# *Eco*RI DNA Methyltransferase–DNA Interactions†

Norbert O. Reich* and Michael J. Danzitz, Jr.
*Department of Chemistry, University of California, Santa Barbara, California 93106*
*Received July 19, 1991; Revised Manuscript Received October 11, 1991*

ABSTRACT: We present a novel strategy with synthetic hemimethylated DNA substrates containing uracil for thymine and inosine for guanosine replacements and *Eco*RI DNA methyltransferase to characterize the importance of major groove hydrophobic groups to the sequence-specific modification of DNA. The bacterial Mtase uses *S*-adenosyl-L-methionine to methylate the double-stranded DNA site 5′GAATTC3′ at the N6 position of the central adenosine of each strand. Uracil substitution in either strand at the outer thymine (5′GAATUC3′) causes 2.2- and 1.7-fold *improvements* in specificity ($k_{cat}/K_m^{DNA}$). The fact that the specificity constant for the substrate containing uracil in both strands is identical to the value expected for noninteracting substitutions suggests that no significant methyltransferase–DNA interactions are altered beyond the site of either substitution. Similar analysis of the internal thymine (5′GAAUTC3′) also shows these methyl groups to make a negative contribution to specificity, although the observed nonadditivity with the doubly modified substrate clearly shows methyltransferase–DNA interactions beyond the site of substitution to be affected in this case. To further probe the effect of analogue incorporation on methyltransferase–DNA interactions beyond the site of substitution, the relatively "silent" and additive uracil changes (5′GAATUC3′) were combined with inosine for guanosine substitutions (e.g., 5′IAATTC3′) known to have significant negative effects on specificity. In contrast to the additivity observed with the outer thymines, these studies show significant changes in methyltransferase–DNA interactions caused by the removal of the thymine methyls. Our results implicate a complex and flexible methyltransferase–DNA interface in which subtle structural changes in the substrate are transmitted over the entire canonical site. Thermal stability analyses and determination of $\Delta H°$ and $\Delta S°$ for the double- to single-stranded transition for single and doubly substituted substrates show no additivity. This suggests that structural changes in the DNA alone may occur beyond the site of substitution. Interestingly, substrates with widely varying enthalpy and entropy terms show similar specificity with the Mtase, suggesting the Mtase is insensitive to the underlying conformational differences.

What is the molecular basis for the extraordinary specificity demonstrated by DNA binding and modifying proteins? Over the past decade and a half our understanding of these systems has increased dramatically, largely as a result of detailed structural studies (Kim et al., 1990; Jordan & Pabo, 1988; Anderson et al., 1987; Otinowski et al., 1988; Wolberger et al., 1988). Still, several issues obscure a mechanistic understanding of specificity: the quantitative contribution toward specificity deriving from individual protein–DNA interactions; the relative roles played by direct protein–DNA interactions (e.g., hydrogen bonds) versus detection of less direct, localized conformational variations (e.g., DNA-helical parameters such as twist); the contribution to enzymic discrimination of sites deriving from catalysis versus binding phenomena.

We are investigating the *Eco*RI DNA Mtase to better understand enzymic discrimination of DNA sequences. The Mtase differs from most previously investigated systems in that it is functional as a monomer, requires a cofactor for activity, and covalently modifies its DNA substrate. The Mtase, a 38 050-dalton enzyme, is part of the well-studied *Eco*RI restriction-modification system; the Mtase requires *S*-adenosyl-L-methionine (AdoMet) and its DNA substrate for activity (Rubin & Modrich, 1977). The methyl group of the AdoMet's sulfonium moiety is transferred to the N6 position of the central adenosine residue in the double-stranded DNA sequence 5′GAATTC3′. The Mtase is a highly efficient enzyme with a specificity constant ($k_{cat}/K_m^{DNA}$) for plasmid DNA of over $10^8$ $s^{-1}$ $M^{-1}$ (Reich & Mashhoon, 1991). The enzyme binds AdoMet and noncanonical DNA randomly, but recognition and methylation of the canonical sequence require AdoMet to be bound. Steps after methylation limit catalysis since the methyl-transfer step is at least 10 times faster than $k_{cat}$ (Reich & Mashhoon, 1991). Structural data exist for both the DNA substrate and methylated product, providing a good starting point for investigating the details of sequence-specific recognition. The X-ray structure of the product in which the central two adenosine residues are methylated at their N6 positions highlights the proximity of these methyl groups within the major groove and reveals minimal structural changes caused by the methylation (Frederick et al. 1988).

We describe a novel strategy to assess the importance to specificity of the major groove-residing thymine methyl groups present in the *Eco*RI recognition site. Previous studies have shown thymine methyl groups to be important for both the *Eco*RI endonuclease and Mtase (Brennan et al., 1986a,b) as well as other DNA-binding proteins (Caruthers, 1980; Anderson et al., 1987; Wharton & Ptashne, 1987; Ha et al., 1989). However, the studies with the *Eco*RI Mtase did not allow investigation of contributions toward specificity of individual methyl groups. Using comparative specificity analyses of singly and doubly modified substrates, we have determined the contributions toward specificity deriving from all four thymine methyl groups. Moreover, we have analyzed the validity of this approach with numerous structural and functional controls.

## Materials and Methods

*Materials.* Phosphoramidites and ancillary DNA synthesis reagents were obtained from Milligen/Biosearch and American Bionetics; β-cyanoethyl $N^6$-methyladenosine was obtained from Pharmacia. [γ-$^{32}$P]ATP (5000 Ci/mmol) was purchased

† This work was supported by National Science Foundation Grant DMB 89-11401 to N.O.R.
* To whom all correspondence should be addressed.

0006-2960/92/0431-1937$03.00/0 

from Amersham and [*methyl*-$^3$H]-AdoMet (78.5 Ci/mmol) 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 (50- × 7.8-mm) cation-exchange column (Bio-Rad) according to a modified, published procedure (Zappia et al., 1980). Ammonium acetate buffer (0.1 M, pH 4.0) was used to elute AdoMet at a flow rate of 1.0 mL/min, and the peak corresponding to AdoMet was collected, concentrated on a Bio-Rex 70 cation-exchange column (Bio-Rad), and stored at −20 °C in 0.1 M HCl (Reich & Everett, 1990). *Eco*RI Mtase was obtained from an overproducing strain of *Escherichia coli* and purified in our laboratory (Greene et al., 1978).

*Oligonucleotide Synthesis and Purification.* Oligonucleotides were synthesized on a Biosearch 3810 DNA synthesizer using β-cyanoethyl phosphoramidites. Oligonucleotides were purified on a 10-μm 4.6- × 250-mm Vydac C-18 column, eluted with a 15:85 to 40:60 acetonitrile/triethylammonium acetate, pH 7.0 (TEAAc), 20-min gradient. The 5′-(dimethoxytrityl) blocking groups were removed from the purified oligonucleotides with 80% acetic acid in water, and the DNA was precipitated with 2.5 M ammonium acetate and 75% ethanol and dissolved in 10 mM Tris–1 mM EDTA, pH 7.5 (TE).

The double-stranded DNA substrates listed in Tables I and II were made by annealing the two appropriate single strands using a 2-fold molar excess of the bottom strand to enhance the formation of the desired double-stranded substrates. All double-stranded oligonucleoides used for the melting analysis were annealed using equimolar amounts of the component strands. The mixture of single strands was heated to 95 °C and allowed to cool to room temperature over a 3–3$^1/_2$-h period.

*Oligonucleotide Characterization.* Oligonucleotide purity was confirmed by denaturing polyacrylamide gel electrophoresis of $^{32}$P-labeled oligonucleotides. Oligonucleotides were labeled by reaction with a 4-fold molar excess of [γ-$^{32}$P]ATP and 0.1 unit of T4 polynucleotide kinase per picomole of oligonucleotide using published reaction conditions (Maniatis et al., 1982). Samples were then electrophoresed through a 25% denaturing polyacrylamide sequencing gel, and an autoradiograph of the gel was used to estimate the homogeneity of each sample.

Double-stranded DNA formation was confirmed by annealing a radiolabeled oligonucleotide (top strand) to an unlabeled oligonucleotide (bottom strand) using the method described earlier. Samples were electrophoresed through a 14% nondenaturing polyacrylamide gel, and autoradiography was performed to determine the percentage of oligonucleotide present in double-stranded form.

*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 dictated by a homemade "temperature ramp generator" which was previously calibrated to produce a 0.25 °C/min ramp from 15 to 88 °C. The oligonucleotide solutions analyzed consisted of 10 mM phosphate, 0.1 mM EDTA, 20 mM NaCl, pH7.0, and double-stranded oligonucleotide at various concentrations (from 6 to 9 different concentrations) and were stirred throughout the analysis. The resulting $T_m$ values were determined using the sloping base line–sloping plateau method (Breslauer, 1986). The enthalpy and entropy terms were determined from the slopes and $y$-intercepts, respectively, of $R \ln (4/C)$ vs $1/T_m$ plots. Our regression analysis generated standard errors.

*Kinetic Analysis. Eco*RI Mtase activity was determined by an assay which monitors the transfer of the tritiated methyl group from [methyl-$^3$H]AdoMet to DNA (Reich & Mashhoon, 1990, 1991). Mtase reactions contained 200 μg/mL BSA, 100 mM Tris, pH 8.0, 5.0 mM EDTA, 5.0 mM DTT, 20 mM NaCl, and 0.2 nM Mtase with various concentrations of the cofactor [methyl-$^3$H]AdoMet (0.05, 0.10, 0.30, 1.0, and 3.0 μM at 10 000–50 000 cpm/pmol) in a volume of 45 μL. The methyl-transfer reactions were initiated by adding 5 μL of DNA substrate at the appropriate concentration (5 different DNA concentrations ranging from 0.5–10 times $K_m^{DNA}$). These reactions were allowed to proceed for 5–7 min at 37 °C and stopped by tranferring 40 μL of the reaction mixtures onto Whatman DE81 filter papers. In all cases product formation was linear as a function of time. The filter papers were washed three times with 50 mM $KH_2PO_4$ and one time each with 80% ethanol, 95% ethanol, and anhydrous ethyl ether for 20 min each (150 mL for each wash). The amount of radioactively labeled DNA bound to the filter papers was quantified by determining the tritium content with a Beckman LS-1701 liquid scintillation counter and correcting for signal loss caused by the filter papers (2.1-fold).

*Data Analysis.* All kinetic data were determined at least in duplicate and averaged. The raw cpm data were converted to velocity ($V$) in picomoles of methyl groups transferred per minute and fit to the Michaelis–Menten equation using a Basic version of Cleland's kinetics program (Cleland, 1979). Direct curve fitting was done with Cleland's HYPER equation to produce $1/V$ vs 1/[substrate] (1/[S]) plots which were used to determine $V_{max}$ vs 1/[S] replots, yielding true $k_{cat}$ and $K_m$ values. The results of replots for two independent assays were then averaged to produce the kinetic data in Tables I and II. Our regression analysis generated standard errors shown with all kinetic data; these were computed from the slope and intercept replots.

## Experimental Design

All substrates are non-self-complementary and methylated at the N6 position of the central adenosine residue of one strand. We previously determined that the methyl group on one strand has minimal effects on the specificity of the Mtase-catalyzed reaction. Hemimethylated substrates allow us to make single substitutions within the recognition site, thereby limiting the enzyme to a single productive binding orientation. These options provide more detailed information about the investigated interactions than was possible in the past. Possible structural alterations of the DNA substrate and Mtase–DNA–AdoMet complex resulting from the incorporation of modified bases are assessed with melting studies, "second site" analysis, and comparisons of $K_m^{AdoMet}$.

All oligonucleotides are 14 base pairs in length to assure stability of the double-stranded form at our assay conditions and to provide enough DNA flanking the recognition site to allow for any nonspecific contacts to be formed. Oligonucleotides were designed around the "Dickerson dodecamer" which has been characterized structurally with both NMR (Nerdal et al., 1989) and X-ray crystallographic (Wing et al., 1980) techniques.

## Results

Our study focuses on the functional significance of the thymine methyl moieties residing in the major groove of the

Table I: Steady-State Kinetic Parameters Determined for the Hemimethylated Control and Substrates Containing Deoxyuridine at the Outer Thymine Positions (U9)[a]

| abbreviation | sequence | $K_m$ DNA (nM) | $K_m$ AdoMet ($\mu$M) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m^{DNA}$ ($s^{-1}$ $M^{-1}$) |
|---|---|---|---|---|---|
| CBM | 5′ 3′<br>GGCGGAATTCGCGG<br>CCGCCTTAAGCGCC<br>* | 43.0 (3.03) | 0.33 (0.107) | 0.13 (0.006) | $3.0 \times 10^6$ |
| TU9-BM | GGCGGAATUCGCGG<br>CCGCCTTAAGCGCC<br>* | 25.4 (0.64) | 1.18 (0.67) | 0.17 (0.03) | $6.7 \times 10^6$ |
| T-BMU9 | GGCGGAATTCGCGG<br>CCGCCUTAAGCGCC<br>* | 63.3 (2.54) | 0.68 (0.65) | 0.33 (0.08) | $5.2 \times 10^6$ |
| TU9-BMU9 | GGCGGAATUCGCGG<br>CCGCCUTAAGCGCC<br>* | 13.3 (1.60) | 0.465 (0.090) | 0.16 (0.02) | $12 \times 10^6$ |

[a] Parameters were determined from double-reciprocal data; standard errors are shown in parentheses. (* denotes methylated adenine.)

Table II: Steady-State Kinetic Parameters Determined for the Hemimethylated Control and Substrates Containing Deoxyuridine at the Inner Thymine Positions (U8)[a]

| abbreviation | sequence | $K_m$ DNA (nM) | $K_m$ AdoMet ($\mu$M) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m^{DNA}$ ($s^{-1}$ $M^{-1}$) |
|---|---|---|---|---|---|
| CBM | 5′ 3′<br>GGCGGAATTCGCGG<br>CCGCCTTAAGCGCC<br>* | 43.0 (3.03) | 0.33 (0.107) | 0.13 (0.006) | $3.0 \times 10^6$ |
| TU8-BM | GGCGGAAUTCGCGG<br>CCGCCTTAAGCGCC<br>* | 7.89 (1.89) | 0.130 (0.004) | 0.060 (0.003) | $7.6 \times 10^6$ |
| T-BMU8 | GGCGGAATTCGCGG<br>CCGCCTUAAGCGCC<br>* | 28.0 (3.65) | 0.244 (0.051) | 0.14 (0.01) | $5.0 \times 10^6$ |
| TU8-BMU8 | GGCGGAAUTCGCGG<br>CCGCCTUAAGCGCC<br>* | 4.40 (1.16) | 0.409 (0.116) | 0.12 (0.013) | $28 \times 10^6$ |

[a] Parameters were determined from double-reciprocal data; standard errors are shown in parentheses. (* denotes methylated adenine.)

*Eco*RI Mtase recognition site. A related issue is the extent to which Mtase–DNA (methyl group) interactions are coupled to previously identified interactions involving the minor groove amino functionalities on guanosine and the Mtase (Reich & Danzitz, 1991). We define function as the catalytic turnover constant ($k_{cat}$), Michaelis constants ($K_m^{DNA}$ and $K_m^{AdoMet}$), and the specificity constant ($k_{cat}/K_m^{DNA}$) for each substrate (Fersht, 1985). The DNA substrates are shown in Tables I–III; they are all hemimethylated and 14 base pairs in length. The abbreviations are as follows: CBM, control bottom strand methylated at the internal adenine, as are all substrates; substitutions are numbered from the 5′ end, for example TU9-BM contains uracil within the top strand at the ninth position, bottom strand methylated; substitution of both strands is indicated as follows: TU9-BMU9, TU8-BMU8, TI5-BMU9, etc.

*Preliminary Characterization of DNA Substrates and Assay Conditions.* DNA substrates were shown to be greater than 98% pure by denaturing polyacrylamide gel electrophoresis. Annealings were shown to be at least 95% complete as determined by nondenaturing polyacrylamide gel electrophoresis of $^{32}P$-labeled substrates with a 2-fold excess of the unlabeled strand. Enzyme activity is first order with respect to enzyme concentration and was shown to be maximal at pH 7.8 and 10–20 mM NaCl. The presence of single strands was shown not to affect the steady-state parameters. Substrate concentration is in terms of canonical sites.

*DNA Melting Analysis.* Thermodynamic constants were derived from the temperature-dependent double-to single-strand transition. In all cases we observed a single spectroscopically detected transition, consistent with a two-state system, in contrast to the previously detected complex melting behavior of entirely self-complementary oligonucleotides (Breslauer et al., 1987). Each substrate was shown to be at least 90% double stranded in nature at our assay conditions. The dependence of melting temperature on oligonucleotide concentration was used to determine the entropy, enthalpy, and the resulting Gibbs free energy associated with the double-to single-stranded transition for each substrate (Table IV). All substrates are considerably more stable than our control substrate with the exception of TU9-BM, which is slightly less stable than CBM. Figure 3 shows that compensating changes in enthalpy and entropy are observed for all substrates relative to CBM.

*Kinetic Analysis.* As a result of our double-reciprocal analysis in which both AdoMet and DNA concentrations are varied, we were able to extract $K_m$, $k_{cat}$, and $k_{cat}/K_m$ values for DNA and AdoMet, shown in Tables I–III. The data in Table I show the Mtase is relatively insensitive to the absence of either methyl group normally present at the two outer thymine positions (U9 substrates). In both cases (TU9-BM and T-BMU9) the specificity constants are slightly increased relative to the control (CBM). However, the underlying contributions are different: for TU9-BM the increase derives from a small decrease in $K_m^{DNA}$, while for T-BMU9, $k_{cat}$ is increased 2.5-fold. The specificity constant for the doubly modified substrate (TU9-BMU9) is 4.0-fold greater than for CBM; this compares favorably with the 3.9-fold increase

Table III: Steady-State Kinetic Parameters Determined for the Hemimethylated Control (CBM), Substrates Containing Deoxyuridine at the Outer Thymine Positions (U9) in Combination with Inosine Replacement for Guanosine (I5), and Substrates Previously Characterized with Inosine Substitutions Alone (Reich & Danzitz, 1991)[a]

| abbreviation | sequence | $K_m$ DNA (nM) | $K_m$ AdoMet ($\mu$M) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m^{DNA}$ ($s^{-1}$ $M^{-1}$) |
|---|---|---|---|---|---|
| CBM | 5′ 3′<br>GGCGGAATTCGCGG<br>CCGCCTTAAGCGCC<br>* | 43.0 (3.03) | 0.33 (0.107) | 0.13 (0.006) | $3.0 \times 10^6$ |
| TI5-BM | GGCGIAATTCGCGG<br>CCGCCTTAAGCGCC<br>* | 554 (6.4) | 0.256 (0.062) | 0.13 (0.01) | $0.24 \times 10^6$ |
| T-BMI5 | GGCGGAATTCGCGG<br>CCGCCTTAAICGCC<br>* | 1430 (14) | 0.322 (0.018) | 0.11 (0.03) | $0.077 \times 10^6$ |
| TI5U9-BM | GGCGIAATUCGCGG<br>CCGCCTTAAGCGCC<br>* | 26.2 (6.1) | 0.44 (0.037) | 0.081 (0.009) | $3.1 \times 10^6$ |
| TI5-BMU9 | GGCGIAATTCGCGG<br>CCGCCUTAAGCGCC<br>* | 13.4 (3.8) | 0.25 (0.025) | 0.13 (0.004) | $9.7 \times 10^6$ |
| TU9-BMI5 | GGCGGAATUCGCGG<br>CCGCCTTAAICGCC<br>* | 213 (6.2) | 0.43 (0.16) | 0.135 (0.08) | $0.63 \times 10^6$ |
| T-BMI5U9 | GGCGGAATTCGCGG<br>CCGCCUTAAICGCC<br>* | 265 (8.5) | 0.21 (0.68) | 0.12 (0.01) | $0.45 \times 10^6$ |

[a] Parameters were determined from double-reciprocal data; standard errors are shown in parentheses. (* denotes methylated adenine.)

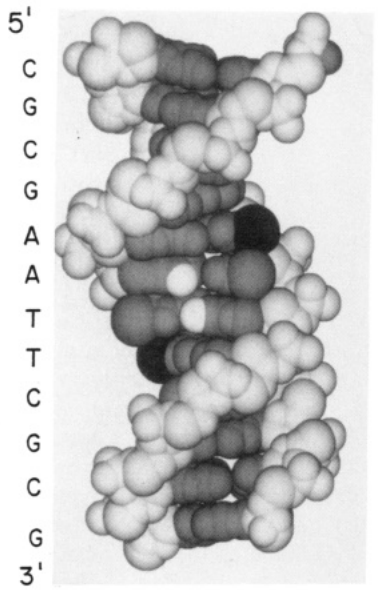


FIGURE 1: Graphical representation of the *Eco*RI recognition site containing "Dickerson dodecamer". The view is into the major groove of the recognition site, highlighting the C5 methyl groups of the outer thymine residues within the recognition site (U9). The methyl groups are black while the sugar phosphate backbone is light gray and the bases are dark gray.

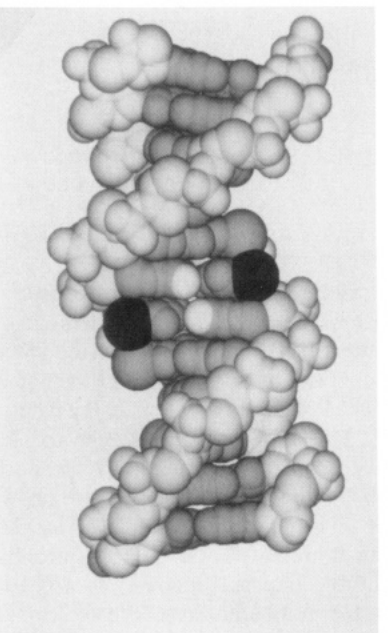

FIGURE 2: Graphical representation of the *Eco*RI recognition site containing "Dickerson dodecamer". The view is into the major groove of the recognition site, highlighting the C5 methyl groups of the inner thymine residues within the recognition site (U8). The methyl groups are black while the sugar phosphate backbone is light gray and the bases are dark gray.

predicted for the combined effect of the individual substitutions. Although associated with larger errors, the $K_m^{AdoMet}$ values for these U9 substrates are only slightly altered when compared to $K_m^{AdoMet}$ obtained with CBM.

Table II shows the results of removing the methyl group of either central thymine. The specificity constants of TU8-BM and T-BMU8 are increased 2.5- and 1.6-fold, respectively. The majority of this effect comes from the 5.4- and 1.5-fold decreases in $K_m^{DNA}$. As for the U9 substrates, comparison of $k_{cat}/K_m^{DNA}$ for the doubly substituted substrate can be used to assess the degree of additivity of individual substitutions (Carter et al., 1984). Unlike the U9 substrates, the predicted decrease relative to CBM for the sum of the individual changes (4.2-fold) is exceeded by the observed 9.2-fold decrease in TU8-BMU8, implicating structural changes beyond the site of substitution. In contrast, other than the 2.5-fold decrease in $K_m^{AdoMet}$ seen for TU8-BM, the minimal changes in this parameter suggest that AdoMet-binding determinants remain unchanged.

Table II shows the kinetic results when the uracil (U9) substitutions are combined with inosine substitutions (Reich & Danzitz, 1991). For comparison the effects of each inosine substitution alone (TI5-BM and T-BMI5) are shown (Reich

Table IV: Thermal Stability Data Determined for the Hemimethylated Control and All Deoxyuridine-Containing Oligonucleotides

| | $\Delta H°$ | $\Delta S°$ | $\Delta G°$ | $\Delta\Delta H°$ | $\Delta\Delta S°$ | $\Delta\Delta G°$ |
|---|---|---|---|---|---|---|
| CBM | 80.74 (1.8) | 227.8 (5.5) | 10.09 (3.5) | | | |
| TU9-BM | 53.37 (3.4) | 141.1 (10.6) | 9.62 (1.3) | 27.37 | 86.7 | 0.57 |
| T-BMU9 | 105.95 (3.8) | 302.9 (11.9) | 11.99 (0.9) | −25.21 | −75.1 | −1.90 |
| TU9-BMU9 | 97.71 (4.4) | 277.8 (13.7) | 11.53 (1.1) | −16.97 | −50.1 | −1.44 |
| TU8-BM | 107.84 (6.7) | 309.4 (20.9) | 11.89 (1.5) | −27.10 | −81.6 | −1.80 |
| T-BMU8 | 134.33 (6.3) | 393.8 (19.6) | 12.19 (1.2) | −53.59 | −166.0 | −2.10 |
| TU8-BMU8 | 143.31 (7.0) | 423.4 (21.9) | 12.00 (1.2) | −62.57 | −195.6 | −1.91 |
| TI5U9-BM | 90.12 (5.3) | 257.0 (18.2) | 10.41 (1.1) | −9.38 | −29.2 | −0.32 |
| TI5-BMU9 | 67.56 (4.1) | 184.9 (12.2) | 10.21 (0.9) | 13.18 | 42.9 | −0.12 |
| TU9-BMI5 | 116.50 (6.8) | 340.6 (21.0) | 10.85 (1.0) | −35.76 | −112.9 | −0.76 |
| T-BMI5U9 | 75.61 (2.6) | 211.3 (4.9) | 10.07 (1.1) | 5.13 | 16.5 | 0.02 |

[a] Units for $\Delta H°_{obs}$ and $\Delta\Delta H°_{obs}$ are kcal/mol, for $\Delta S°_{obs}$ and $\Delta\Delta S°_{obs}$, cal mol$^{-1}$ K$^{-1}$, and for $\Delta G°_{obs}$ and $\Delta\Delta G°_{obs}$, kcal/mol. All Gibbs free energy values were calculated from $\Delta H°_{obs}$ and $\Delta s°_{obs}$ at 310 K. Data were collected at pH 7.0 and 20 mM NaCl.

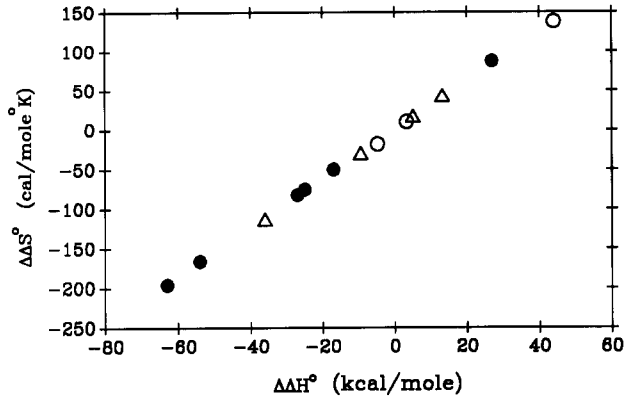


FIGURE 3: Change in entropy ($\Delta\Delta S°$) versus change in enthalpy ($\Delta\Delta H°$) associated with the double- to single-stranded transition for all oligonucleotides shown in Tables I–III relative to CBM. Filled circles are for uracil-substituted analogues (see Tables I and II), open circles are for inosine-substituted analogues (see Table III and footnote 1), and triangles are for mixed substitutions (inosine and uracil, Table III). All data were obtained from Table IV.

& Danzitz, 1991). The minor *improvement* in specificity with the U9 substrates relative to CBM contrasts with the 13- and 39-fold *decrease* in specificity observed with the inosine substrates. Introduction of a uracil (top or bottom strand) into either inosine-substituted substrate results in a significant improvement in specificity. This is seen most dramatically when the specificity constants for TI5-BM and TI5-BMU9 are compared; uracil incorporation into TI5-BM results in a 41-fold increase in specificity while the same uracil substitution into CBM (T-BMU9, Table I) causes only a 1.7-fold increase. Alternatively, starting from the two functionally similar U9 substrates, inosine incorporation into the bottom strand (TU9-BMI5 and T-BMI5U9) results in large *decreases* in specificity (11- and 15-fold, respectively) while inosine incorporation into the top strand (TI5U9-BM and TI5-BMU9) causes only a slight decrease (2.2-fold) and an *increase* (1.9-fold), respectively. The increase seen with TI5-BMU9 is particularly significant considering that the same inosine substitution into CBM results in a 13-fold decrease in specificity. The small changes in $K_m^{AdoMet}$ for TI5U9-BM and TI5-BMU9 relative to our previously tested substrates suggest that only minor structural changes in those functionalities critical for AdoMet binding have occurred.

## Discussion

We are attempting to characterize the contribution of individual functionalities toward DNA recognition by *Eco*RI DNA Mtase. Our novel strategy provides quantitative specificity data on DNA substrates lacking single functionalities implicated in DNA recognition (Reich & Danzitz, 1991). Prior application of this strategy to synthetic substrates containing inosine in place of guanine showed that (1) the Mtase is sensitive to removal of either guanine-associated amino group which resides in the minor groove of the canonical site and (2) these Mtase–DNA interactions are not additive, suggesting that sequence discrimination by *Eco*RI DNA Mtase may involve more complex mechanisms than those described for repressor–DNA interactions (Reich & Danzitz, 1991).

We sought to use this specificity analysis to determine the importance of the four methyl moieties located in the major groove (Figures 1 and 2). Hydrophobic interactions, and in particular protein–DNA interactions involving the thymine methyl group, are known to be critical for stabilizing sequence-specific protein–DNA complexes (Ha et al., 1989) and for sequence-specific recognition (Caruthers, 1980; Wharton & Ptashne, 1987; Anderson et al., 1987). However, past work has not led to a quantitative assessment of the contribution to specificity made by individual interactions involving thymine methyls. The importance of thymine methyl groups to *Eco*RI DNA Mtase specificity was previously investigated with synthetic oligonucleotides as well as phage DNA substrates. For example, Brennan et al. (1986b) showed that uracil replacement at the second thymine resulted in an inactive substrate. However, these studies were carried out with octanucleotides in which two thymines were substituted with uracil (Brennan et al., 1986b). As noted by Brennan et al. (1986b), the observed lack of activity may be due to the shortness of the oligonucleotide since a slightly longer substrate showed some activity. Mtase–DNA interactions with bases outside the recognition sequence may allow the enzyme to compensate for alterations in the substrate,[1] as appears to be the case with the endonuclease (Brennan et al., 1986a). Brennan et al. (1986b) also reported that substitution of uracil for both centrally located thymines reduced catalytic efficiency largely due to a reduction in $k_{cat}$. The phage investigations used DNA in which every thymine was replaced with uracil, making quantitative structure–function analysis difficult (Berkner & Folk, 1977). In contrast to these past studies, our strategy can be used to quantitatively assess the importance of putative Mtase–DNA and AdoMet–DNA interactions mediated through *individual* thymine methyls. Our additivity analysis with multiple analogue substitutions provides a check for isolated or interactive Mtase–DNA interactions.

*Effect of Deoxyuridine Substitutions at the Outer Thymine Positions (U9).* Substrates lacking either of the outer two methyl groups are methylated by the Mtase with an efficiency

[1] The Mtase interacts with at least one and up to two bases beyond the recognition hexanucleotide by DNA-footprinting analyses (N. O. Reich, and M. Danzitz, unpublished observations).

similar to that of the control (CBM, see Table I). This suggests that only minor Mtase–DNA interactions occur at either altered site. Although the specificity constants are slightly increased relative to CBM, the individual contributions ($K_m^{DNA}$ and $k_{cat}$) are different for TU9-BM and T-BMU9. The small changes in kinetic values (U9-substituted substrates) are nevertheless positionally related, consistent with a monomeric Mtase interacting with the hemimethylated substrate in a single functional binding orientation.

Our results contrast with the previous findings by Brennan et al. (1986b) using a shorter, nonmethylated substrate containing both U9 substitutions (see TU9-BMU9, discussed below). However, the work by Brennan et al. (1986b) was done at 10 °C, the data were collected at a single concentration of AdoMet, and as mentioned earlier, the octanucleotide substrates may have lacked stabilizing nonspecific Mtase–DNA interactions.[1] These differences make any direct comparison obscure.

A critical issue in our structure–function analysis is whether individual structural changes in the substrate perturb the DNA conformation beyond the site of substitution. Past structure–function analyses of DNA Mtases (Berkner & Folk, 1979; Brennan et al., 1986b; Modrich & Rubin, 1977; Newman et al., 1990a,b) and restriction endonucleases (Berkner & Folk, 1979; Brennan et al., 1986a; Diekmann & McLaughlin, 1988; Fliess et al., 1986; McLaughlin et al., 1987; Seela & Kehne, 1987; Bodnar et al., 1983; Wolfes et al., 1986; Newman et al., 1990a,b; Cosstick et al., 1990; Lesser et al., 1990) using base analogues have not addressed this issue. Therefore, in addition to comparisons of $K_m^{AdoMet}$, we determined if individual substitutions result in additive effects in $k_{cat}/K_m^{DNA}$. This powerful probe of conformational changes has been used in mutant analyses (Carter et al., 1984) but has been overlooked in studies of protein–DNA interactions and analogue substitutions. In contrast to our findings with inosine-substituted substrates (Reich & Danzitz, 1991), the doubly substituted substrate (TU9-BMU9) shows the individual uracil substitutions (TU9-BM and T-BMU9) result in additive effects. Additivity in this context implies that any Mtase–DNA interactions removed (or altered) by analogue (e.g., uracil) replacement do not result in altered Mtase–DNA interactions at the reciprocal position within the substrate. Also, this suggests that the small specificity enhancements observed with TU9-BM and T-BMU9 are real and derive from the removal of localized *negative* interactions between the Mtase and DNA.

*Effect of Deoxyuridine Substitutions at the Central Thymine Positions (U8).* Substrates lacking either of the two central methyl groups, TU8-BM and T-BMU8 (Table II), are methylated more efficiently than CBM by a factor of 2.5- and 1.6-fold, respectively. The changes in $K_m^{DNA}$ are larger, but the corresponding changes in $k_{cat}$ result in lower specificity differences. Again, the increased specificities suggest that negative interactions between the Mtase and DNA have been removed. However, analysis of the doubly substituted substrate (TU8-BMU8) shows these effects to be nonadditive since the predicted gain in specificity is 4.1-fold and TU8-BMU8 has a specificity constant 9.2-fold greater than CBM (Table II). This suggests that methyl group removal from either internal thymine alters Mtase–DNA interactions beyond the site of substitution and obscures quantitative interpretation by singly modified substrates. Further evidence for conformational changes beyond the site of substitution comes from comparison of $K_m^{AdoMet}$ for the substrates in Table II. The 2.5-fold enhanced affinity for AdoMet observed with TU8-BM versus

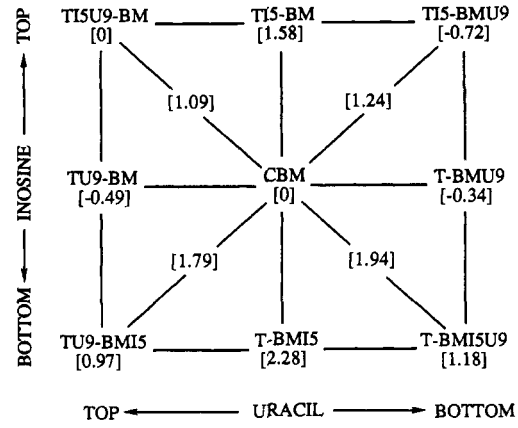


FIGURE 4: Change in free energy associated with differences in $k_{cat}/K_m^{DNA}$ of indicated substrates relative to CBM. Values in brackets (kcal/mol) refer to the following: $\Delta\Delta G° = -RT \ln [(k_{cat}/K_m^{analogue})/(k_{cat}/K_m^{CBM})]$. The $k_{cat}/K_m^{DNA}$ values were obtained from Tables I–III. A positive $\Delta\Delta G°$ reflects a loss in specificity in going from CBM to the analogue of interest. Values on diagonals are the predicted effects (sum) resulting from the observed individual substitutions.

CBM suggests either that the cofactor interacts directly with the thymine methyl or that Mtase–DNA interactions in this region also form part of the AdoMet-binding site. Given the proximity of the exocyclic amine of the target adenine to the thymine methyl group (Figure 2), either of these explanations seems plausible.

*Effect of Combining "Silent" Changes at U9 with Previously Characterized Inosine Substitutions.* As a further probe of conformational changes beyond the site of substitution, we combined relatively silent changes (U9) with inosine substitutions shown to have significant negative effects on specificity (Table III) (Reich & Danzitz, 1991). The four substrates (Table III) allow characterization of changes in Mtase–DNA interactions resulting from thymine–uracil replacements beyond those investigated with our previous additivity analyses (Tables I and II) (Reich & Danzitz, 1991). The observed minimal changes in $K_m^{AdoMet}$ again suggest that the Mtase–DNA–AdoMet complexes undergo only limited conformational changes with the four doubly substituted substrates. The free energy differences in $k_{cat}/K_m^{DNA}$ relative to CBM are schematically shown in Figure 4. The predicted changes in free energy which derive from combining two substitutions are listed on the diagonals (Carter et al., 1984); in no case are the observed changes additive. The data representation used in Figure 4 provides facile assessment of analogue-induced alterations in Mtase–DNA interactions beyond the site of substitution. For example, the two analogues and CBM along the horizontal axis (e.g., TU9-BM, CBM, and T-BMU9) can be compared to the set of three analogues above or below (e.g., TI5U9-BM, TI5-BM, and TI5-BMU9) to determine the effect of uracil for thymine replacement (U9) in the bottom or top strands. Theoretically, if removal of the thymine methyl in either strand left all other Mtase–DNA unchanged, then the *difference* in free energy change (e.g., TU9-BM to TI5U9-BM; CBM to TI5-BM; T-BMU9 to TI5-BMU9) would be the same for each pair. Hence, removal of the thymine methyl (U9) clearly alters Mtase–DNA interactions in the minor groove since inosine incorporation (top strand) into TU9-BM, CBM, and T-BMU9 results in a wide range of free energy changes (0.49, 1.58, and −0.38 kcal/ mol, respectively). The related values for inosine incorporation into the bottom strand (1.46, 2.28, and 1.52 kcal/mol) vary less widely and show a significantly greater average effect (1.76 versus 0.56 kcal/mol for the top strand). Thus, the difference between these two

(1.20 kcal/mol) is greater than the analogous difference observed between TI5-BM (1.58 kcal/mol) and T-BMI5 (2.28 kcal/mol) relative to CBM. As described below, this may derive from differences in the sequence context between the top and bottom strand substitutions.

Our previous additivity analysis involving inosine replacements showed that this analogue altered Mtase–DNA interactions at least five bases removed from the site of substitution (Reich & Danzitz, 1991). The results in Table III and Figure 4 clearly support our previous findings. For example, the effect of replacing the thymine on the bottom strand with uracil is significantly different when the top strand contains inosine (TI5-BM, −2.30 kcal/mol), when no prior substitution has been made (CBM, −0.34 kcal/mol), and when the bottom strand contains inosine (T-BMI5, −1.10 kcal/mol). The related comparison involving thymine replacement of the top strand shows the same trend. These results not only confirm our earlier results, but also demonstrate that Mtase–DNA interactions are altered at various positions along the DNA as a result of inosine substitution.

The generally observed "recovery" of decreased specificity observed with both inosine-containing substrates upon thymine replacement with uracil is particularly obvious with TI5-BMU9 when compared with TI5-BM. Why the removal of the thymine methyl should cause a gain of 2.3 kcal/mol in $k_{cat}/k_m^{DNA}$ for TI5-BM and only 0.34 kcal/mol for CBM is unclear. Moreover, the fact that TI5-BMU9 is a *better* substrate than CBM and T-BMU9 suggests that removal of the guanosine minor groove amino group actually enhances specificity in this case. Clearly, Mtase–DNA interactions in the vicinity of the guanosine minor groove have been altered as a result of uracil replacement for thymine.

Unlike the relatively equivalent effects on specificity seen when comparing previous top and bottom strand substitutions (Table I and II) (Reich & Danzitz, 1991), the substrates described in Table III and Figure 4 clearly show differential effects with respect to which strand is modified. Thus, relative to CBM, TI5-BM and T-BMI5 have free energy terms which are 1.58 and 2.28 kcal/mol lower (Figure 4), corresponding to a difference of 0.70 kcal/mol. In contrast, the differences for TI5U9-BM versus TU9-BMI5 (0.97 kcal/mol) and TI5-BMU9 and T-BMI5U9 (1.90 kcal/mol) are significantly larger. The difference between TI5-BMU9 and T-BMI5U9 is particularly large and may be related to our demonstration that replacement of guanosine with inosine in the bottom strand dramatically alters the enthalpic and entropic contributions to DNA stability, while only minor changes are detected when the top strand is altered (Reich & Danzitz, 1991). We previously proposed this differential behavior derives from the bottom strand substitution occurring within the trinucleotide sequence 5′CGA3′, while in the bottom strand the sequence 5′GGA3′ is involved (site of inosine replacement is underlined). As described below, the 5′CGA3′ sequence is structurally anomalous.

Although our use of multiple analogue substitutions to assess changes in enzyme–DNA interactions is novel, as stated earlier, a similar strategy applied to mutant enzymes led to quantitative insights into whether changes in conformation occur beyond the site of substitution (Carter et al., 1984). Our additivity analyses with one (Tables I and II) and two (Table III, Figure 4) types of base substitution clearly show extensive "communication" between single sites of Mtase–DNA interaction, consistent with a flexible and complex Mtase–DNA interface. Since previous application of a related approach to DNA-binding proteins involved multiple switches of entire base pairs (Sarai & Takeda, 1989; Lehming et al., 1990; Takeda et al., 1989; Ebright et al., 1987) and largely showed additivity, it is unclear whether Mtase–DNA interactions are fundamentally different from these previously characterized systems or whether similar interactions might be found if these systems where submitted to our analysis. In any event, our results suggest that quantitative interpretation of such data is simplistic without an understanding of any communication *between* enzyme–ligand interactions (Carter et al., 1984).

*Thermodynamic Analysis of DNA Substrates.* The thermal stabilities listed in Table IV demonstrate that, with the exception of TU9-BM, all substrates analyzed in Tables I–III are more stable than the unmodified CBM and are largely double stranded under our assay conditions. In all cases the melting behavior showed a cooperative thermal transition, supporting a quantitative interpretation of the derived thermodynamic parameters (see below) (Breslauer, 1986; Borer et al., 1974; Aboul-ela et al., 1985). As part of our effort to understand the structural basis for the specificity differences described above, we determined the enthalpic and entropic contributions toward the free energy term associated with the double to single strand transition (Table IV) (Breslauer, 1986; Borer et al., 1974; Aboul-ela et al., 1985). Enthalpy and entropy analysis of ligand–DNA (Breslauer et al., 1987; Marky & Breslauer, 1987), enzyme–DNA (Petruska et al., 1988), and protein–protein interactions (Brandts & Hunt, 1967; Shortle et al., 1988) have contributed to an understanding of the forces involved in these systems. Therefore, although unable to identify the structural origins of any differences in thermal stability, our thermodynamic analysis is used to implicate conformational differences in the free DNA substrates and associated waters.

In addition to being unique in its relative destabilization, TU9-BM has significantly altered enthalpic and entropic components which make up the free energy term for the double- to single-stranded transition. The decreases in enthalpy and entropy in comparison to CBM are dramatic, particularly when compared to the common, and frequently large, increases observed with the remaining substrates (e.g., the symmetry-related T-BMU9). Variations in enthalpy and entropy of this magnitude suggest significant changes in DNA conformation (Freier et al., 1983; Aboul-ela et al., 1985). A plausible explanation for values shown for TU9-BM may be that the 5′CGA3′/5′TCG3′ portion of the oligonucleotide has been altered to 5′CGA3′/5′UCG3′. Various analyses of the sequence-related "Dickerson" dodecamer (Wing et al., 1980) have shown the 5′CGA3′/5′TCG3′ segment to have unusual helical parameters. DNase digestion (Lomonossoff et al., 1981), NMR analysis (Nerdal et al., 1989; Chary et al., 1987), and X-ray crystallography (Wing et al., 1980; Dickerson & Drew, 1981; Frederick et al., 1988) all suggest that the helical twist in this region varies greatly. We previously showed that inosine substitution in this region also caused aberrant melting behavior while an identical substitution in the 5′GGA3′/5′TCC3′ region did not show unusual effects (Reich & Danzitz, 1991). Although the structural origins for the unusual alterations in stability remain to be determined,[2] changes in the helical parameters of the DNA (e.g., helical twist) and/or interactions with associated waters seem highly plausible. Particularly interesting are the cases in which substrates with similar functional characteristics (Tables I–III) show large differences in the thermodynamic parameters (e.g., TU9-BM,

[2] The crystal structures of several substrates described in this report are being characterized (N. O. Reich, and C. A. Frederick, unpublished observations).

T-BMU9, and T-BMU8 in Table IV). Even in the absence of detailed structural information these results suggest that the specificity of the Mtase-catalyzed reaction is relatively insensitive to the underlying conformational differences in the various DNA molecules leading to the changes in thermodynamic parameters.

Application of the additivity analysis described for our functional comparisons (Table I–III) to the data in Table IV clearly shows the enthalpy and entropy terms to be nonadditive. A similar additivity analysis of thermodynamic parameters has previously been used to assess long-range conformational changes in mutant proteins (Shortle et al., 1988). Our data suggest that, in the absence of the Mtase, some form of structural communication occurs between sites of modification within the DNA. A contributing factor to the nonadditivity observed with our functional analysis (Tables I–III and Figure 4) may therefore be conformational changes in the DNA itself.

Since the large variations in enthalpy and entropy shown in Table IV do not lead to significant free energy changes, compensation must be occurring. This is clearly evident in Figure 3 showing the changes in enthalpy and entropy relative to CBM for all of the substrates listed in Table IV. This remarkably good correlation, which spans changes in enthalpy of over 100 kcal/mol and entropy changes of 350 cal $mol^{-1}$ $K^{-1}$, is similar to trends observed in drug–DNA interactions (Breslauer et al., 1987; Marky & Breslauer, 1987), enzyme–DNA interactions (Petruska et al., 1988), and protein stability studies (Shortle et al., 1988). A common feature in the interpretation of these studies is the importance of water, and, in particular, water molecules directly associated with the macromolecule (Breslauer et al., 1987; Marky & Breslauer, 1987; Petruska et al., 1988; Shortle et al., 1988). Lumry and Rajender (1970) describe two criteria to determine if such compensation is real and if water plays a significant role: (1) the linear dependence of enthalpy and entropy changes over a significant range and (2) a "compensation temperature" ($T_c$, obtained from the slope of such plots) of between 250 and 320 K. A critical role for water in the compensation behavior observed in our system is therefore supported by a $T_c$ value of 320 K (derived from Figure 3), the ranges of enthalpy and entropy changes are greater than those cited by Lumry and Rajender (1970), and the linearity of the data in Figure 3 is excellent. Unfortunately, the precise nature of that role remains uncertain, although water-mediated stabilization of DNA structure via minor groove interactions (Drew & Dickerson, 1981) and phosphate oxygens (Saenger et al., 1986) has been proposed. As stated earlier, what is remarkable about our results is the limited functional consequence to the Mtase specificity resulting from large differences in the forces responsible for maintaining DNA structure.

## Acknowledgments

We thank Neda Mashhoon for purifying AdoMet, Chris Jenecki for the design and manufacture of our temperature ramp generator, and Prof. T. C. Bruice for the use of his molecular graphics facility.

## References

Aboul-ela, F., Koh, D., Tinoco, I. Jr., & Martin, F. H. (1985) *Nucleic Acids Res. 13*, 4811–4824.

Anderson, J. E., Ptashne, M., & Harrison, S. C. (1987) *Nature 327*, 846–852.

Berkner, K. L., & Folk, W. R. (1977) *J. Biol. Chem. 252*, 3185–3193.

Borer, P. N., Dengler, B., Tinoco, I., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843–853.

Brandts, J. F., & Hunt, L. (1967) *J. Am. Chem. Soc. 89*, 4826–4838.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986a) *J. Biol. Chem. 161*, 7270–7278.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986b) *J. Biol. Chem. 161*, 7279–7286.

Breslauer, K. J. (1986) in *Thermodynamic Data for Biochemistry and Biotechnology* (Hinz, H.-J., Ed.) pp 402–427, Springer-Verlag, Berlin.

Breslauer, K. J., Remeta, D. P., Chou, W.-Y. C., Ferrante, R., Curry, J., Zaunczkowski, D., Snyder, J. G., & Marky, L. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8922–8926.

Carter, P. J., Winter, G., Wilkinson, A. J., & Fersht, A. R. (1984) *Cell 38*, 835–840.

Caruthers, M. H. (1980) *Acc. Chem. Res. 13*, 155–160.

Chary, K. V. R., Hosur, R. V., Govil, G., Zu-kun, T., & Miles, H. T. (1987) *Biochemistry 26*, 1315–1322.

Cleland, W. W. (1979) *Methods Enzymol. 66*, 103–117.

Cosstick, R., Li, X., Tuli, D. K., Williams, D. M., Connolly, B. A., & Newman, P. C. (1990) *Nucleic Acids Res. 18*, 4771–4778.

Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.

Diekmann, S., & McLaughlin, L. W. (1988) *J. Mol. Biol. 202*, 823–834.

Drew, H. R., & Dickerson, R. E. (1981) *J. Mol. Biol. 151*, 535–556.

Ebright, R. H., Kolb, A., Buc, H., Kunkel, T. A., Krakow, J. S., & Beckwith, J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6083–6087.

Fersht, A. (1985) *Enzyme Structure and Mechanism*, W. H. Freeman and Co., New York.

Fliess, A., Wolfes, H., Rosenthal, A., Schwellnus, K., Blocker, H., Frank, R., & Pingoud, A. (1986) *Nucleic Acids Res. 14*, 3463–3474.

Frederick, C. A., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., & Rich, A. (1988) *J. Biol. Chem. 263*, 17872–17879.

Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., & Turner, D. H. (1983) *Biochemistry 22*, 6198–6206.

Greene, P. J., Heyneher, H. L., Bolivar, F., Rodriguez, R. L., Betlach, M. C., Covarrubias, A. A., Backman, K., Russel, D. J., Tait, R., & Boyer, H. W. (1978) *Nucleic Acids Res. 5*, 2373–2380.

Ha, J.-H., Spolar, R. S., & Record, M. T., Jr. (1989) *J. Mol. Biol. 209*, 801–816.

Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893–907.

Kim, Y., Grable, J. C., Love, R., Greene, P. J., & Rosenberg, J. M. (1990) *Science 249*, 1307–1309.

Lehming, N., Sartorius, J., Kisters-Woike, B., Wilcken-Bergmann, B., & Müller-Hill, B. (1990) *EMBO J. 9* (3), 615–621.

Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776–786.

Lomonossoff, G. P., Butler, P. J. G., & Klug, A. (1981) *J. Mol. Biol. 149*, 745–760.

Lumry, R., & Rajender, S. (1970) *Biopolymers 9*, 1125–1227.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marky, L. A., & Breslauer, K. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4359–4363.

McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtissek, S. (1987) *Biochemistry 26*, 7238-7245.

Modrich, P., & Rubin, R. A. (1977) *J. Biol. Chem. 252*, 7273–7278.
Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008–10021.
Newman, P. C., Nwosu, V. U., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990a) *Biochemistry 29*, 9891–9901.
Newman, P. C., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990b) *Biochemistry 29*, 9902–9909.
Otinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., *Marmorstein*, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 355*, 321–329.
Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., & Tinoco, I. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6252–6256.
Reich, N. O., & Everett, E. (1990) *J. Biol. Chem. 265*, 8929–8934.
Reich, N. O., & Mashhoon, N. (1990) *J. Biol. Chem. 265*, 8966–8970.
Reich, N. O., & Danzitz, M. (1991) *Nucleic Acids Res. 19* (23), 6587–6594.
Reich, N. O., & Mashhoon, N. (1991) *Biochemistry 30*, 2933–2939.
Rubin, R. A., & Modrich, P. (1977) *J. Biol. Chem. 252*, 7265–7272.
Saenger, W., Hunter, W. N., & Kennard, O. (1986) *Nature 324*, 385–388.
Seela, F., & Kehne, A. (1987) *Biochemistry 26*, 2232–2238.
Shortle, D., Meeker, A. K., & Freire, E. (1988) *Biochemistry 27*, 4761–4768.
Takeda, Y., Sarai, A., & Rivera, V. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 439–443.
Wharton, R. P., & Ptashne, M. (1987) *Nature 326*, 888–891.
Wing, R., Drew, H. R., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature 287*, 755–758.
Wolberger, C., Dong, Y., Ptashne, M., & Harrison, S. C. (1988) *Nature 335*, 789–795.
Wolfes, H., Alves, J., Fliess, A., Geiger, R., & Pingoud, A. (1986) *Nucleic Acids Res. 14*, 9063–9080.
Zappia, V., Galletti, P., Porcelli, M., Manna, C., & Ragione, F. D. (1980) *J. Chromatogr. 189*, 399–405.

# Structural Basis of Human Erythrocyte Glucose Transporter Function: pH Effects on Intrinsic Fluorescence†

James J. Chin, Byung H. Jhun, and Chan Y. Jung*
*Department of Biophysical Sciences, State University of New York at Buffalo, and the Biophysical Laboratory, Veterans Administration Medical Center, Buffalo, New York 14215*
*Received August 20, 1991; Revised Manuscript Received December 4, 1991*

ABSTRACT: The effects of pH on the intrinsic fluorescence of purified human erythrocyte glucose transporter (HEGT) were studied to deduce the structure and the ligand-induced dynamics of this protein. D-Glucose increases tryptophan fluorescence of HEGT at a 320-nm peak with a concomitant reduction in a 350-nm peak, suggesting that glucose shifts a tryptophan residue from a polar to a nonpolar environment. Cytochalasin B or forskolin, on the other hand, only produces a reduction at the 350-nm peak. The pH titration of the intrinsic fluorescence of HEGT revealed that at least two tryptophan residues are quenched, one with a $pK_a$ of 5.5, the other with a $pK_a$ of 8.2, indicating involvement of histidine and cysteine protonation, respectively. D-Glucose abolishes both of these quenchings. Cytochalasin B or forskolin, on the other hand, abolishes the histidine quenching but not the cysteine quenching and induces a new pH quenching with a $pK_a$ of about 4, implicating involvement of a carboxyl group. These results, together with the known primary structure and the transmembrane disposition of this protein, predict the dynamic interactions between Trp388 and His337, Trp412 and Cys347, and Trp412 and Glu380, depending on liganded state of HEGT, and suggest the importance of the transmembrane helices 9, 10, and 11 in transport function.

A family of specific transmembrane proteins facilitates the movement of glucose across the plasma membrane in animal cells (Mueckler et al., 1985; Birmbaum et al., 1986; Charron et al., 1989; James et al., 1989). Four tissue-specific isoforms are identified in this family, including the erythrocyte type (GLUT-1), liver and pancreatic $\beta$-cell type (GLUT-2), and muscle and fat cell type (GLUT-4) (Bell et al., 1990). All of these isoforms have 12 transmembrane $\alpha$-helices (TMH 1–12, numbered in sequence from the N-terminal to the C-terminal ends) of highly conserved amino acid sequences (Bell et al., 1990) thought to form a transmembrane aqueous channel for glucose (Mueckler et al., 1985; Jung et al., 1986; Chin et al., 1986). These isoforms also show similar non-membrane domain structures, which may play important roles in tissue–specific regulations of transporter function (Bell et al., 1990).

GLUT-1 of human erythrocytes is the only isoform that is available as a purified and functional protein (HEGT)[1] in this family (Kasahara & Hinkle, 1977; Baldwin et al. 1982; Rampal et al., 1985). HEGT has been used extensively for the biochemical and biophysical characterization of this iso-

† This work was supported in part by National Institute of Health Grant DK13376 and by the Veterans Administration Medical Center, Buffalo, NY.
* Address correspondence to this author at the Veterans Administration Medical Center.

[1] Abbreviations: HEGT, human erythrocyte glucose transporter; TMH, transmembrane $\alpha$ helix.

0006-2960/92/0431-1945$03.00/0 