of 5 μ M AdoMet, a concentration sufficient to saturate T4 Dam (24). It can be seen that in the range of low enzyme concentrations a complex (designated C1) was formed, which contained T4 Dam:oligo duplex at a stoichiometry of 1:1. At elevated concentrations of T4 Dam, a larger complex (designated C2) was formed, which we estimated to contain no less than two T4 Dam per oligo duplex. To calculate the fraction of duplex 1 bound by enzyme, we summed the radioactivities of all complexes observed in the lane.

The presence of AdoMet in the binding mixture enhanced complex formation ~3-fold (Table 1). The calculated apparent K_D values were comparable with those obtained earlier using a native

20mer duplex with different flanking nucleotide sequences (23), indicating that different sequences flanking GATC did not influence complex formation. The titration curves in the absence and presence of AdoMet (Fig. 1D) were well fitted to a simple hyperbola corresponding to a Langmuir isotherm, $P = P_{\max}/(1 + K_D/[E])$. However, we also used the more general Hill equation, $P = P_{\max}/[1 + (K_D/[E])^n]$ (30); this was necessary for some modified substrates, as will be seen below.

In a control experiment, complex formation between T4 Dam and the non-specific duplex 0, containing the inverse palindrome, CTAG/CTAG (Table 1), was shown to be negligible up to 1.5 μ M T4 Dam (Fig. 1C). At higher concentrations, aggregates of T4 Dam formed non-specific complexes (Fig. 1C, lane 6), although this was reduced in the presence of AdoMet (Fig. 1C, lanes 7–12).

Binding of T4 Dam to Ade-substituted duplexes

Two different Ade substitutions were introduced into the upper or lower strand GATC, namely purine (A→P) or 2-aminopurine (A→N) (Table 1). Both substitutions delete a potential contact point (A-N6) with protein and one of the two major groove Watson–Crick hydrogen bonds; the A→N substitution, however,

introduces a hydrogen bond in the minor groove (31). Figures 2 and 3 show the results of gel shift assays with A→P and A→N substituted duplexes, respectively. It can be seen that these substitutions did not greatly reduce binding capability of the modified substrates, i.e. the apparent K_d was increased from 2- to 6-fold, with the A→N substitution being relatively poorer in comparison with A→P (Table 1). As with the canonical duplex 1, presence of AdoMet enhanced binding 2- to 3-fold for both substitutions.

Surprisingly, the shapes of the binding curves for all the Ade-substituted substrates (Figs 2C and F and 3C and F) were poorly described by a simple hyperbola, as compared with that for duplex 1 (Fig. 1D). Instead, they had a distinct sigmoidal character and they were well fitted to the Hill equation, $P = P_{\max} / [1 + (K_d/[E])^n]$, where n is the Hill coefficient and K_d is equal to $[E]$ at $P = P_{\max}/2$. The n values for duplexes 2–5 were from 2.5 to 3 in the absence of AdoMet and from 3 to 5 in its presence, suggesting some kind of cooperativity in binding of T4 Dam. However, this is not consistent with the presence of a complex containing only a T4 Dam monomer over a range of concentrations sufficient to have seen cooperativity; also, there were no multimeric complexes observed for those interactions characterized by n values >1 . Thus, we do not yet understand the significance of the Hill coefficients being greater than unity, so this question requires further investigation. In this regard, a sigmoidal dependence on enzyme concentration was observed for binding by the murine DNA-[C5-Cyt] MTase with oligonucleotide duplexes (32).

Binding of T4 Dam to Gua-substituted duplexes

We also investigated the influence of two Gua substitutions on

complex formation with T4 Dam, namely 2-aminopurine (G→N) and 7-deazaG (G→D). The G→N substitution deletes an O6-keto group, which normally participates in hydrogen bonding; the G→D substitution replaces the imidazole ring N7 with a carbon, but that does not perturb hydrogen bonding. Figures 4 and 5 show the results of gel shift assays for complexes of T4 Dam with these modified duplexes in the presence and absence of AdoMet. It can be seen that all these substitutions strongly decreased complex formation, ~40- to 50-fold, compared with native duplex 1 (Table 1). In the absence of AdoMet the K_d values for duplexes 6–9 were comparable with one another (1400–1700 nM); presence of AdoMet in the reaction mixture had little or no effect on binding. Since binding was, in all cases, far from saturation, we used the Hill equation for curve fitting, with a physically justified limitation of $P_{\max} \leq 100\%$. It can be seen that the Hill coefficients were all near to unity for these duplexes.

Steady-state kinetics of methylation

The effect of the above Ade and Gua substitutions on T4 Dam methylation of these duplexes was also investigated. For all substrates, the time dependence of methyl group transfer was linear under steady-state conditions, with a distinct burst of product formation at the initial time of reaction (Fig. 6). This is in agreement with earlier results using a different set of defective oligonucleotide duplexes (24). The initial rates of methylation (M) were linear up to at least 40 nM T4 Dam (Fig. 7). In order to ensure linearity and real measures of the initial rates, assays were carried out such that product formation was $<20\%$ of the initial target A residues. All results fit well to the Michaelis equation and steady-state kinetic parameters for AdoMet and DNA were calculated in the standard fashion.

Table 2. Steady-state kinetic parameters and bursts of T4 Dam determined for various 24mer duplexes^a

Duplex no.	Recognition site ^b	K_m (nM)	k_{cat} (per s $\times 10^{-3}$)	k_{cat}/K_m (per M/s $\times 10^{-6}$)	k_{cat}/K_m (relative)
1	-G-A-T-C-	7.7	14	1.80	(1.0)
	-C-T-A-G-	(2.1)	(3.0)		
2	-G-P-T-C-	2.5	0.99	0.40	0.22
	-C-T-A-G-	(1.2)	(0.15)		
3	-G-A-T-C-	4.3	0.89	0.21	0.11
	-C-T-P-G-	(1.0)	(0.13)		
4	-G-N-T-C-	8.7	2.6	0.30	0.16
	-C-T-A-G-	(3.0)	(0.5)		
5	-G-A-T-C-	12.9	2.9	0.22	0.12
	-C-T-N-G-	(3.8)	(1.0)		
6	-N-A-T-C-	108	5.7	0.05	0.03
	-C-T-A-G-	(50)	(3.7)		
7	-G-A-T-C-	59	5.6	0.09	0.05
	-C-T-A-N-	(15.3)	(1.3)		
8	-D-A-T-C-	12.3	13	1.10	0.58
	-C-T-A-G-	(5.7)	(3.0)		
9	-G-A-T-C-	10.9	14	1.30	0.71
	-C-T-A-D-	(4.4)	(2.0)		

^aStandard deviations in parentheses were determined from at least two independent experiments.

^bP, purine; D, 7-deazaguanine; N, 2-aminopurine.

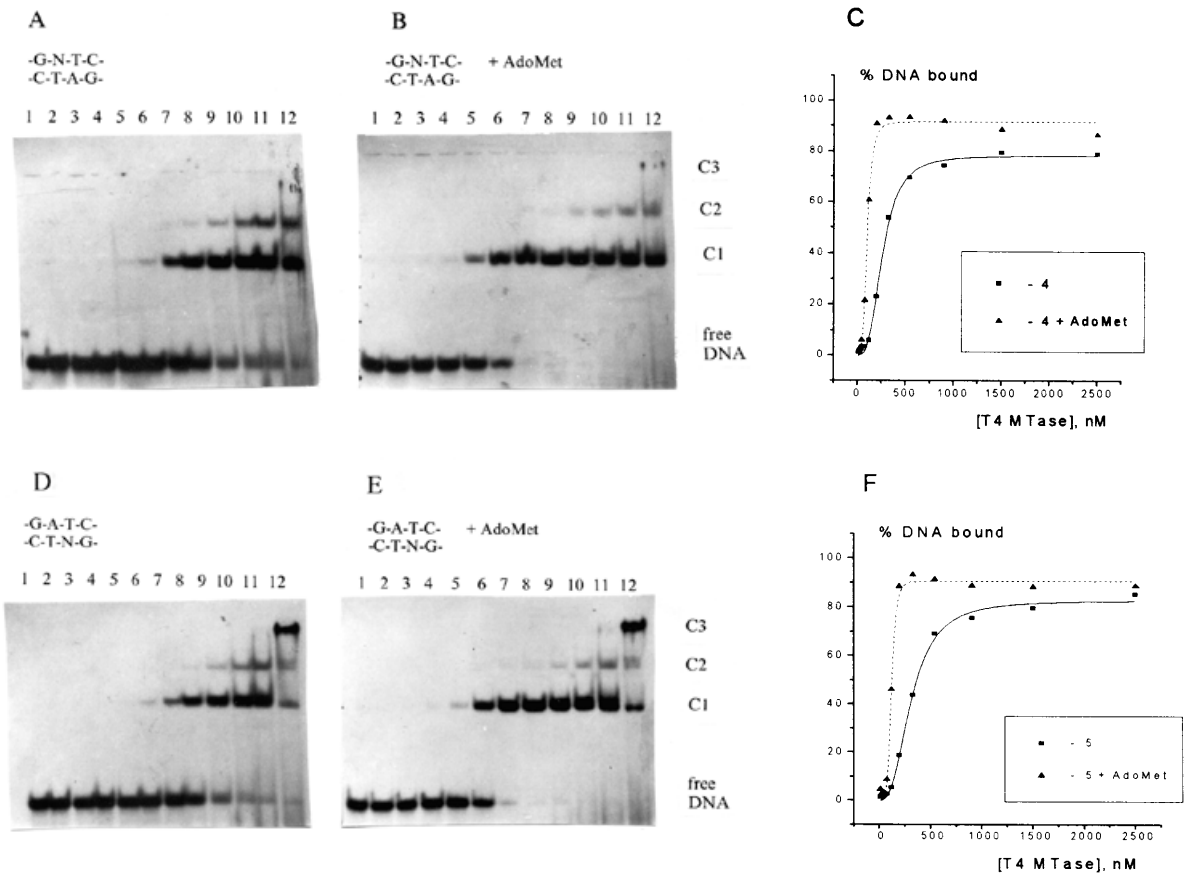


Figure 3. Binding of T4 Dam with the duplexes 4 and 5 containing an Ade residue substituted by 2-aminopurine (N) in the absence (A and D) or presence (B and E) of AdoMet (5 μ M). Lanes 1–12 had T4 Dam concentrations of 9, 15, 25, 42, 70, 117, 194, 324, 540, 900, 1500 and 2500 nM, respectively.

In a preliminary experiment using unmodified duplex 1 (Fig. 8A), the $K_m(\text{AdoMet})$ was calculated to be 0.33 μ M, therefore, all subsequent kinetic experiments were performed in the presence of 2 μ M AdoMet. For comparison, $K_m(\text{AdoMet})$ values of 0.1 μ M (28) and 0.49 μ M (24) were previously determined using polymeric T4 *gt^r dam^r* DNA and 20mer oligonucleotide duplex, respectively. It would seem that lengthening of substrate DNA might be accompanied by some lowering in $K_m(\text{AdoMet})$, but this was not investigated further. Initial reaction rates, as a function of DNA concentration, are presented in Figure 8B and steady-state kinetic parameters for various 24mer duplexes are summarized in Table 2. Parameters for T4 Dam methylation of duplex 1 were $K_m(\text{DNA}) = 7.7$ nM and $k_{\text{cat}} = 0.014/\text{s}$; these values coincided quite well with those reported for a perfect 20mer, $K_m(\text{DNA}) = 6.3$ nM and $k_{\text{cat}} = 0.015/\text{s}$ (24).

The salient findings are summarized as follows. (i) Steady-state parameters were independent of which strand contained a base modification. (ii) Values of $K_m(\text{DNA})$ increased as follows: A \rightarrow P < unmodified duplex < A \rightarrow N = G \rightarrow D < G \rightarrow N. Note that K_m values for the A \rightarrow P substitution were lower than for the unmodified DNA. Values of k_{cat} decreased as follows: G(unmodified) = G \rightarrow D, G \rightarrow N, A \rightarrow N, A \rightarrow P. (iii) The G \rightarrow N substitution exhibited the most pronounced decline in specificity coefficient ($k_{\text{cat}}/K_m < 0.1 \times 10^6/\text{M/s}$). (iv) The A \rightarrow P substituted duplexes exhibited a lowering in $K_m(\text{DNA})$, but a 15-fold decrease in k_{cat} compared with duplex 1. It should be noted that there is only one target Ade

residue in duplexes 2–5, which contain the A \rightarrow P and A \rightarrow N substitutions. Since unmethylated and hemimethylated (unmodified) 20mers showed similar $K_m(\text{DNA})$ values and only a 1.5-fold difference in k_{cat} (24), then having only one strand to methylate *per se* in duplexes 2–5 cannot account for the changes in steady-state methylation parameters observed here.

DISCUSSION

Effect of base analog substitutions on T4 Dam binding and complex formation

We showed earlier that substrate modification may have the opposite effects on stable complex formation versus kinetics of T4 Dam methylation (23,24). In this regard, there are only a limited number of published reports concerning the binding of MTases to specific oligonucleotide duplexes. For example, Klimasauskas and Roberts changed one of the target cytosine residues in the palindromic M \cdot *HhaI* site, GCGC, to a variety of other bases (33). They found that the MTase bound more tightly to duplexes containing a mismatch. This is analogous to our observations that certain mismatches, or absence of a base, in the central A-T enhanced binding of T4 Dam (23). In the present study, however, we did not observe any enhancement of binding when Ade was substituted by either purine or 2-aminopurine. Szczelkun *et al.* (34) observed complete inhibition of both

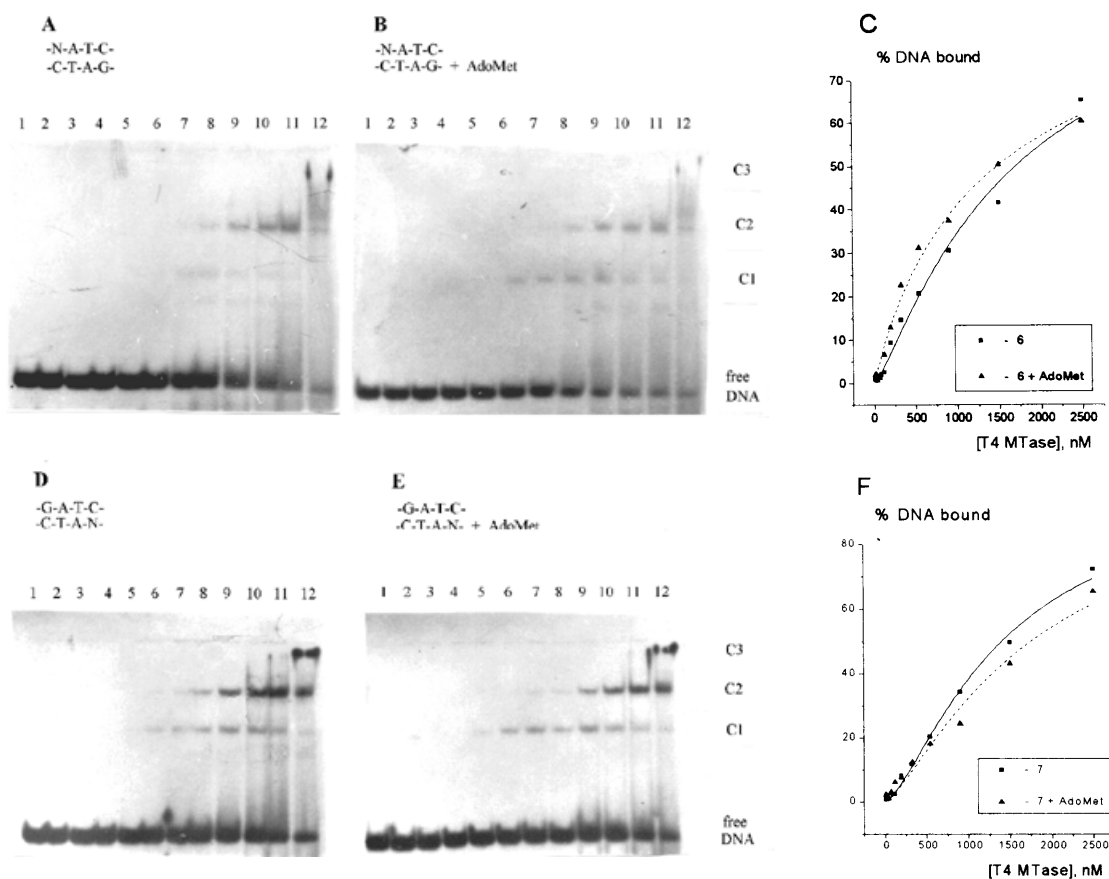


Figure 4. Binding of T4 Dam with the duplexes 6 and 7 containing the Gua residue substituted by 2-aminopurine (N) in the absence (**A** and **D**) or presence (**B** and **E**) of AdoMet (5 μ M). Lanes 1–12 had T4 Dam concentrations of 9, 15, 25, 42, 70, 117, 194, 324, 540, 900, 1500 and 2500 nM, respectively.

binding and catalysis when Gua was replaced by 7-deazaG (G \rightarrow D) in the *EcoRV* MTase palindromic recognition site, GATATC. Thus, the N7 position of both Gua residues is important in target recognition by *EcoRV*. We also observed that Gua substitutions gave 40- to 50-fold poorer binding of T4 Dam (duplexes 6–9, Table 1), which is consistent with previous evidence for the importance of this residue in T4 Dam binding and catalysis (23,24). It should be noted that in the case of T4 Dam, not only the N7 position, but also the O6-keto group seems important for specific enzyme–substrate interaction. This suggests that T4 Dam makes hydrogen bonds with both these substituents in forming the complex.

Energetic penalties for modified sites

For the present, there are no three-dimensional structures solved for T4 Dam complexed with its substrates. Nevertheless, we can analyze our data following the logic proposed by Lesser *et al.* (18), who studied the cost of modified duplex recognition by the *EcoRI* endonuclease, for which X-ray crystallographic structures are known. They calculated the difference in standard binding free energy between the canonical site and each modified site, $\Delta\Delta G^{\circ}_{\text{bind}} = -RT\ln(K_d/K_{d \text{ mod}})$. These values may include a variety of changes in interactions, including protein–base ($\Delta\Delta G^{\circ}_{\text{base}}$), protein–phosphate ($\Delta\Delta G^{\circ}_{\text{phos}}$) and a more general ‘reorganization’ term ($\Delta\Delta G^{\circ}_{\text{reorg}}$), for changes in conformational

contribution to the interactions. From a number of lines of evidence and in comparison with the X-ray structure of the *EcoRI*–DNA complex, Lesser *et al.* (18) established that the energetic contribution of a single protein–base hydrogen bond to binding free energy was \sim 1.4 kcal/mol.

Following this analysis, we evaluated the energetic consequences of different purine modifications to T4 Dam interaction. Table 1 shows the $\Delta\Delta G^{\circ}_{\text{bind}}$ values for single substitutions in the GATC site. We found small penalties in $\Delta\Delta G^{\circ}_{\text{bind}}$ values of from \sim +0.4 to +1.1 kcal/mol for an A \rightarrow P or A \rightarrow N substitution, each of which removes a potential T4 Dam contact point (A–N6) (Table 1, duplexes 2–5). These small penalties are likely the net of an unfavorable contribution of deletion of the hydrogen bond to A–N6, estimated at +1.4 kcal/mol (18), and a compensating favorable effect on the reorganization energy ($\Delta\Delta G^{\circ}_{\text{reorg}}$). Removal of a major groove Watson–Crick constraint may facilitate achieving the precise local DNA structural features required for maximum protein–DNA complementarity at the interface. The addition of a 2-amino group for the A \rightarrow N substitution introduces a Watson–Crick constraint in the minor groove; this might account for the slightly greater penalty (\sim +1.0 kcal/mol) observed than that for the A \rightarrow P substitution (\sim +0.5 kcal/mol). For Ade substitutions, the penalties were the same in the presence or absence of AdoMet, thus enhancement of complex formation by AdoMet is not related to any direct influence on T4 Dam interaction with the Ade residue.

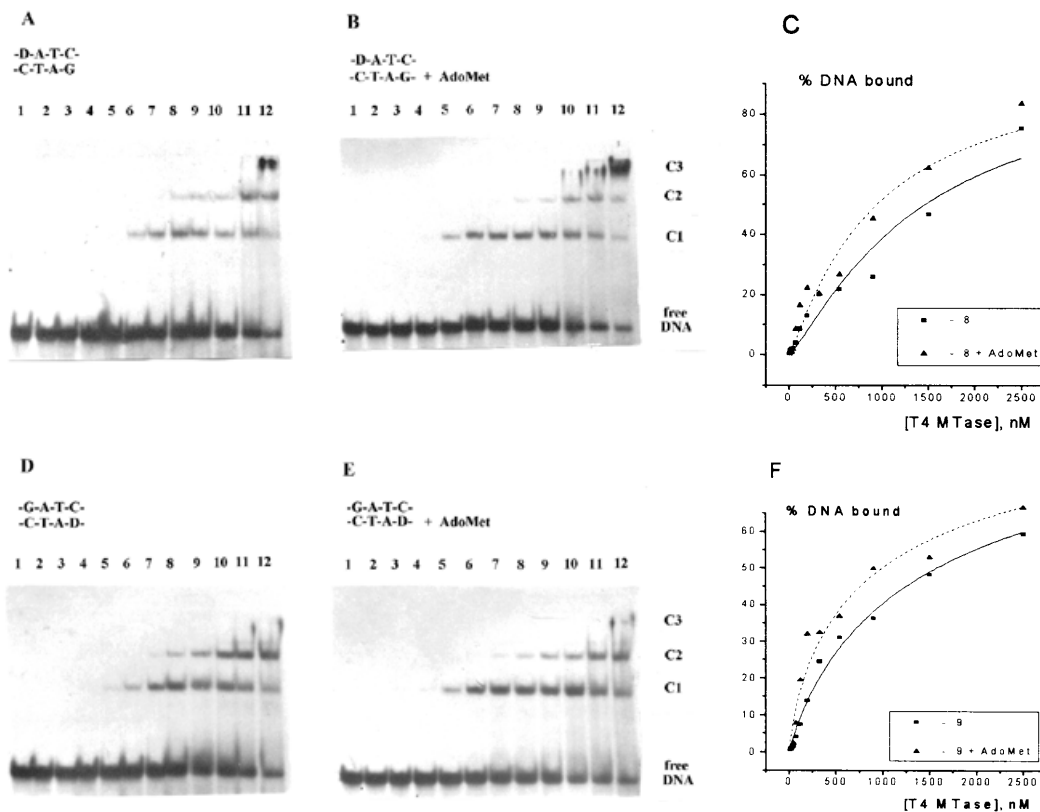


Figure 5. Binding of T4 Dam with the duplexes 8 and 9 containing the Gua residue substituted by 7-deazaG (D) in the absence (A and D) or presence (B and E) of AdoMet (5 μ M). Lanes 1–12 had T4 Dam concentrations of 9, 15, 25, 42, 70, 117, 194, 324, 540, 900, 1500 and 2500 nM, respectively.

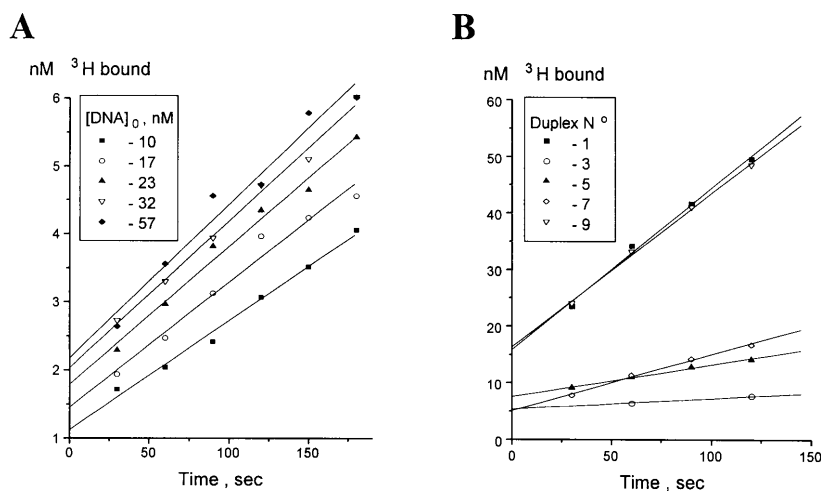


Figure 6. Time dependence of methylation of oligonucleotide duplexes by T4 Dam: (A) perfect duplex 1 at different concentrations (2.25 nM T4 Dam); (B) comparison of methylation of various substrates (Table 1) at 18 nM T4 Dam and saturating DNA concentrations (duplexes 1, 3 and 5, 200 nM; duplex 7, 500 nM; duplex 9, 300 nM).

By contrast, we found significantly greater binding free energy penalties ($\Delta\Delta G^\circ_{\text{bind}}$ of $\sim +1.9$ to $+2.6$ kcal/mol) for 2-aminopurine and 7-deazaG substitutions of Gua in the GATC site (Table 1, duplexes 6–9). These larger penalties could result from changes in protein–phosphate contacts ($\Delta\Delta G^\circ_{\text{phos}}$) and/or conformational

factors ($\Delta\Delta G^\circ_{\text{reorg}}$) in addition to the removal of the potential contact points O6 or N7 ($\Delta\Delta G^\circ_{\text{base}}$). The $\Delta\Delta G^\circ_{\text{phos}}$ and $\Delta\Delta G^\circ_{\text{reorg}}$ are the same for any substitution at a given base position, a 1.3 kcal/mol incremental contribution to the binding energy (17). At present, we have not obtained other lines of

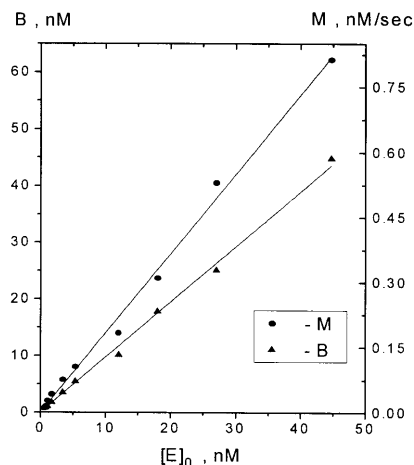


Figure 7. Initial reaction rates (M) and bursts (B) of perfect duplex 1 (210 nM) methylation by T4 Dam as a function of enzyme concentration.

evidence, such as ethylation interference footprints, which would help us to assess whether, in addition to protein–base contacts, protein–phosphate contacts have also been perturbed, as was found for 7-deazaG and 2-aminopurine substitutions of G in the *EcoRI* recognition site, GAATTC (17), and for 2-aminopurine substitutions of either G in the *BamHI* recognition site, GGATCC (L.Engler and L.Jen-Jacobson, unpublished results).

Effects of base analog substitutions on methylation

The 24mers investigated here contain the T4 Dam recognition site at the center of an 8 bp palindromic sequence 5'-CG(GATC)CG. Earlier, we studied 20mers with a central GGATCC flanked by other sequences (24). In all these studies, we did not observe any modulation of methylation parameters by the flanking sequences; this is in contrast to observations reported (21,22) for *EcoDam*, another GATC-specific MTase. Our results with T4 Dam indicate

that the second (and beyond) base pairs flanking either side of the target site have a negligible effect on the reaction.

Previously, the importance of the Gua residue in the GATC target site for both binding and catalysis was shown for T4 Dam (23,24). In order to delineate the possible role in catalysis by each chemical substituent of Gua, we investigated the effect of replacing the imidazole N7 with a carbon atom (G→D, duplexes 8 and 9) or deleting the O6-keto group (G→N, duplexes 6 and 7), as these are the most likely positions for major groove hydrogen bonding with T4 Dam (18,34). The G→D substitution was found to have a weak effect, expressed as a slight decrease in specificity coefficient compared with unmodified canonical duplex 1. This can be attributed to a single non-critical protein–DNA contact loss, weakly affecting the conversion step and overall course of reaction. In contrast, the G→N substitution greatly affected the kinetic parameters (25-fold reduction in specificity coefficient). This is almost comparable with the effect of deleting a Gua residue in one strand of the recognition site (24). Thus, we conclude that T4 Dam forms a critical contact with Gua at the O6 position.

The effects of base analog substitutions on catalysis have also been studied with other restriction and modification enzymes, e.g. *EcoRI* ENase (17,18) and *EcoRV* MTase (34). For *EcoRI* ENase, G→D and G→N substitutions produce 'structural adaptations' at the protein–DNA interface, resulting in large reductions in cleavage rates (17). As for T4 Dam, the importance of the GAT sequence for effective catalysis was also recognized for the GATATC-specific *EcoRV* MTase (34). However, the roles of the Gua N7 and O6 positions appears to be different for the two enzymes, e.g. G→D was very inhibitory for the *EcoRV* MTase, while G→N had little influence on the methylation rate.

Although not inhibiting binding, the absence of an Ade residue in one strand was shown to completely block T4 Dam methylation of the other strand (24). In this study, A→P (duplexes 2 and 3) and A→N (duplexes 4 and 5) substitutions were introduced to investigate the effect of losing one target N6 amino group and one Watson–Crick hydrogen bond. In contrast to absence of an Ade, the A→P and A→N substituted duplexes remained substrates for

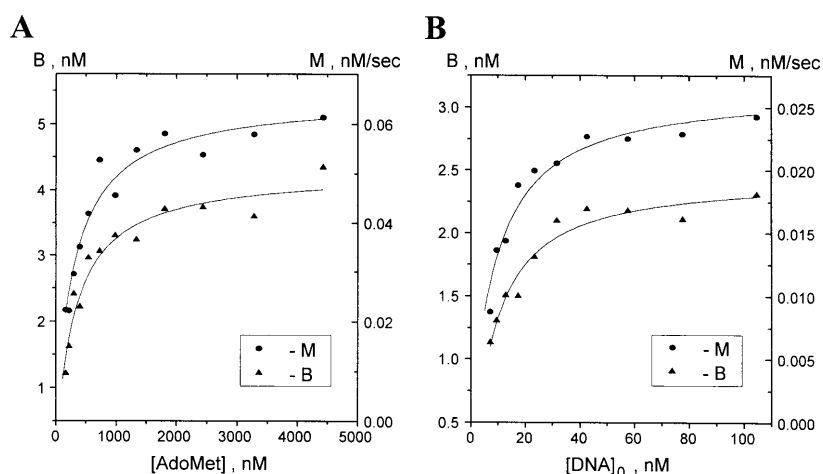


Figure 8. Initial reaction rates (M) and bursts (B) of perfect duplex 1 methylation: (A) as a function of AdoMet concentration (110 nM DNA, 4.5 nM T4 Dam); (B) as a function of DNA concentration (2.25 nM T4 Dam).

methylation, although eliminating the N6 amino group reduced the specificity constant 6- to 7-fold (Table 2).

In conclusion, the results presented here extend our previous findings (23,24) that the presence of standard G:C base pairs at both ends of the palindromic GATC methylation site was critical for stable complex formation with T4 Dam, while the central A:T base pairs had much less influence. As for structural modifications in the target site, the effects of base analog substitutions on stable complex formation did not correlate well with changes in steady-state methylation parameters, K_m and k_{cat} .

ACKNOWLEDGEMENTS

The authors are grateful to L. G. Ovechkina and A. V. Zinoviev for their help in the experiments. This work was supported by a US Public Health Service grant from the Fogarty International Center (no. TW00529), a grant from the Russian Foundation for Fundamental Researches (no. 97-04-49404) and a US Public Health Service grant GM29227 from the National Institutes of Health (to S.H.).

REFERENCES

- Cheng, X. (1995) *Annu. Rev. Biophys. Biomol. Struct.*, **24**, 293–318.
- Anderson, J.E. (1993) *Curr. Opin. Struct. Biol.*, **3**, 24–30.
- Wu, J.C. and Santi, D.V. (1987) *J. Biol. Chem.*, **262**, 4778–4786.
- Cheng, X., Kumar, S., Posfai, J., Pflugrath, J.W. and Roberts, R.J. (1993) *Cell*, **74**, 299–307.
- Cheng, X., Kumar, S., Klimasauskas, S. and Roberts, R.J. (1993) *Cold Spring Harbor Symp. Quant. Biol.*, **58**, 331–338.
- Klimasauskas, S., Kumar, S., Roberts, R.J. and Cheng, X. (1994) *Cell*, **76**, 357–369.
- Reinisch, K.M., Chen, L., Verdine, G.L. and Lipscomb, N. (1995) *Cell*, **82**, 143–153.
- Labahn, J., Granzin, J., Schluckebier, G., Robinson, D.P., Jack, W.E., Schildkraut, I. and Saenger, W. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 10957–10961.
- Gong, W., O'Gara, M., Blumenthal, R.M. and Cheng, X. (1997) *Nucleic Acids Res.*, **25**, 2702–2715.
- Allan, B.W. and Reich, N.O. (1996) *Biochemistry*, **35**, 14757–14762.
- Allan, B.W., Beechem, J.M., Lindstrom, W.M. and Reich, N.O. (1998) *J. Biol. Chem.*, **273**, 2368–2373.
- Holz, B., Klimasauskas, S., Serva, S. and Weinhold, E. (1998) *Nucleic Acids Res.*, **26**, 1076–1083.
- Cal, S. and Connolly, B.A. (1997) *J. Biol. Chem.*, **271**, 1008–1015.
- Schluckebier, G., Labahn, J., Granzin, J., Schildkraut, I. and Saenger, W. (1995) *Gene*, **157**, 131–134.
- Kosykh, V.G., Schlagman, S.L. and Hattman, S. (1993) *Nucleic Acids Res.*, **21**, 4659–4662.
- Kosykh, V.G., Schlagman, S.L. and Hattman, S. (1997) *J. Bacteriol.*, **179**, 3239–3243.
- Lesser, D.R., Kurpiewski, M.R. and Jen-Jacobson, L. (1990) *Science*, **250**, 776–785.
- Lesser, D.R., Kurpiewski, M.R., Waters, T., Connolly, B.A. and Jen-Jacobson, L. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 7548–7552.
- Jen-Jacobson, L. (1995) *Methods Enzymol.*, **259**, 305–344.
- Gromova, E.S., Oretskaya, T.S., Eritja, R. and Guschlbauer, W. (1994) *Biochem. Mol. Biol. Int.*, **36**, 247–255.
- Marzabal, S., Dubois, S., Thielking, V., Cano, A., Eritja, R. and Guschlbauer, W. (1995) *Nucleic Acids Res.*, **23**, 3648–3655.
- Thielking, V., Du Bois, S., Eritja, R. and Guschlbauer, W. (1997) *Biol. Chem.*, **378**, 407–415.
- Malygin, E.G., Petrov, N.A., Gorbunov, Yu.A., Kosykh, V.G. and Hattman, S. (1997) *Nucleic Acids Res.*, **25**, 4393–4399.
- Zinoviev, V.V., Evdokimov, A.A., Gorbunov, Yu.A., Malygin, E.G., Kosykh, V.G. and Hattman, S. (1998) *Biol. Chem.*, **379**, 481–488.
- Zinoviev, V.V., Ovechkina, L.G. and Malygin, E.G. (1996) *Mol. Biol.*, **30**, 724–726.
- Mirzabekov, A.D., Lastity, D., Levina, E.S., Undrizov, I.M. and Bayev, A.A. (1972) *Molekulyar Biol.*, **6**, 87–105.
- Zinoviev, V.V., Gorbunov, Yu.A., Baclanov, M.M., Popov, S.G. and Malygin, E.G. (1983) *FEBS Lett.*, **154**, 282–284.
- Kosykh, V.G., Schlagman, S.L. and Hattman, S. (1995) *J. Biol. Chem.*, **270**, 14389–14393.
- Bradford, M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
- Kourganov, B.I. (1978) In Yakovlev, V.A. (ed.), *Allosteric Enzymes*. Nauka, Moscow, Russia, pp. 27–30.
- Freese, E. (1959) *J. Mol. Biol.*, **1**, 87–105.
- Flynn, J., Azzam, R. and Reich, N. (1998) *J. Mol. Biol.*, **279**, 101–116.
- Klimasauskas, S. and Roberts, R.J. (1995) *Nucleic Acids Res.*, **23**, 1388–1395.
- Szczelkun, M.D., Jones, H. and Connolly, B.A. (1995) *Biochemistry*, **34**, 10734–10743.