Non-additivity of sequence-specific enzyme-DNA interactions in the EcoRI DNA methyltransferase

Norbert O. Reich and Michael J. Danzitz Jr
Department of Chemistry, University of California, Santa Barbara, CA 93106, USA

Received July 12, 1991; Revised and Accepted October 17, 1991

ABSTRACT
We describe a novel strategy to characterize protein-DNA interactions involving monomeric enzymes such as DNA methyltransferases (Mtases). This strategy is applied to our investigation of the EcoRI DNA Mtase, which binds its double stranded recognition site 5'-G-AATTC-3' and methylates the central adenosine of each strand using S-adenosyl-L-methionine as the methyl donor. We show that prior methylation of adenosine in either strand does not perturb catalysis. In contrast, substrates substituted with deoxyinosine at either guanosine position (T-BMI5 and T15-BM) show the minor groove residing N2 amino group of both guanosines contribute to DNA recognition since specificity constants for the modified substrates are reduced 13 and 39 fold. Similar analysis of a substrate containing deoxyinosine at both positions (T15-BMI5) clearly shows that some communication occurs between the sites. To determine the extent to which structural changes in the DNA alone contribute to this lack of additivity, we performed DNA melting analysis of the singly and doubly substituted substrates, and also found nonadditivity. Although our functional and structural analyses suggest that deoxyinosine incorporation causes long range conformational effects, the similarity of KMAdoMet for all substrates suggests that no large-scale structural changes occur in the Mtase-DNA-AdoMet complex. Our results support the following conclusions: 1) The non-additivity shown in this system contrasts with the widespread demonstration of additivity involving repressors [Lehming et al., 1990; Takeda et al., 1989; Ebright et al., 1987], suggesting that sequence discrimination by enzymes may involve more complex mechanisms. Further, this non-additivity precludes quantitative assignment of individual interactions and we suggest that future analyses of this and related enzyme systems with base analogs include detailed information about the long range structural consequences of individual substitutions. 2) Although T15-BM and T-BMI5 are shown to be radically different by thermodynamic analysis, the similar specificity constants with the Mtase suggest that the underlying structural differences (e.g., altered helical parameters of the DNA) are not critical for sequence-recognition. 3) The significance of minor groove Mtase-DNA interactions to specificity is confirmed.

INTRODUCTION
How do DNA binding and modifying proteins discriminate between DNA sequences with such incredible specificity? Our understanding of the underlying mechanisms to this question has progressed considerably, particularly with regards to prokaryotic repressors [Anderson et al., 1987; Otwinowski et al., 1988; Jordan and Pabo, 1988; Wolberger et al., 1988; Rafferty et al., 1989]. A common element of protein-mediated sequence-specific DNA binding is a complex network of hydrogen bonds and hydrophobic and ionic interactions [Seeman et al., 1976; von Hippel & Berg, 1986], yet the diversity of observed recognition mechanisms suggests that no set of simple rules apply [Matthews, 1988]. Also, a quantitative understanding of how individual protein-DNA interactions contribute to specificity remains to be elucidated. This is particularly true for DNA-modifying enzymes whose sequence-specificity need not be determined solely by binding phenomena.

We are investigating the EcoRI1 DNA Mtase to understand the molecular basis of enzymic DNA sequence discrimination. Detailed information about the protein-DNA interface for the EcoRI Mtase, any other DNA Mtase or any monomeric DNA-binding protein remain to be uncovered. The Mtase is part of the well studied EcoRI restriction-modification system. The Mtase is a 38,050 dalton protein which is active as a monomer and requires S-adenosylmethionine (AdoMet) and its DNA substrate for activity [Rubin and Modrich, 1977]. Methylation occurs at the N6 position of the second adenine in the double stranded DNA sequence 5' GAATTC3'. We recently demonstrated that initial binding of the enzyme to AdoMet and noncanonical DNA is random, but that canonical site recognition and subsequent catalysis requires AdoMet to be bound [Reich and Mashhoon, 1991]. The Mtase is an extremely efficient enzyme with a specificity constant (kcat/KmDNA) greater than 10^8 s^{-1} M^{-1}, and catalysis is limited by product release step(s) with the methyl transfer step being at least ten times faster than kcat.

Structural data for the EcoRI DNA Mtase substrate and product provide an excellent basis for investigating the details of sequence-specific recognition. The X-ray [Wing et al., 1980] and solution structures [Nerdal et al., 1989] of a DNA sequence containing the canonical site are available (see Figure 1). The X-ray structure of the identical sequence in which both central adenosines are methylated shows that methylation causes minimal structural changes and highlights the proximity of the two methyl groups in the major groove [Frederick et al., 1988] (see Figure 2).

Our approach to understanding Mtase-DNA interactions is by functional analysis of DNA substrates lacking structural features proposed to be critical for DNA recognition. Previous
investigations have implicated the exocyclic amine of guanosine, located in the minor groove, as being critical to the EcoRI Mtase [Modrich and Rubin, 1977; Brennan et al., 1986b]. (See Figures 1, 2 and 3). This is a provocative finding since sequence-recognition by proteins is thought to occur largely through interactions in the major groove [Dickerson et al., 1987]. Moreover, since methylation occurs in the major groove, the Mtase would be required to interact with both grooves simultaneously (Figure 1). For these reasons, as well as the prospect of testing our strategy with a conservative substitution (deoxyinosine replacement of guanosine), we investigated the contribution toward specificity which derives from Mtase-DNA interactions at this position.

We designed synthetic oligonucleotide substrates that allow investigation of a single binding orientation of the Mtase. Incorporation of a modified base (e.g., deoxyinosine for guanosine) within the canonical site should allow assessment of individual Mtase-DNA interactions. Our functional analyses of these modified substrates include true steady state parameters for AdoMet and DNA. The structural consequences of the functional group deletion are probed by comparisons of $K_{m}^{AdoMet}$, DNA stability using DNA melting and 'second site' analysis [Carter et al., 1984].

MATERIALS AND METHODS

Materials
Phosphoramidites and ancillary DNA synthesis reagents were obtained from Milligen/Biosearch and American Bionetics, betacyanoethyl-$N^6$-methyladenine from Pharmacia. Gamma $^{32}$P-ATP (5,000 Ci/mmole) was purchased from Amersham, and [methyl $^3$H]-AdoMet (78.5 Ci/mmole) from New England Nuclear. T4 polynucleotide kinase was from New England Biolabs. AdoMet was purchased from Sigma and DE81 filter papers were from Whatman. Unlabeled AdoMet was purified using an HRLC MA7S (50x7.8 mm) cation exchange column according to a modified procedure of Zappia et al. [1980]. Ammonium acetate buffer (0.1 M, pH 4.0) was used to elute AdoMet at a flow rate of 1 ml/min. The peak corresponding to AdoMet was collected, concentrated on a BioRex 70 cation exchange column (BioRad) and stored at $-20\degree C$ in 0.1 N HCl [Reich and Everett, 1990]. EcoRI Mtase was obtained from an overproducing strain of E.coli and purified in our laboratory [Greene, et al., 1978].

Oligonucleotide synthesis and purification
Oligonucleotides were synthesized on a Biosearch 3810 DNA synthesizer using beta-cyanoethyl phosphoramidites. Oligonucleotides were purified on a 10 micron 4.6x250 mm Vydac C-18 column, eluted with a 15%-85% to 40%-60% acetonitrile:triethylammonium acetate pH 7.0 (TEAAc) 20 minute generator' which was previously calibrated, and produced a 0.25°C/minute ramp from 15 to 88°C. Oligonucleotide solutions consisted of 10 mM phosphate, 0.1 mM EDTA, 20 mM NaCl pH 7.0 and double stranded oligonucleotide at various concentrations (from 6 to 9 different concentrations). Solutions were stirred throughout the analysis. The resulting melting temperatures ($T_m$) were determined using the sloping baseline-sloping plateau method [Breslauer, 1986]. The enthalpy and entropy terms were determined from the slopes and Y-intercepts respectively of $R\ln(I/4C)$ vs. $1/T_m$ plots (R is 1.987 cal/mole/°K), such as those depicted in Figure 5, where C is the total concentration of DNA expressed as single strands and temperature is in degrees Kelvin [Aboul-ela et al., 1985]. Our

Figure 1: Graphical representation of the EcoRI recognition site containing 'Dickerson dodecamer' [Dickerson and Drew, 1981]. The N2 amino group of one guanosine is shaded black while the sugar phosphate backbone is shaded light grey. Bases are dark grey. The two N6 amino groups of the central adenines are white. The methyl group of one thymine is also shaded black to identify the major groove.

Oligonucleotide purity was confirmed by denaturing polyacrylamide gel electrophoresis (PAGE) of $^{32}$P labeled oligonucleotides. Oligonucleotides were labeled by reaction with a 4 fold molar excess of gamma $^{32}$P-ATP and 0.1 units of T4 polynucleotide kinase per picomole of oligonucleotide using published reaction conditions [Maniatis et al., 1982]. The labeled oligonucleotides were electrophoresed through a 25% denaturing polyacrylamide sequencing gel and an autoradiogram of the gel was used to estimate the homogeneity of each sample. Double stranded substrate formation was confirmed by annealing a radiolabeled oligonucleotide to the complementary unlabeled oligonucleotide using the method described earlier. Samples were electrophoresed through a 14% nondenaturing polyacrylamide gel and analyzed by autoradiography.

Oligonucleotide melting analysis
The temperature dependent denaturation of double stranded oligonucleotides to single strands was monitored at 260 nm with a Shimadzu UV-265 UV/Vis spectrophotometer equipped with a water jacketed cuvette holder. The temperature of the oligonucleotide solution was controlled with a Haake A80 circulating bath connected to the cuvette holder. The output of this bath was regulated by a homemade 'temperature ramp generator' which was previously calibrated, and produced a 0.25°C/minute ramp from 15 to 88°C. Oligonucleotide solutions consisted of 10 mM phosphate, 0.1 mM EDTA, 20 mM NaCl pH 7.0 and double stranded oligonucleotide at various concentrations (from 6 to 9 different concentrations). Solutions were stirred throughout the analysis. The resulting melting temperatures ($T_m$) were determined using the sloping baseline-sloping plateau method [Breslauer, 1986]. The enthalpy and entropy terms were determined from the slopes and Y-intercepts respectively of $R\ln(I/4C)$ vs. $1/T_m$ plots (R is 1.987 cal/mole/°K), such as those depicted in Figure 5, where C is the total concentration of DNA expressed as single strands and temperature is in degrees Kelvin [Aboul-ela et al., 1985]. Our
EcoKl methylases methylate their palindromic hexanucleotide and Modrich, 1977. This observation and the fact that the Mtase recognition site on both strands is separated in intimate catalytic events [Rubin et al., 1988].

**EXPERIMENTAL DESIGN**

EcoRI Mtase methylates its palindromic hexanucleotide recognition site on both strands in separate catalytic events [Rubin and Modrich, 1977]. This observation and the fact that the Mtase is active as a monomer led us to speculate that each strand is methylated from a unique binding orientation. This is schematically depicted in Figure 4 in which the Mtase binds the native substrate (A) in two orientations (B and C). If true, then functional parameters such as $k_{cat}$, $K_m$ and $k_{cat}/K_m$ will result from both orientations, making it impossible to comment on unique interactions. We addressed this problem by using hemi-methylated DNA substrates. Although nonproductive binding is expected to occur based on the inhibition constant for a doubly methylated site [Reich and Mashhoon, 1991], contributions to $k_{cat}$ and $K_m$ will cancel so that $k_{cat}/K_m$ will not be influenced [Fersht, 1985]. The Mtase interacting with hemi-methylated substrates is shown in Figures 4D and 4E.

The initial substrate is shown in Figure 4A, and like all substrates used in this study, is non-self-complementary in sequence and fourteen basepairs in length. The use of hemi-methylated substrates described above requires non-self-complementary DNA strands which was achieved with the basepairs flanking the canonical hexanucleotide. Since our kinetic analyses are performed at 37°C and 20 mM NaCl, the substrates had to be long enough to ensure thermal stability. Moreover, Mtase-DNA interactions beyond the canonical hexanucleotide should be satisfied; preliminary DNA-footprinting results from our laboratory have shown a minimum of eight, and possibly ten basepairs are protected by the Mtase. Finally, to facilitate our structural analysis of modified substrates, the DNA sequence flanking the canonical site (Figure 4A) was selected based on the ‘Dickerson dodecamer’ which has been structurally characterized by both X-ray (Figure 2) [Wing et al., 1980] and NMR [Nerdal et al., 1989] techniques.

Our strategy should lead to assignment of DNA functionalities critical for recognition and catalysis by the EcoRI Mtase. More optimistically, assessment of the quantitative contributions toward specificity may be obtained [Fersht et al., 1986]. The most compelling criticism of such quantitative interpretation is lack of structural information concerning the modified DNA and Mtase-DNA complex. We therefore studied the solution structures of the modified DNA substrates with thermal stability studies; comparisons were at the level of enthalpy and entropy changes. $K_{cat}^{AdoMet}$ was also determined to assess structural perturbations in the ternary complex (Mtase-DNA-AdoMet). Finally, we used 'second site' substitutions in the canonical site to determine any structural communication between sites [Carter et al., 1984].

**RESULTS**

We submitted several synthetic DNA substrates of the EcoRI DNA Mtase to a structural and functional analysis. The substrates are all fourteen basepairs in length and the sequences shown in Tables I and II. The set in Table I was designed to determine
the effect of hemi-methylation on Mtase function. We define function as the catalytic turnover constant ($k_{cat}$), Michaelis constants ($K_m^{DNA}$ and $K_m^{AdoMet}$) and specificity constant ($k_{cat}/K_m^{DNA}$) [Fersht, 1985]. The set in Table II was designed to assess the importance of the guanosine minor groove amine, which we approached by substituting deoxyinosine for guanosine and thereby replacing the amine with a hydrogen atom.

### Preliminary characterization of DNA substrates and assay conditions

DNA substrates were shown to be greater than 98% pure by denaturing polyacrylamide gel electrophoresis of $^{32}$P labeled oligonucleotides. Extent of annealings were shown to be greater than 95% by non-denaturing PAGE of $^{32}$P labeled substrates annealed with a 2 fold molar excess of the bottom strand. *EcoRI* Mtase was shown to have maximal activity at pH 7.8 and 10–20 mM NaCl. For all conditions of DNA and AdoMet used in this report, the formation of methylated DNA was shown to be constant with respect to time. Enzyme activity was first order with respect to enzyme concentration. Under our assay conditions, no methylation was detected with single-stranded oligonucleotides. Enzyme activity was monitored as a function of annealing ratio (1:1, 1:2, 1:3 and 1:4; top:bottom strand), and followed a saturation curve saturating at a 1:2 ratio. (Data not shown).

### Kinetic analysis

Figure 6 shows a representative double reciprocal analysis for the control substrate, CBM. This plot, the related double reciprocal analysis of velocity versus AdoMet concentration and the plots for the remaining DNA substrates are all consistent with the proposed kinetic mechanism [Reich and Mashhoon, 1991]. The resultant steady-state parameters are shown in Tables I and II. Our previous characterization of substrate C showed a significantly lower $K_m^{DNA}$ [Reich and Mashhoon, 1991]. This previous analysis required a larger set of data and used Cleland’s methods (1979) rather than the slope replot approach used in the present study. Application of Cleland’s methods to the data collected for C in this study generates parameters comparable to our previously published values. Because we are obtaining specificity data for numerous substrates we used the replot approach which requires fewer data points. In any event, all substrates in the present study have been submitted to identical data analysis methods and our specificity analysis is thus valid.

To assess the effect that hemi-methylation has on specificity, we determined the kinetic parameters for a native substrate (C) and the corresponding hemi-methylated substrates (CBM and C-TM), as shown in Table I. The data indicate that C, CBM and CTM have essentially identical $k_{cat}$ values while $K_m^{DNA}$ for both hemi-methylated substrates are approximately two times greater than that of the non-methylated substrate. The difference in available sites may be responsible for the observed $K_m^{DNA}$ differences. The similarity of the corresponding specificity constants indicates that the Mtase is relatively insensitive to the N6 methyl group on the strand opposite to the site of methylation. Further support for the lack of significant structural changes comes from the minimal changes in $K_m^{AdoMet}$ values. Thus, although canonical site recognition requires prior AdoMet binding, $K_m^{AdoMet}$ is 250 fold smaller than $K_m^{DNA}$ for both hemi-methylated substrates are approximately two times greater than that of the non-methylated substrate. The difference in available sites may be responsible for the observed $K_m^{DNA}$ differences. The similarity of the corresponding specificity constants indicates that the Mtase is relatively insensitive to the N6 methyl group on the strand opposite to the site of methylation.

Further support for the lack of significant structural changes comes from the minimal changes in $K_m^{AdoMet}$ values. Thus, although canonical site recognition requires prior AdoMet binding, $K_m^{AdoMet}$ is 250 fold smaller than $K_m^{DNA}$ for both hemi-methylated substrates are approximately two times greater than that of the non-methylated substrate. The difference in available sites may be responsible for the observed $K_m^{DNA}$ differences. The similarity of the corresponding specificity constants indicates that the Mtase is relatively insensitive to the N6 methyl group on the strand opposite to the site of methylation.
Table I. Steady state kinetic parameters determined for the non- and hemi-
by methylated DNA substrates. Parameters were determined from double reciprocal data as shown in Figure 6; standard errors shown in parentheses. (* denotes methylated adenine).

| Abbreviation | Sequence | V (pM) | kcat/KmDNA (pM) | kcat/Km | kcat/νDNA acc-1 | kcat/νDNA acc-1m 2
|---------------|----------|--------|-----------------|--------|-----------------|----------------
| CBM           | 5' GCGAATTCCGGCCGCGCCTTTAAGCGCC 3' | 24.4 (10.3) | 0.280 (0.198) | 0.14 | 4.13 X 10^5 |
|               | 3' GCGAATTCCGGCCGCGCCTTTAAGCGCC 5' | (10.3) | (0.198) | (0.14) | (4.13 X 10^5) |
| T5-BM         | 5' GCGAATTCCGGCCGCGCCTTTAAGCGCC 3' | 43.9 (12.9) | 0.184 (0.099) | 0.105 | 3.45 X 10^6 |
|               | 3' GCGAATTCCGGCCGCGCCTTTAAGCGCC 5' | (12.9) | (0.099) | (0.105) | (3.45 X 10^6) |
| T5-BMIS       | 5' GCGAATTCCGGCCGCGCCTTTAAGCGCC 3' | 43.0 (3.0) | 0.322 (0.107) | 0.13 | 3.05 X 10^6 |
|               | 3' GCGAATTCCGGCCGCGCCTTTAAGCGCC 5' | (3.0) | (0.107) | (0.13) | (3.05 X 10^6) |

Table II. Steady state kinetic parameters determined for the hemi-methylated control and the deoxyinosine containing substrates. Parameters were determined from double reciprocal data as shown in Figure 6; standard errors shown in parentheses. (* denotes methylated adenine).

| Abbreviation | Sequence | V (pM) | kcat/KmDNA (pM) | kcat/Km | kcat/νDNA acc-1 | kcat/νDNA acc-1m 2
|---------------|----------|--------|-----------------|--------|-----------------|----------------
| CBM           | 5' GCGAATTCCGGCCGCGCCTTTAAGCGCC 3' | 24.4 (10.3) | 0.280 (0.198) | 0.14 | 4.13 X 10^5 |
|               | 3' GCGAATTCCGGCCGCGCCTTTAAGCGCC 5' | (10.3) | (0.198) | (0.14) | (4.13 X 10^5) |
| T5-BM         | 5' GCGAATTCCGGCCGCGCCTTTAAGCGCC 3' | 43.9 (12.9) | 0.184 (0.099) | 0.105 | 3.45 X 10^6 |
|               | 3' GCGAATTCCGGCCGCGCCTTTAAGCGCC 5' | (12.9) | (0.099) | (0.105) | (3.45 X 10^6) |
| T5-BMIS       | 5' GCGAATTCCGGCCGCGCCTTTAAGCGCC 3' | 43.0 (3.0) | 0.322 (0.107) | 0.13 | 3.05 X 10^6 |
|               | 3' GCGAATTCCGGCCGCGCCTTTAAGCGCC 5' | (3.0) | (0.107) | (0.13) | (3.05 X 10^6) |

that communication occurs between each of the two individual sites. To quantitate this nonadditivity at the DNA level, we characterized the thermal stability of each substrate.

Substrate melting analysis

All thermodynamic constants were derived from the temperature-
dependent double to single-stranded transition of the oligonucleotides assuming a true two state system. In all cases the absorbance versus temperature profiles were consistent with this assumption, in contrast to the complex behaviour observed with completely self-complementary oligonucleotides [Breslauer et al., 1987]. Our analysis was initially used to confirm that greater than 80% of each substrate remains in the double stranded form under our assay conditions. As shown in Table III, the deoxyinosine substitutions change the Gibbs free energy for the double- to single-stranded transition by less than one kcal/mole (at 37°C). As with the kcat/KmDNA analysis (Table II), comparison of free energy values shows the interdependence of each deoxyinosine substitution since T5-BM15 is more stable than T-BM15. Hence, removal of one amino group perturbs the DNA structure enough to affect the imino/cytosine basepair five bases away. To further probe the structural origins of this effect we determined the enthalpic and entropic components of the free energy term by characterizing the dependency of melting temperature on the DNA concentration [Breslauer, 1986; Petruska et al., 1988]. As shown in Figure 5 and Table III the melting transitions of T15-BM and T-BM15 have decreased enthalpy and entropy terms relative to CBM. This type of 'compensation' is well known [Breslauer et al., 1987; Petruska et al., 1988], although an explanation at the molecular level is not available. Surprisingly, the changes for T-BM15 relative to CBM are extremely large.

Figure 6: 1/V vs. 1/[DNA] plot for the control substrate CBM at 5 different AdoMet concentrations (3.0 μM O, 1.0 μM ●, 0.6 μM △, 0.45 μM ▲, 0.3 μM ▽). The inset shows the 1/V vs. 1/[AdoMet] μM^-1 replot used to determine k_m AdoMet^-1.

Table III. Thermal stability data determined for the hemi-methylated control and the deoxyinosine containing oligonucleotides obtained at pH 7.0 and 20 mM NaCl. Units for ΔH° and ΔAΔH° are Kcal/mole, ΔS° and ΔAΔS° are cal/(K mole)^-1, and ΔG° and ΔAΔG° are Kcal/mole. All ΔG° values were calculated from ΔH° and ΔS° using ΔG° = ΔH° - TΔS° at 310°C. Derivation of standard errors (parentheses) is described in Materials and Methods.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>ΔH° (Kcal)</th>
<th>ΔAΔH° (cal)</th>
<th>ΔS° (Kcal)</th>
<th>ΔAΔS° (cal)</th>
<th>ΔG° (Kcal)</th>
<th>ΔAΔG° (cal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBM</td>
<td>80.74</td>
<td>227.8</td>
<td>10.9</td>
<td>-</td>
<td>80.74</td>
<td>-</td>
</tr>
<tr>
<td>T5-BM</td>
<td>77.42</td>
<td>217.4</td>
<td>9.87</td>
<td>3.22</td>
<td>80.74</td>
<td>-</td>
</tr>
<tr>
<td>T5-BM15</td>
<td>39.69</td>
<td>88.8</td>
<td>9.04</td>
<td>44.05</td>
<td>80.74</td>
<td>-</td>
</tr>
<tr>
<td>T5-BM15</td>
<td>85.55</td>
<td>245.5</td>
<td>9.41</td>
<td>-8.81</td>
<td>80.74</td>
<td>-</td>
</tr>
</tbody>
</table>

DISCUSSION

Investigations of enzymic specificity frequently exploit substrate and enzyme modifications. Modifications in either type of molecule have numerous attendant problems, although investigations with 'small' substrates have led to mechanistic insights [Fersht et al., 1986; Kati and Wolfenden, 1989]. It is clear from studies with small substrate systems that a mechanistic understanding of specificity requires detailed quantitative data. The use of modified bases to assess contributions of individual protein-DNA interactions toward specificity has been widespread, particularly in the context of bacterial Type II restriction-modification enzymes [Brennan et al., 1986a; Diekmann and McLaughlin, 1988; Modrich and Rubin, 1977; Fliess et al., 1986; McLaughlin et al., 1987b; Seela and Kehne, 1987; Bodnar et al., 1983; Berkner and Wolfenden, 1979; Wolfes et al., 1986; Newman et al., 1990a; Cossick et al., 1990]. However, for a variety of reasons these past studies have been limited to a qualitative analysis. A major concern is the extent to which the desired structural modification causes long range conformational
changes, thereby complicating assignment of individual interactions. Also, in the absence of knowledge about the kinetic mechanism and rate limiting step(s), interpretation of specificity data is hampered. Two additional concerns are specific to studies of DNA Mases: 1) These enzymes all use AdoMet as a cofactor and thus an accurate specificity analysis requires comparison of true kinetic constants, in contrast to the reported apparent values [Brennan et al., 1986b; Modrich and Rubin, 1977; Newman et al., 1990a, b]. 2) Most DNA Mases are functional monomers which can bind to their canonical sites, and more importantly, modified sites, in two orientations. This has precluded investigations of unique Mtase-DNA complexes and an understanding of individual Mtase-DNA interactions.

Our strategy (see Results) addresses these concerns and was applied to an investigation of the contribution of the minor groove amino functionality of guanosine to substrate specificity in the EcoRI DNA Mtase system. The position of the amino group within the GC basepair is shown in Figure 3, and its location in the minor groove is shown in Figure 1. Figures 1 and 2 show structures derived from X-ray crystallographic analyses of the EcoRI Mtase DNA substrate [Wing et al., 1980] and methylated product [Frederick et al., 1988] respectively. Both dodecameric sequences contain the canonical site and three flanking basepairs identical to those found in our DNA substrates (see Figures 1, 2 and 4, and Tables I and II). Figure 1 shows the minor groove is quite narrow in this region, with the amino group near the bottom of the groove facing away from the site of methylation in the major groove (N6 position of centrally located adenines).

To determine the contribution of Mtase interactions with the guanosine amino group we substituted deoxyinosine for guanosine, effectively replacing the amine with a hydrogen. This is a conservative replacement, as suggested by the insensitivity of the EcoRI endonuclease to incorporation of deoxyinosine into the canonical site [Modrich and Rubin, 1977; Brennan et al., 1986a]. We selected this analog to test our system since previous studies with the EcoRI Mtase have also used deoxyinosine [Modrich and Rubin, 1977; Brennan et al., 1986a]. Modrich and Rubin [1977] used a plasmid DNA substrate with a single recognition site and found that activity with the deoxyinosine containing substrate was reduced 25 fold. Brennan et al. [1986a] found that a deoxyinosine containing octameric substrate was completely inactive. These investigations involved substrates in which both guanosines were replaced with deoxyinosine, analogous to our substrate T15-BMI5.

Effect of hemi-methylation on Mtase function

Since our strategy requires the use of hemi-methylated substrates, we analyzed the functional characteristics of both hemi-methylated substrates (CTM and CBM) and compared these with an unmethylated substrate (C). Our kinetic analyses, shown in Table I, demonstrate no effects on $k_{cat}$ and only minor changes on $K_m$ (AdoMet and DNA). However, due to the methylation rate constant ($k_{methylation}$) being at least ten times faster than $k_{cat}$ [Reich and Mashhoon, 1991], changes in $k_{methylation}$ less than ten fold would not have altered $k_{cat}$.

The apparent insensitivity of the Mtase to prior adenine methylation of one strand is surprising given the proximity of the two N6 methyl groups, shown in Figure 2 for the doubly methylated dodecameric product. The functional similarity of the hemi-methylated and native substrates supports a mechanism in which methylation of each strand occurs from a unique binding orientation. The alternative mechanism (methylation of either strand from one binding orientation) is more likely to be sensitive to the presence of a methyl group on one of the potential sites of methylation.

Effect of minor groove guanosine amino on Mtase function

The results with the hemi-methylated substrates (CBM and C- TM) supported the use of one of these as a control substrate in our specificity analysis. We selected CBM, and Table II shows that $k_{cat}$ for the two singly-modified deoxyinosine-containing substrates (T-BMI5 and T-I5-BM) are not significantly perturbed in comparison to CBM. In contrast, the significant increases in $K_m$ with these substrates suggests that both guanosine amino groups contribute to Mtase-DNA interactions. A comparison of the specificity constants confirms that T15-BM and T-BMI5 are poorer substrates by a factor of of 13 and 39 fold respectively, equivalent to a loss in binding energy of the enzyme for the transition state of 1.56 and 2.23 kcal/mole. Although these apparent binding energies might be over- or under-estimations of the true binding energy, they do reflect differences in specificity [Fersht, 1986]. Thus, differences in $k_{cat}/K_m$ directly implicate the deleted minor groove functionality in Mtase-DNA interactions since $k_{cat}/K_m$ relates the conversion of free enzyme and free DNA to the transition state [Fersht, 1986]. Most structural data [Anderson et al., 1987; Otwinowski et al., 1988; Jordan and Pabo, 1988; Wolberger et al., 1988; Rafferty et al., 1989] and predicted models for sequence recognition [Seeman et al., 1977; von Hippel and Berg, 1986; Dickerson et al., 1987] implicate the DNA major groove as the site of protein-DNA interaction. Nevertheless, results from a variety of systems suggest that minor groove interactions do occur. Phage 434 repressor and Cro proteins both have an arginine which fits in the minor groove of the DNA, although a mutant repressor with an alanine at this position has wild-type specificity [Anderson et al., 1987; Wolberger et al., 1988]. X-ray crystallographic analyses of DNase I [Suck et al., 1988] and the histones in the nucleosome core particle [Richmond et al., 1984] show minor groove interactions. Finally, the tetrapeptide (Ser-Pro-Lys-Lys), found in histone H1, forms a structural motif which binds to the minor groove [Churchill and Suzuki, 1989]. Our results with the EcoRI DNA Mtase also support minor groove interactions and further implicate such interactions as determinants of sequence specificity.

Our analysis depends on the modified DNA substrate and Mtase-DNA-AdoMet complex remaining relatively structurally unperturbed. As for the hemi-methylated substrates, our functional analysis included comparisons of $K_m$ and $k_{cat}$. In support of the protein conformation remaining unaffected (in the Mtase-DNA-AdoMet complex), the $K_m$ and $k_{cat}$ values for the different DNA substrates are essentially identical (Table II). An additional approach to determine the structural integrity of the modified substrates and/or Mtase-DNA complexes relies on the extent to which multiple substitutions cause additive functional changes. Previous application of this approach, limited to DNA-binding proteins and substitutions of entire basepairs, demonstrated additivity in the majority of cases [Sarai and Takeda, 1989; Lehming et al., 1990; Takeda et al., 1989; Ebright et al., 1987], although nonadditivity has also been observed [Mossing and Record, 1985]. A more subtle application by Carter et al. [1984] with modified proteins was used to elucidate relatively minor structural changes. Here we combined the single modifications introduced in T15-BM and T-BMI5 to generate T15-BMI5 (Table II). Clearly, effects of the two substitutions are not additive since
k_{\text{cat}}/K_{\text{m, DNA}} is greater than for T-BMI5. This implies that modification at one position alters the structure of the DNA, the enzyme or both sufficiently to perturb enzyme-DNA interactions at the second site. This precludes quantitative assignment of interactions probed with the individually substituted substrates. Further, our demonstration of the lack of additivity, at least with these substrates implicates a DNA recognition mechanism which is more complex than that found for the previously characterized repressor systems [Sarai and Takeda, 1989; Lehming et al., 1990; Takeda et al., 1989; Ebright et al., 1987]. To investigate the nonadditivity at the level of the DNA substrate we undertook a detailed comparison of the thermodynamic parameters associated with the double- to single-stranded transition.

**Thermodynamic analysis of DNA substrates**

Our initial interest with the DNA melting analyses was to confirm the thermal stability of the modified substrates. This was verified (Table III) and the deoxyinosine substitutions change the Gibbs free energy for the double- to single-stranded transition by less than one kcal/mole at 37°C. Additivity analysis at this level also shows the interdependence of the individual substitutions since T15-BMI5 is more stable than T-BMI5. Therefore, although the effects are small, our thermodynamic analysis suggests that the relatively minor single substitutions perturb the structure of the DNA sufficiently to alter the conformation at the non-substituted position. To characterize these effects further we determined the enthalpic and entropic components of the free energy term (Figure 5 and Table III) [Breslauer, 1986; Petruska et al., 1988]. In comparison to CBM, T15-BM and T15-BMI5 show minor compensating alterations in enthalpy and entropy. Although well known [Breslauer et al., 1987; Petruska et al., 1988], we have found no generally accepted explanation for this type of compensation. Because the effects on the enthalpy and entropy for these substrates (and T-BMI5, discussed below) are larger than the free energy changes, the non-additivity is even more obvious.

In contrast to T15-BM and T15-BMI5, T-BMI5 when compared to CBM shows a large decrease in enthalpy and increase in entropy. A plausible explanation of this unusual behaviour assumes that the denatured strands of both CBM and T-BMI5 have comparable enthalpy and entropy terms. This is reasonable since although the structures of the denatured form(s) are not known, the most significant difference between the deoxyinosine- and guanosine-containing single strands should be the ability to form a small number of additional hydrogen bonds. The differences in enthalpy and entropy therefore derive predominantly from differences in the double stranded forms; hence, the chemical bonds comprising the T-BMI5 structure (and associated waters) are overall weaker by approximately 44 kcal/mole than those in the CBM structure. The smaller entropy change implies that the T-BMI5 structure (and associated waters) is significantly more disordered than the CBM structure (Table III).

Although extrapolation of our thermodynamic data toward a structural understanding of the modified DNA is not straightforward, our results show that conclusions based solely on free energy of melting data can be misleading [McLaughlin et al., 1987; Cosstick et al., 1990; Seela and Kehne, 1987; Newman et al., 1990a, b; Seela and Driller, 1986; Fliess et al., 1986, 1988]. Similar results have been reported for DNA-ligand interactions [Breslauer et al., 1987]. Potential structural explanations of our thermodynamic data are based on the extensive structural characterization of the EcoRI Mtase DNA substrate (Figure 1) [Wing et al., 1980; Nerdal et al., 1989]. The DNA has an extremely narrow minor groove in the central tetranucleotide region (AATT) in which a 'spine' of water molecules has been identified [Drew and Dickerson, 1981]. These water molecules are extensively hydrogen bonded with DNA functionalities and the guanosine amino group in a GC basepair precludes a water molecule from such interactions. This constraint is removed when deoxyinosine is used in place of guanosine, making the minor grooves of inosine-cytosine and adenine-thymine basepairs similar. A likely outcome of this is the extension of the spine of hydration to include five (T15-BM and T-BMI5) or six (T15-BMI5) basepairs. Such constrained waters might contribute to an increased degree of order in the inosine substituted substrates; our observed decreases in entropy (smaller changes in entropy, see Table II) with TI-BM and T-BMI5 are at odds with this prediction, suggesting the importance of additional factors.

The EcoRI DNA Mtase substrate shown in Figure 1 is completely self-complementary and contains two symmetrically positioned 5'CGA3'/5'TCG3' sequences [Wing et al., 1980; Nerdal et al., 1989]. This trinucleotide sequence has unusual helical parameters as identified by crystallographic analysis [Wing et al., 1980; Dickerson and Drew, 1981; Frederick et al., 1988], DNase digestion [Lomonossoff et al., 1981], and N.M.R. analysis [Nerdal et al., 1989; Chary et al., 1987]. The large variations in helical twist (between the AT and GC basepair, 39 degrees; between GC and CG, 33 degrees) result from movement of the cytosine away from the adjacent thymine to increase its overlap with the adjacent guanosine [Dickerson and Drew, 1981]. In contrast to the dodecamer (Figure 1), the substrates studied in this report are not self-complementary outside the recognition hexanucleotide (Figure 4 and Tables I and II). Thus, 5'CGA3'/5'TCG3' appears only once, while 5'GG-A3'/5'TCC3' occurs at the opposite end. Structural analysis of the 5'GG-A3'/5'TCC3' sequence shows no unusual features in the helical twist parameter [Nilges et al., 1987]. This presumably results from the cytosine being adjacent to another cytosine rather than a guanosine. In the case of T15-BM, the deoxyinosine replacement creates a 5'GIA3'/5'TCG3' segment, while in T-BMI5, deoxyinosine replacement occurs within the structurally anomalous 5'CGA3'/5'TCG3' segment, generating 5'CIA3'/5'T- CG3'. Since each substitution only removes a single basepairing hydrogen bond, the differential changes in thermal parameters imply that the DNA (T-BMI5) and/or associated waters are conformationally altered. We are investigating the solution structures of CBM, T-BMI5 and T15-BM with N.M.R. methods.

Although deoxyinosine incorporation into the 5'CGA3'/5'T- CG3' segment (T-BMI5) may cause significant structural changes in the DNA, the specificity constants for T-BMI5 and T15-BM vary only threefold (Table II). How, or the extent to which the 5'CGA3'/5'TCG3' segment is altered is not known, although some structural features of the DNA (e.g., helical twist, propeller twist, groove width, flexibility) will likely be affected. These structural features of DNA have been proposed to contribute to the mechanisms whereby proteins discriminate DNA sequences [Dickerson and Drew, 1981a; Dickerson and Drew, 1981b; Calladine, 1982; Sarai et al., 1989]. Unfortunately, this intriguing suggestion is experimentally difficult to investigate. The functional similarity of T-BMI5 and T15-BM provides evidence that these features within the canonical site may not be primary determinants of sequence specificity for EcoRI Mtase.
Relationship to previous analysis of the EcoRI DNA Mtase

Previous studies with the EcoRI Mtase relied on recognition sites containing deoxyinosine at both guanosine positions, equivalent to a nonmethylated version of T15-BM15 [Modrich and Rubin, 1977; Brennan et al., 1986a]. The plasmid substrate used by Modrich and Rubin contained a single canonical site and was 25 fold less active when guanosine was replaced with deoxyinosine. These results are surprisingly similar to ours (14 fold, Table II), since the plasmid data was obtained at single concentrations of DNA and AdoMet. Based on the 26 fold poorer binding of the T15-BM15 substrate (Table II), the difference is easily explained by the lack of saturation with the inosine-containing plasmid. In contrast, the work by Brennan et al. [1986a] showed that a deoxyinosine containing octameric oligonucleotide was completely inactive. Comparison with our work is compromised by several factors. Their assays were done at 10°C, and catalytic turnover with an unmodified substrate was 80 times slower than substrate C (Table I). Based on our footprinting data, their octameric substrates may preclude some Mtase-DNA interactions with the flanking DNA. This in turn may have altered Mtase interactions with the minor groove of the canonical site.

ACKNOWLEDGEMENTS

We wish to thank Neda Mashhoon for purifying AdoMet, Chris Janecki for the design and manufacture of our temperature ramp generator, and Prof. T.C. Bruice for the use of his molecular graphics device. This work was supported by National Science Foundation Grant DMB 89–11404 to N.O.R.

REFERENCES