Dynamic Impact of Methylation at the M. *Hha*I Target Site: A Solid-State Deuterium NMR Study†

Gary A. Meints and Gary P. Drobny*

Departments of Chemistry and Physics, University of Washington, Seattle, Washington 98195

Received February 6, 2001; Revised Manuscript Received June 14, 2001

ABSTRACT: Base methylation plays an important role in numerous biological functions of DNA, from inhibition of cleavage by endonucleases to inhibition of transcription factor binding. Studies of nucleic acid structure have shown little differences in unmethylated DNAs and the identical sequence containing methylated analogues. We have investigated changes in the local dynamics of DNA upon substitution of a methylated cytosine analogue for cytosine using solid-state deuterium NMR. In particular, we have observed changes in the local dynamics at the target site of the M. *Hha*I restriction system. These studies observe changes in the amplitudes of the local backbone dynamics at the actual target site of the *Hha*I methyltransferase. This conclusion is another indication that the significant result of base methylation is to perturb the local dynamics, and therefore the local conformational flexibility, of the DNA helix, inhibiting or restricting the protein's ability to manipulate the DNA helix in order to perform its chemical alterations.

Methylation of the heterocyclic bases in DNA[1] has diverse effects on its biological function, depending on the specific system in question. Many transcription factors no longer efficiently bind DNA when their recognition sites are methylated (*1*−*3*). Gene silencing in eukaryotes is thought to result from methylation either directly by interfering with the interaction of a transcription factor with its recognition site, or indirectly by attracting proteins that have a high affinity for methylated DNA (*4*, *5*). For example, methylation-induced structural and/or dynamical changes within a triplet of CpG dinucleotides have been shown to directly contribute to translational positioning of nucleosomes by affecting binding of the histone octamer (*6*). Methylated CpG dinucleotides are also sites for a high percentage of point mutations (*7*), and abnormal methylation patterns in DNA have been linked to various cancers and to genetic disorders such as fragile X syndrome (*1*, *2*, *8*). In prokaryotic cells, methylation is typically used in systems of restriction and modification as a defense mechanism against extracellular infection. It now becomes necessary to determine the relationship between these functional differences and any structural or dynamic changes between methylated and unmethylated DNA.

No clear relationship exists between the functional impact of base methylation in DNA and any structural alteration. Although cytosine methylation increases the thermal stability of the DNA and gel electrophoresis assays show CpG methylation alters DNA bending slightly (*9*), X-ray diffraction (*5*, *10*, *11*), solution NMR (*12*−*14*), and solid-state NMR studies (*12*) of singly CpG-methylated DNA oligomers indicate minimal structural alteration. However, in a crystallography investigation by Vargason et al., the DNA sequence $d(GGCGCC)_2$ exhibits traditional B-form characteristics, whereas $d(GGCGm^5CC)_2$, where m^5C is a 5-methylcytosine, forms a new conformation called E-DNA, which contains structural characteristics of both A-DNA and B-DNA (*15*). Experiments examining sensitivity to DNase I conclude that methylation has only a small effect on DNA structure in the *Hha*I binding site (*16*).

However, certain CpG steps have been shown to display unique dynamic properties that are affected by methylation. Hatcher et al. have shown large amplitude dynamics at the 2″ position in the sugar ring and the 5′ backbone methylene group in the dC (C*) at a CpG step in the *Eco*RI binding site (underlined) in the DNA sequence [d-(CGCGAATTC*GCG)]$_2$ (*17*), henceforth known as the Dickerson dodecamer for the group who originally crystallized the sequence in B-form (*18*). A solution NMR study of a similar sequence indicated the C* furanose ring was puckering between two conformations (*19*). Additional work by Hatcher et al. showed large amplitude motion in the dC furanose ring immediately adjacent to the *Eco*RI binding site in the same DNA sequence (C**), [d-(CGC**GAATTCGCG)]$_2$, which is also at a CpG step (*20*). The dynamics at this site, not observed in the solution study by Aramini and co-workers (*19*), have been determined to also be a repuckering of the cytidine furanose ring between conformations, with a puckering amplitude of ∼0.4 Å (*21*). Large amplitude motions have also been observed by solid-state deuterium NMR to be localized to deoxycytidines of the *Hha*I binding site, -GCGC- (*22*). All of these motions are present at CpG steps or within a restriction site. Also, experimental measurements of localized dynamics at CpG dinucleotides are supported by molecular dynamics and energy minimization studies that indicate that CpG steps are

† This research was supported by NIH Grant RO1 GM58914-01. G.A.M. acknowledges the support from NIH Training Grant GM32681.

* To whom correspondence should be addressed at the Department of Chemistry, University of Washington, Box 351700, Seattle, WA 98195. E-mail: drobny@macmail.chem.washington.edu. Fax: (206) 685-8665.

[1] Abbreviations: NMR, nuclear magnetic resonance; DNA, deoxyribonucleic acid; dC, 2′-deoxycytidine; T, thymidine; ^2H, deuterium.

10.1021/bi0102555 CCC: $20.00
Published on Web 09/21/2001

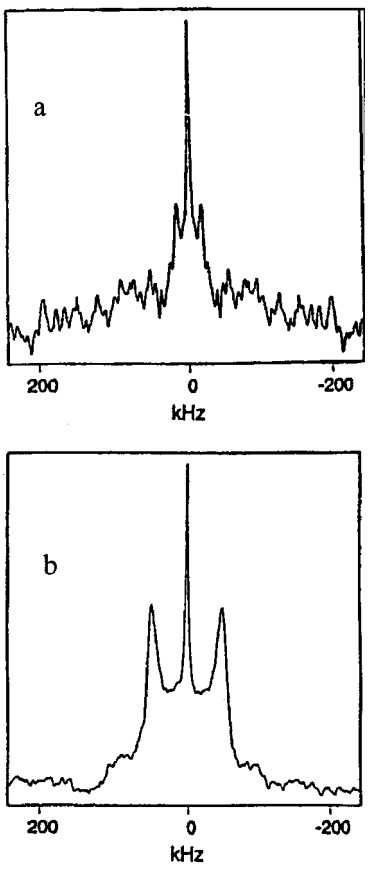

FIGURE 1: Comparison of line shape spectra for (a) [5′/5″-^{2}H]-C9 from the Dickerson sequence [d(CGCGAATTCGCG)]$_2$ (adapted from ref *17*) and (b) [5′/5″-^{2}H]-5-methyl-C9 from the Dickerson sequence (adapted from ref *28*). The central isotopic component is due to residual HDO.

more malleable than other steps (*23*). This notion is sustained by solution NMR studies, which have observed structural variations within DNAs containing CpG steps and conclude that there is an increased malleability at CpG steps (*24*−*27*).

There is recent evidence that methylation has an effect on the internal dynamics of CpG dinucleotides in DNA. In the *Eco*RI binding site, -GAATTC-, it was observed that the large amplitude librations of the 5′-methylene group in the dC were effectively quenched when 5-methylcytosine is substituted for cytosine (*28*) (see Figure 1). Methylation at this position, albeit not the target of the methyltransferase, inhibited cleavage by the *Eco*RI endonuclease (*29*). Overall, these observations suggest a direct link between suppression of the amplitudes of localized, internal motions of the sugar−phosphate backbone of the DNA and inhibition of restriction enzyme cleavage, and base methylation may play a role in suppression of these local motions. The purpose of this current study is to investigate whether methylation has a similar effect on the internal dynamics in the *Hha*I binding sequence, in particular at the target site of the methyltransferase.

The *Hha*I restriction endonuclease and associated methyltransferases bind to the sequence 5′-G↓CGC-3′, where the arrow indicates the cutting site and the underlined dC is the target for the methyltransferase. This sequence has been investigated in great detail but primarily to elucidate the methylation mechanism, whose reaction mechanism was first solved by Wu and Santi (*30*). The M. *Hha*I methyltransferase is a cytosine-5-type methyltransferase, which catalyzes the transfer of a methyl group from *S*-adenosyl-L-methionine (AdoMet) to the 5 position of cytosine in DNA. There is a formation of a covalent bond between a cysteine residue in the catalytic site of the enzyme and the 6 position of the cytosine. This generates a delocalized carbanion that can accept the methyl group from the sulfonium center of AdoMet, forming a methylated dihydrocytosine intermediate. The loss of the hydrogen from the 5 position and the release of the enzyme leads to the methylated product.

A crystal structure of a ternary complex containing bound M. *Hha*I, an oligonucleotide containing 5-fluorocytosine at the target position and the end product of the reaction, *S*-adenosylhomocysteine, showed that the target cytosine was flipped out of the DNA helix and into the binding pocket of the protein (*31*), but there was no overall bend to the DNA helix itself. Also, crystal studies of the same complex with mismatches at the target site in the DNA, including an abasic site, showed that base flipping still occurs (*32*).

In addition to the crystal structure studies, the *Hha*I system has been investigated using numerous other techniques. Footprinting studies show extensive protein−DNA contacts along the sugar−phosphate backbone (*33*). This same study indicated no protein−base contacts between the methyltransferase and the target base. Binding studies have shown that not only does the methyltransferase still bind and perform the base flipping with a methylated substrate (*34*, *35*) but it also preferentially binds mismatches (*35*). Footprinting studies indicate little structural deformation of the DNA helix by 5-methylation of the target cytosine (*16*). No dramatic structural differences are seen between sequences containing the -GCGC- sequence and standard B-form DNA from crystal or solution NMR structures (*36*, *37*). Unfortunately, there are no crystal or solution structures of the DNA sequence complexed with the endonuclease to determine if any structural variations occur in this complex.

As is typical of restriction and modification systems, methylation of the target base in the *Hha*I binding site completely inhibits the cleavage by the endonuclease. Additionally, in work done by Klimasauskas et al. (*35*), it has been shown that methylation at the 5 position on the target cytosine reduces by 2−4-fold the binding affinity of the methyltransferase.

The current investigation is intended to determine the effect that methylation has on the internal dynamics of the DNA backbone at the site of the target dC. The particular sequence of interest in this study is [d-($G_1A_2T_3A_4\underline{G}_5\underline{C}_6\underline{G}_7\underline{C}_8T_9A_{10}T_{11}C_{12}$)]$_2$, where the underlined residues constitute the binding site for the *Hha*I restriction system and the numbering sequence is for convenient referral to specific residues. This sequence was chosen because it is the same sequence used in the crystallographic studies of the methyltransferase/DNA complexes, mentioned previously.

MATERIALS AND METHODS

Chemical Synthesis of Selectively Deuterated DNAs. To investigate the sequence specificity of furanose ring dynamics, selectively deuterated nucleosides were synthesized (Figure 2) and incorporated into the DNA dodecamer using

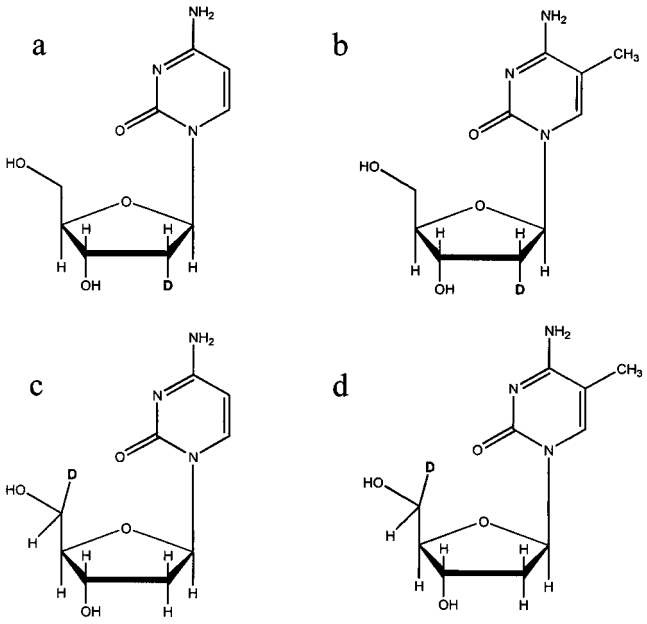

FIGURE 2: Deuterated nucleosides used in this study: (a) [2″-^{2}H]-2′-deoxycytidine, (b) [2″-^{2}H]-5-methyl-2′-deoxycytidine, (c) [5′/5″-^{2}H]-2′-deoxycytidine, and (d) [5′/5″-^{2}H]-5-methyl-2′-deoxycytidine.

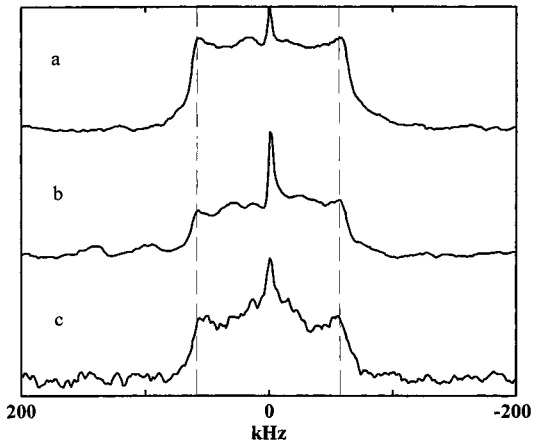

FIGURE 3: Comparison of line shapes for (a) [2″-^{2}H]-C6 from the DNA sequence [d(GATAGCGCTATC)]$_2$ with W = 9.8 and (b) W = 11.5 and (c) [2″-^{2}H]-5-methyl-C6 from the DNA sequence [d(GATAG^{5Me}CGCTATC)]$_2$ with W = 11. Note that the change in signal-to-noise ratio for (a), (b), and (c) is due to the different number of experimental scans and sample size. The central isotopic component is due to residual HDO. Dashed lines indicate constant frequency and that there is no change in spectral width upon methylation.

phosphoramidite chemistry. Deuterated DNA oligonucleotides were synthesized as follows. [2″-^{2}H]-dC and [2″-^{2}H]-T was prepared by the method of Robins et al. with some minor modifications to the solvent systems (*38*). [5′/5″-^{2}H]-dC and T were synthesized with a method similar to that for T, as described by Orban and Reid (*39*). The deuteration is nonstereospecific, which will not affect the NMR spectrum. Deuterated dC nucleosides were converted to 5′-*O*-(DMT)-2′-dC-3′-CED phosphoramidites as described previously (*40*). Deuterated thymidine nucleosides were converted to 5′-*O*-(DMT)-2′-T-3′-CED-phosphoramidites as described previously and converted to N4-triazole derivatives using the procedure of Cowart et al. (*41*). N4-triazole phosphoramidites were incorporated at the 6 position of the growing DNA oligonucleotide sequence using an ABI Model 394 automated DNA/RNA synthesizer. Deprotection of the DNA in concentrated ammonia (2 days, 55 °C) converted the N4-protected thymidine derivative to the 5-methyl-2′-dC nucleotide. The presence of 5-methyl-2′-dC was determined using both solution NMR and digestion followed by HPLC analysis (data not shown). Oligonucleotides were purified as described previously, salted (10% NaCl by weight), packed into a 5 mm solid-state NMR Kel-F sample chamber, and hydrated by vapor diffusion in a humidity chamber containing saturated salts in ^{2}H-depleted water (75% relative humidity at 20 °C) (*42*). Water content was quantified gravimetrically by the parameter W (number of water molecules per nucleotide) and is accurate to ±1 waters per nucleotide.

Solid-State NMR Spectroscopy. All ^{2}H NMR experiments were performed on a home-built NMR spectrometer (11.75 T, deuterium Larmor frequency of 76.776 MHz). A quadrupolar echo pulse sequence with an eight-step phase cycling scheme was implemented with a delay of 40 μs between 90° pulses (typically, 2.4−2.8 μs in duration) and a dwell time of 200 ns during acquisition. Data acquisition was initiated prior to the echo maximum. The time domain data were left-shifted and apodized with 3000 Hz Lorentzian line broadening prior to Fourier transformation. Spin−lattice relaxation times were determined using an inversion recovery pulse sequence, which incorporated a 180° composite pulse to ensure broad-band excitation (*43*). To obtain powder-averaged Zeeman spin−lattice relaxation times, $\langle T_{1Z} \rangle$, the integrated intensity of the powder spectrum was monitored as a function of recovery time and analyzed using a nonlinear least-squares fitting routine (*44*). All experimental line shape and Zeeman spin−lattice relaxation data were collected at room temperature. Dynamically averaged deuterium line shapes were calculated utilizing the MXET1 computer code written by the Vold research group (*45*, *46*).

RESULTS

Deuterium Line Shape and Relaxation Data for the 2″ Deuteron in C6. The quadrupolar echo line shapes for C6 at 76.775 MHz with W = 9.8 and W = 11.5 are shown in Figure 3a and b, respectively, for a hydration study of the line shape. There is significant motional averaging at the intermediate time scale (~10^6−10^7 Hz) seen in these spectra, as is evidenced by deviation from the Pake doublet form.

Table 1 shows the spin−lattice relaxation times for [2″-^{2}H]-C6 at the hydration levels discussed above and compares them to C9 from the Dickerson dodecamer. The $\langle T_{1Z} \rangle$ for C6 with W = 9.8 at 76.775 MHz is 28 ± 5 ms and with W = 11.5 at 76.775 MHz is 30 ± 8 ms. The $\langle T_{1Z} \rangle$'s are comparable to the $\langle T_{1Z} \rangle$ observed for the mobile C9 position in the Dickerson sequence, which has a $\langle T_{1Z} \rangle$ of 30 ± 7 (*17*).

Deuterium Line Shape and Relaxation Data for the 2″ Deuteron in 5-Methyl-C6. The quadrupolar echo line shape for [2"-^{2}H]-5-methyl-C6 (^{5Me}C6) with hydration W = 11 is shown in Figure 3c, in comparison to that of unmethylated C6 with W = 9.8 and 11.5. There is some perturbation on the line shape in the methylated sequence as compared to the unmethylated sequence, but it is not significant, and the difference may be only a result of differential hydration. As shown in Table 1, the spin−lattice relaxation times at comparable hydrations are comparable within the experi-

Table 1: Spin−Lattice Relaxation Times for Furanose-Labeled Sites[a]

sample	hydration (waters/nucleotide)	$\langle T_{1Z} \rangle$ (ms)
[2″-^{2}H]-C6 (from *Hha*I sequence)	W = 9.8 ± 1	28 ± 5
[2″-^{2}H]-C6 (from *Hha*I sequence)	W = 11.5 ± 1	30 ± 8
[2″-^{2}H]-5-methyl-C6 (from *Hha*I sequence)	W = 11.0 ± 1	21 ± 4
[2″-^{2}H]-C9 (from *Eco*RI sequence)	W = 12.3 ± 1	30 ± 7[b]

[a] Spin−lattice relaxation times for C6 from the DNA sequence containing the *Hha*I restriction site, $[d(GATAGCGCTATC)]_2$, for several hydration levels and field strengths and compared to values from C9 from the Dickerson dodecamer mentioned in the text, $[d(CGCGAATTCGCG)]_2$, which contains the *Eco*RI restriction site. [b] Data modified from ref *17*.

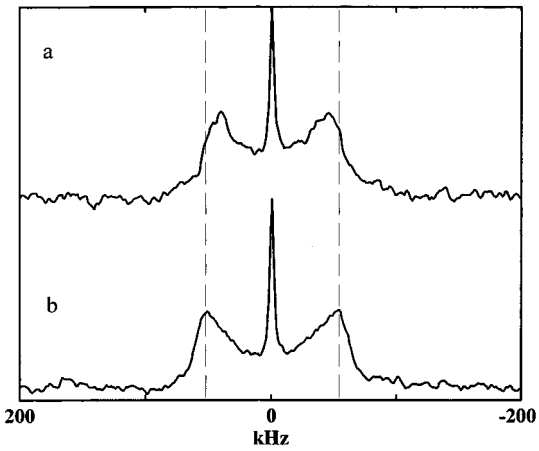

FIGURE 4: Comparison of line shapes for (a) [5′/5″-^{2}H]-C6 from the DNA sequence $[d(GATAGCGCTATC)]_2$ with W = 11.8 and (b) [5′/5″-^{2}H]-5-methyl-C6 from the DNA sequence $[d(GATAG^{5Me}CGCTATC)]_2$ with W = 11.2, both at 76.775 MHz. Note that the dashed lines are of constant frequency, indicating the widening of the spectrum upon methylation.

mental error. The [2″-^{2}H]-5-methyl-C6 sample has a $\langle T_{1Z} \rangle$ of 21 ± 4 ms, whereas its unmethylated counterpart has a $\langle T_{1Z} \rangle$ of 30 ± 8 ms, at a similar hydration level. Therefore, methylation has little effect on the relaxation time of the furanose ring for M. *Hha*I methyltransferase target dC, a result also observed in the Dickerson dodecamer.

Deuterium Line Shape and Relaxation Data for the 5′/5″ Deuterons in C6. The quadrupolar echo line shape for C6 with hydration W = 11.8 at 76.775 MHz is shown in Figure 4a. The line shape deviates from the Pake doublet that would be seen in a sample with minimal motional averaging. There is a bending of the horns toward the center of the spectrum, and two shoulders that appear to be above the experimental noise are also present. But the most interesting feature is the narrowing of the spectrum.

Alam et al. (*47*) determined that the static quadrupolar coupling constant (QCC_{static}, $^4/_3$ multiplied by the horn-to-horn splitting in a line shape with no dynamic averaging) for the 5′/5″ methylene deuterons in the thymidine samples from the Dickerson sequence is 175 ± 1 kHz. Dynamic amplitudes of motion can be extrapolated from QCC_{eff}, $^4/_3$ multiplied by the horn-to-horn splitting in a dynamically averaged line shape. Measuring the QCC_{eff} for the [5′/5″-^{2}H]-C6 sample gives a value of 120 ± 2 kHz. For comparison, the QCC_{eff} of the [5′/5″-^{2}H]-C9 sample from the Dickerson dodecamer (shown in Figure 1a) hydrated to W = 10.1 has a QCC_{eff} of 60 kHz, and the QCC_{eff} for the thymidine sites in the same sequence hydrated to W = 11.9 is 150 ± 2 kHz. These data are summarized in Table 3.

Table 2: Spin−Lattice Relaxation Times for Backbone-Labeled Sites[a]

sample	hydration (waters/nucleotide)	$\langle T_{1Z} \rangle$ (ms)
[5′/5″-^{2}H]-C6 (from *Hha*I sequence)	W = 11.8 ± 1	47 ± 5
[5′/5″-^{2}H]-5-methyl-C6 (from *Hha*I sequence)	W = 11.3 ± 1	68 ± 8
[5′/5″-^{2}H]-C9 (from *Eco*RI sequence)	W = 10.5 ± 2	30 ± 10[b]
[5′/5″-^{2}H]-5-methyl-C9 (from *Eco*RI sequence)	W = 13.1 ± 1	55[c]
[5′/5″-^{2}H]-T7/*T*8 (from *Eco*RI sequence)	W = 10.5 ± 2	59 ± 10[d]

[a] Spin−lattice relaxation time for [5′/5″-^{2}H]-C6 from the DNA sequence $[d(GATAGCGCTATC)]_2$ and the sequence $[d(GATAG^{5Me}CGCTATC)]_2$. These values are compared to relaxation times from previous solid-state deuterium NMR studies. [b] Data taken from ref *17*. [c] Data taken from ref *28*. [d] Data taken from ref *47*.

Table 3: Values for the QCC_{eff} and Reduction Factor for Backbone-Labeled Samples[a]

sample	hydration (waters/nucleotide)	QCC_{eff} (kHz)	reduction factor, Λ
[5′/5″-^{2}H]-C6 (from *Hha*I sequence)	W = 11.8 ± 1	120 ± 2	0.69
[5′/5″-^{2}H]-5-methyl-C6 (from *Hha*I sequence)	W = 11.3 ± 1	150 ± 2	0.86
[5′/5″-^{2}H]-C9 (from *Eco*RI sequence)	W = 10.5 ± 2	60 ± 2[b]	0.34
[5′/5″-^{2}H]-5-methyl-C9 (from *Eco*RI sequence)	W = (10−12) ± 2	150 ± 2[c]	0.86
[5′/5″-^{2}H]-T7/*T*8 (from *Eco*RI sequence)	W = 10.5 ± 2	150 ± 2[d]	0.86

[a] Values for the effective quadrupolar coupling constants (QCC_{eff}) and reduction factor Λ (defined in eq 1) for [5′/5″-^{2}H]-C6 from the DNA sequence $[d(GATAGCGCTATC)]_2$ and the sequence $[d(GATAG^{5Me}CGCTATC)]_2$. These values are compared to relaxation times from previous studies. [b] Data taken from ref *17*. [c] Data taken from ref *28*. [d] Data taken from ref *47*.

From the inversion recovery experiments, the $\langle T_{1Z} \rangle$ was found to be 47 ± 5 ms. This value is comparable to the values from the thymidines in the *Eco*RI binding site, which have $\langle T_{1Z} \rangle$'s of 59 ± 10 ms with W = 10.5 and 41 ± 10 ms at W = 11.9 (*47*). The value of 47 ms for the [5′/5″-^{2}H]-C6 sample is longer than C9 from Dickerson, which has a $\langle T_{1Z} \rangle$ of 30 ± 10 ms with a hydration of W = 10.5 (*17*). Table 2 summarizes these data.

Deuterium Line Shape and Relaxation Data for the 5′/5″ Deuterons in 5-Methyl-C6. The quadrupolar echo line shape for ^{5Me}C6 at hydrations W = 11.2 is shown in Figures 4b. The form of the line shape is still relatively close to that of the Pake powder pattern. Additionally, the measured QCC_{eff} is 150 ± 2 kHz. This is comparable to the QCC_{eff} of the [5′/5″-^{2}H]-5-methyl-C9 from the Dickerson dodecamer (shown in Figure 1b) with a hydration of W = 13.1 that has a QCC_{eff} of 148 kHz (*28*).

The spin−lattice relaxation time obtained from these experiments is 68 ± 8 ms, and this is even longer than the

values for the thymidines from the Dickerson dodecamer, which have $\langle T_{1Z} \rangle$'s of 59 ± 10 ms with W = 10.5 and 41 ± 10 ms at W = 11.9. Additionally, the [5′/5″-^{2}H]-5-methyl-C9 sample from the Dickerson dodecamer with a W = 13.1 has a $\langle T_{1Z} \rangle$ of 55 ms (*28*). These data are again summarized in Table 2.

DISCUSSION

Effects of Base Methylation on Motion of the Furanose Ring. Three principal methods are used in this study to determine a perturbation of molecular dynamics using solid-state deuterium NMR. The primary method is direct simulation of deuterium line shapes using a well-conceived model for the molecular motion in question. When direct simulation is impossible, comparison of the fully relaxed line shapes and a measurement of their respective QCC_{eff}'s can give an idea of the relative rate and amplitude of molecular motion for two sites. Both [2″-^{2}H]-C6 and [2″-^{2}H]-5-methyl-C6 display significant motional averaging on the intermediate time scale ($\sim 10^6 - 10^7$ Hz), and therefore direct measurement of QCC_{eff} is impossible. Although these line shapes have not been directly simulated, significant work has been done in the past to simulate dynamically averaged furanose ring line shapes using a two-site jump model (*17*) and a model of single axis Brownian diffusion through a potential (*28*). Additionally, we have developed a new model of Brownian diffusion through an angular potential for describing deuterium motion assuming a puckering amplitude and associated angular trajectory and a potential energy surface $U(\varphi)$ (*21*). If we assume that the motions present in this sample are similar in nature to other dynamic furanose rings, we can predict that their amplitude of motion will be on the order of 38°, or they have a puckering amplitude of about 0.4 Å. Upon inspection it appears there is minimal change in the overall form of the line shapes in the methylated versus unmethylated samples, as can be seen in Figure 3, indicating that neither the amplitude nor the rate of the furanose ring motion changes appreciably upon base methylation.

Comparison of the relaxation times is another good measure of change in the molecular motions. Alam and Drobny showed that for 12 base pair DNAs, the ^{2}H spin−lattice relaxation time is near a minimum and does not change markedly in the hydration range W = 10−16 (*47*). Table 1 indicates that there is negligible change in the spin−lattice relaxation times between [2″-^{2}H]-C6 and [2″-^{2}H]-5-methyl-C6, within the experimental error. From these two pieces of information, it can be concluded that base methylation of the target dC in the *Hha*I binding site has a minimal effect on the furanose ring motion. A similar result was obtained for the methylation of the *Eco*RI DNA binding site dC (*28*).

Effects of Base Methylation on Motion of the Backbone Methylene Group. The criteria discussed in the previous section can now be applied to the experimental data for the backbone-labeled samples. First, a comparison of the line shapes and QCC_{eff} is necessary. Figure 4 shows the line shapes for [5′/5″-^{2}H]-C6 (a) and [5′/5″-^{2}H]-5-methyl-C6 (b) in order to show the change in the horn-to-horn splitting of the backbone line shapes upon base methylation. The dashed lines are lines of constant frequency, and as can be seen, there is a noticeable widening of the line shape and a straightening of the edge of the horns in the methylated versus the unmethylated samples.

As stated previously, it has been shown (*47*) that QCC_{static} for a backbone methylene deuteron in 175 ± 1 kHz. The amplitude reduction factor (ARF) Λ, used to assess motional averaging of the QCC (*48*), is defined as

$$\Lambda = \frac{QCC_{eff}}{QCC_{static}} = \frac{\left(\frac{e^2qQ}{h}\right)_{eff}}{\left(\frac{e^2qQ}{h}\right)_{static}} \quad (1)$$

For the current work, we assume that QCC_{static} will be identical to that determined previously for the 5′/5″ deuterons in DNA, i.e., 175 ± 1 kHz (*47*). For [5′/5″-^{2}H]-C6, the value of 120 ± 2 kHz for QCC_{eff} gives a value of Λ of 0.69. For [5′/5″-^{2}H]-5-methyl-C6, QCC_{eff} is 150 ± 2 kHz, and Λ is 0.86. This numerical indication of spectral broadening implies differential motional averaging in the methylated versus unmethylated DNAs. These data, as well as those of other samples for comparison, are summarized in Table 3.

The next step is to look at the spin−lattice relaxation times, summarized in Table 2. [5′/5″-^{2}H]-C6 has a $\langle T_{1Z} \rangle$ of 47 ± 5 ms and [5′/5″-^{2}H]-5-methyl-C6 has a $\langle T_{1Z} \rangle$ of 68 ± 8 ms at comparable hydration levels. As can be seen, there is an increase in the relaxation time of the backbone methylene deuterons upon base methylation. As previously stated, there is a little dependence of the spin−lattice relaxation time on the sample hydration, so the changes therein are due to changes in the local motions and not the hydration.

Finally, we have attempted to simulate the line shapes directly. In our simulations, we have considered two types of motion, helical rotation and local motion of the backbone methylene subunits. The physical basis for this model has been established in a number of publications on the dynamics of hydrated, amorphous DNA. These studies and analyses are summarized in ref *48*. In summary, the whole molecule dynamics of short DNA oligomers may be described as a rigid rod motion, where the rotational rates are dependent upon the length-to-width ratio of the oligomer. Line shape studies of base-deuterated, furanose ring-deuterated, and backbone-deuterated DNA dodecamers at low levels of hydration (W < 16−18) indicate that the rigid body motion is dominated by rotation about the helix axis with an amplitude and rate that increase with the hydration level. At higher hydration levels (W > 18) rotation around an axis perpendicular to the helix axis must also be considered.

The exact nature of the rigid body motion, as discussed in ref *48* and in papers cited therein, has been identified in the presence of localized motions of various amplitudes and rates, by applying the same model of whole molecule motion at a given hydration level and temperature to a large number of deuterated sites on the same dodecamer. Although the DNA oligomers studied in ref *48* did not contain the *Hha*I binding site, like the DNA in the present study these prior studies dealt with dodecamers. The assumption that the whole molecule motion depends on the length-to-width ratio allows us to take models of helix motion from prior dodecamer studies and apply it to the present case. At W = 11 helical rotation is simulated using a nearest neighbor six-site jump with an azimuthal angle of ϕ = 0°, 60°, 120°, 180°, 240°, and 300°, a half angle amplitude θ = 40°, which is the angle of the C5′−D bond with respect to the longitudinal DNA

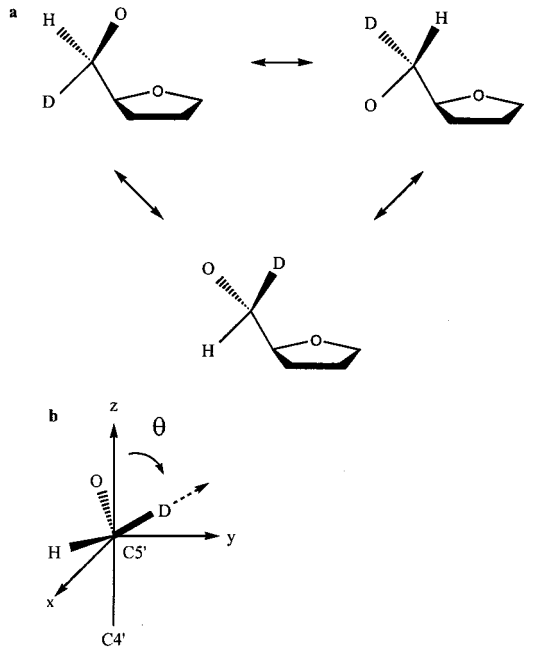

FIGURE 5: (a) Pictorial representation of the trans-gauche isomerization model of backbone motion used in simulations. (b) Pictorial representation of the axis of symmetry for the trans-gauche isomerization model described in the text. The angle θ is the tip angle from the C5′−C4′ axis of symmetry to the actual C5′−D bond.

helix axis, determined previously (*47*), and a rate constant between each site of $k = 10^4$ Hz. The rate parameter was determined from previous work with base-deuterated DNAs of identical overall sequence (*49*). Therefore, any difference in the line shape between the dodecamer CGCGAATTCGCG studied in ref *49* and the dodecamer studied here is attributed to local motions.

For the local motion of the backbone methylene group, we have used a three-site gauche-trans isomerization model to attempt to replicate our experimental data (see Figure 5a). This model, as well as a two-site jump and a biaxial four-site jump, has been used previously (*17*, *47*) to describe the motion of backbone methylene groups. Neither the two-site or biaxial four-site jump could effectively model line shapes whose QCC_{eff} was less than 140 kHz. For the three individual sites, where the axis of symmetry is the C5′−C4′ bond, they have an azimuthal angle of $\phi = 0°$, 120°, and 240° (representing the gauche$^+$, trans, and gauche$^-$ conformations) and a half-angle amplitude $\theta = 60°$ (see Figure 5b), taken from the tetrahedral symmetry of an sp^3-hybridized C5′ carbon. We chose this model for the local motion for several reasons. First, it is a good representation of the local symmetry of the 5′-methylene group. Second, it has been suggested previously that trans-gauche isomerization motions are occurring at these positions in the DNA backbone (*50*). In addition to this simple model, we are currently working on developing a model of diffusion through a potential energy surface for the motions of the backbone, which could better simulate the motion if it is not truly activated exchange between conformations.

Figure 6 shows a series of simulations using the trans-gauche isomerization model described above. A library of simulations has been developed for our model, with our best fits for the line shapes and spin−lattice relaxation times

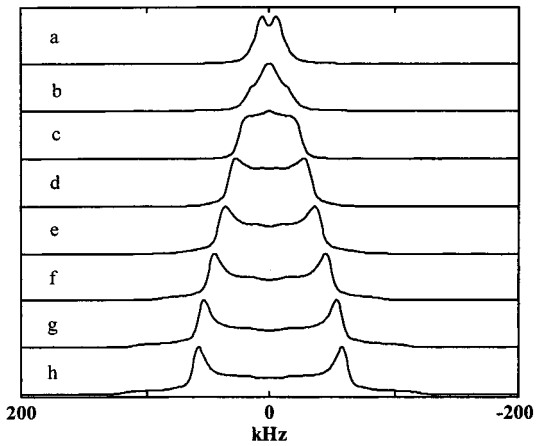

FIGURE 6: Series of simulations of backbone motion. It is a three-site jump model for the backbone motion, replicating trans-gauche isomerization with differential site populations for the three respective sites. For each simulation, the angle ϕ has values of 0°, 120°, and 240°, $\theta = 60°$, and rate constant $k = 5 \times 10^6$ Hz. The site populations for the three sites in question are as follows: (a) equal, (b) 0.40, 0.30, and 0.30, (c) 0.50, 0.25, and 0.25, (d) 0.60, 0.20, and 0.20, (e) 0.70, 0.15, and 0.15, (f) 0.80, 0.10, and 0.10, (g) 0.90, 0.05, and 0.05, and (h) 0.95, 0.025, and 0.025. Their respective values for QCC_{eff} are listed in Table 4.

Table 4: Values for the QCC_{eff} and Reduction Factor for Backbone Simulations[a]

site population distribution	QCC_{eff} (kHz)	reduction factor, Λ
equal	14 ± 1	0.08
0.40, 0.30, 0.30	*b*	*b*
0.50, 0.25, 0.25	*b*	*b*
0.60, 0.20, 0.20	73 ± 1	0.42
0.70, 0.15, 0.15	96 ± 1	0.55
0.80, 0.10, 0.10	120 ± 1	0.69
0.90, 0.05, 0.05	143 ± 1	0.82
0.95, 0.025, 0.025	153 ± 1	0.88

[a] Values for the effective quadrupolar coupling constants (QCC_{eff}) and reduction factor Λ (defined in eq 1) for the simulations described in the text. It is a three-site jump model for the backbone motion, replicating trans-gauche isomerization with differential site populations for the three respective sites. For each simulation, the angle ϕ has values of 0°, 120°, and 240°, $\theta = 60°$, and rate constant $k = 5 \times 10^6$ Hz. [b] Value could not be determined.

shown. We have shown one set to indicate the trend for adjusting the site populations, and effect which is the same for adjusting other parameters, as in the rate constant. For all of the simulations, the angles in question as well as the rate constant, *k*, between the individual sites are kept constant. The effect of changing *k* is most significant on the spin−lattice relaxation time, $\langle T_{1Z} \rangle$, with little change seen in the form of the line shape, except over a wide range of values for *k*. The adjustable parameter is the population of each site. The relative site populations range from equal for each site (top), to a population distribution of 0.95, 0.025, and 0.025 (bottom) for the three sites. As can be seen, there is a tremendous sensitivity to the distribution of populations for the three sites. The spectra widen as the distribution becomes increasingly unequal. The QCC_{eff}'s for each simulation are given in Table 4.

The simulations give an idea of the amplitude of motion, as well as site populations. The narrowing trend seen gives a comparable value for the QCC_{eff}, as compared to the line shape for the [5′/5″-^{2}H]-C6 sample in Figure 4a (QCC_{eff} =

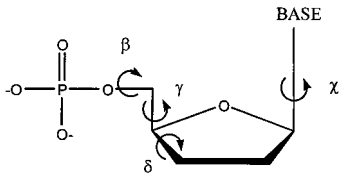

FIGURE 7: Pictorial definition of some DNA backbone torsion angles (β, γ, δ) and the glycosidic torsion angle χ.

120 kHz), for a population distribution of 0.8, 0.1, and 0.1 for the three sites ($QCC_{eff} = 120$ kHz) shown in Figure 6f. Additionally, the model with these populations and a rate constant of $k = 5 \times 10^6$ Hz has a $\langle T_{1Z} \rangle = 50$ ms, which is comparable to [5′/5″-^{2}H]-C6, which has a value of $\langle T_{1Z} \rangle = 47 \pm 5$ ms. Then, upon methylation, the population distribution becomes approximately 0.95, 0.025, and 0.025 ($QCC_{eff} = 153$ kHz), which compares to [5′/5″-^{2}H]-5-methyl-C6 ($QCC_{eff} = 150$ kHz), indicating that one particular conformation of the backbone becomes increasingly energetically favorable. Last, this model has ramifications for [5′/5″-^{2}H]-C9 from the Dickerson dodecamer, seen in Figure 1a. From our simulations we can extrapolate that this site, which has a $QCC_{eff} = 60$ kHz, has site populations between with values between equal for each ($QCC_{eff} = 14$ kHz) to 0.60, 0.20, and 0.20 ($QCC_{eff} = 73$ kHz).

The results from this study indicate that base methylation of the target dC in the *Hha*I binding site has an effect on the dynamics of the backbone methylene group. The reduction factor Λ is increased upon methylation from a value indicating large motional averaging ($\Lambda = 0.69$) to a value observed for all rigid sites ($\Lambda = 0.86$). There is an increase in the spin−lattice relaxation time in the backbone upon methylation, which is another indication of a decrease in local motion.

Conclusions. The results from the present study indicate a significant perturbation of the dC backbone motions upon base methylation, similar to those previously observed in the *Eco*RI recognition site (*28*). It remains to explain how a functional group substitution on the cytosine base influences the motions of the sugar−phosphate backbone. A possible explanation, supported by data in the literature, is that a rearrangement of base stacking occurs to alleviate increased steric hindrance caused by the additional methyl group on the cytosine heterocycle. A change in base stacking in turn influences the conformational flexibility of the backbone in the region of the methylation site. Supporting this hypothesis is work by Marcourt et al. (*13*), who observed changes in the rise, roll, local curvature, and twist parameters in the methyl-CpG versus unmethylated CpG dinucleotide units. Those authors calculated a change in stacking energy to −5.5 kcal/mol from −9 kcal/mol accompanying cytosine methylation. They also observed the most significant changes in torsional angles for γ and δ, as well as a noticeable change in β, where β and γ are the angles most closely associated with the 5′ carbon (see Figure 7). These structural changes alter the orientation between the base and backbone subunits, therefore having a greater steric effect on the backbone from the altered base. This direct steric hindrance then has the effect of perturbing the backbone dynamics.

The local flexibility of the *Hha*I DNA binding site could aid the interactions with the methyltransferase in addition to, or instead of, the endonuclease. The complex with the *Hha*I methyltransferase shows the well-known target site flipped out of the DNA helix and into the binding pocket of the enzyme (*31*). However, there is no evidence of any further deformation of the DNA helix, which emphasizes the importance of the interactions at this particular subunit. Interactions of the protein with the sugar−phosphate backbone may be the dominant interaction for the base-flipping mechanism. Perhaps the dynamics that we observe in the backbone and furanose ring subunits of the target dC play a role in the base-flipping process by again lowering the local rigidity and allowing the methyltransferase to perform its localized distortion of the DNA helix.

The concept of DNA exhibiting local, sequence-dependent flexibility has support in recent literature (*51*, *52*), but the picture is not completely clear (*53*). The results of Hodges-Garcia and Hagerman showed no change in flexibility for multiply methylated sequences, as measured by cyclization by T4 DNA ligase (*54*). They determined that there was little change in such quantities as persistence length, helix repeats, and torsion elastic constants between sequences with methylated and unmethylated cytosines. So although methylation has a somewhat modest effect on long-range properties such as the persistence length, the effects are clearly much more dramatic for properties such as amplitudes and rates of local motions. This is almost certainly a consequence of the fact that if a very small number of base pairs or subunits exhibit greater flexibility in a 500 bp sequence, for example, it will not contribute measurably to the persistence length.

Although the exact nature of the local flexibility of DNA continues to be a subject of discussion, the work presented in this report continues to support the hypothesis that local conformational flexibility of the DNA helix is correlated with aspects of its biological function.

ACKNOWLEDGMENT

We thank Professors J. M. Schurr and B. H. Robinson for helpful discussions. We also thank Dr. Eric Harwood and Karen Lo for assistance in operating the automated DNA synthesizer.

REFERENCES

1. Ahmad, I., and Rao, D. (1996) *Crit. Rev. Biochem. Mol. Biol. 31*, 361−380.
2. Adams, R. L. (1990) *Biochem. J. 265*, 309−320.
3. Bird, A. (1992) *Cell 70*, 5−8.
4. Cedar, H. (1988) *Cell 53*, 3−4.
5. Mayer-Jung, C., Moras, D., and Timsit, Y. (1997) *J. Mol. Biol. 270*, 328−335.
6. Davey, C., Pennings, S., and Allan, J. (1997) *J. Mol. Biol. 267*, 276−288.
7. Coulondre, C., Miller, J. H., Farabaugh, P. J., and Gilbert, W. (1978) *Nature 274*, 775−780.
8. Ramsahoye, B. H., Davies, C. S., and Mills, K. I. (1996) *Blood Rev. 10*, 249−261.
9. Hodges-Garcia, Y., and Hagerman, P. J. (1992) *Biochemistry 31*, 7595−7599.
10. Partridge, B. L., and Salisbury, S. A. (1996) Cambridge University.
11. Heinemann, U., and Hahn, M. (1992) *J. Biol. Chem. 267*, 7332−7341.
12. Geahigan, K. B. (1998) in *Chemistry*, p 263, University of Washington, Seattle, WA.
13. Marcourt, L., Cordier, C., Couesnon, T., and Dodin, G. (1999) *Eur. J. Biochem. 265*, 1032−1042.

14. Lefebvre, A., Mauffret, O., El Antri, S., Monnot, M., Lescot, E., and Fermandjian, S. (1995) *Eur. J. Biochem. 229*, 445−454.
15. Vargason, J. M., Eichman, B. F., and Ho, P. S. (2000) *Nat. Struct. Biol. 7*, 758−761.
16. Fox, K. R. (1986) *Biochem. J. 234*, 213−216.
17. Hatcher, M. E., Mattielo, D. L., Meints, G. A., Orban, J., and Drobny, G. P. (1998) *J. Am. Chem. Soc. 120*, 9850−9862.
18. Dickerson, R. E., and Drew, H. R. (1981) *J. Mol. Biol. 149*, 761−786.
19. Aramini, J. M., Mujeeb, A., and Germann, M. (1998) *Nucleic Acids Res. 26*, 5644−5654.
20. Hatcher, M. E. (1996) in *Chemistry*, p 138, University of Washington, Seattle, WA.
21. Meints, G. A. J., Karlsson, T., and Drobny, G. P. (2001) *J. Am. Chem. Soc.* (in press).
22. Meints, G. A. J. (2000) in *Chemistry*, p 250, University of Washington, Seattle, WA.
23. Betrand, H.-O., Ha-Duong, T., Fermandjian, S., and Hartmann, B. (1998) *Nucleic Acids Res. 26*, 1261−1267.
24. Cordier, C., Marcourt, L., Petitjean, M., and Dodin, G. (1999) *Eur. J. Biochem. 261*, 722−733.
25. Lefebvre, A., Mauffret, O., Hartmann, B., Lescot, E., and Fermandjian, S. (1995) *Biochemistry 34*, 12019−12028.
26. Lefebvre, A., Mauffret, O., Lescot, E., Hartmann, B., and Fermandjian, S. (1996) *Biochemistry 35*, 12560−12569.
27. el Antri, S., Mauffret, O., Monnot, M., Lescot, E., Convert, O., and Fermandjian, S. (1993) *J. Mol. Biol. 230*, 373−378.
28. Geahigan, K. B., Meints, G. A., Hatcher, M. E., Orban, J., and Drobny, G. P. (2000) *Biochemistry 39*, 4939−4946.
29. Brennan, C. A., Van Cleve, M. D., and Gumport, R. I. (1986) *J. Biol. Chem. 261*, 7270−7278.
30. Wu, J. C., and Santi, D. V. (1987) *J. Biol. Chem. 262*, 4778−4786.
31. Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) *Cell 76*, 357−369.
32. O'Gara, M., Horton, J. R., Roberts, R. J., and Cheng, X. (1998) *Nat. Struct. Biol. 5*, 872−877.
33. Renbaum, P., and Razin, A. (1995) *J. Mol. Biol. 248*, 19−26.
34. O'Gara, M., Klimasauskas, S., Roberts, R. J., and Cheng, X. (1996) *J. Mol. Biol. 261*, 634−645.
35. Klimasauskas, S., and Roberts, R. J. (1995) *Nucleic Acids Res. 23*, 1388−1395.
36. Bingman, C., Li, X., Zon, G., and Syundaralingam, M. (1992) *Biochemistry 31*, 12803−12812.
37. Ho, W. C., Steinbeck, C., and Richert, C. (1999) *Biochemistry 38*, 12597−12606.
38. Robins, M. J., Wilson, J. S., and Hansske, F. (1983) *J. Am. Chem. Soc. 105*, 4059−4065.
39. Orban, J., and Reid, B. R. (1989) *J. Labelled Compd. Radiopharm. 27*, 195−198.
40. Gait, M. J. (1984) in *Practical Approach* (Rickwood, D., and Hames, B. D., Eds.) p 217, IRL Press, Oxford and Washington, DC.
41. Cowart, M., Gibson, K. J., Allen, D. J., and Benkovic, S. J. (1989) *Biochemistry 28*, 1975−1983.
42. Weast, R. C. (1979) *CRC Handbook of Chemistry and Physics*, 60th ed., CRC Press, Boca Raton, FL.
43. Tycko, R. (1983) *Phys. Rev. Lett. 51*, 775−777.
44. deFontaine, D. L., Ross, D. L., and Ternai, B. J. (1975) *J. Magn. Reson. 18*, 276.
45. Greenfield, M. S., Ronemus, A. D., Vold, R. L., Vold, R. R., Ellis, P. D., and Raidy, T. E. (1987) *J. Magn. Reson. 72*, 89−107.
46. Vold, R. R., and Vold, R. L. (1991) in *Advances in Magnetic and Optical Resonance*, pp 85−171, Academic Press, San Diego, CA.
47. Alam, T. M., Orban, J., and Drobny, G. P. (1991) *Biochemistry 30*, 9229−9237.
48. Alam, T. M., and Drobny, G. P. (1991) *Chem. Rev. 91*, 1545−1590.
49. Kintanar, A., Hunag, W. C., Schindele, D. C., Wemmer, D. E., and Drobny, G. P. (1989) *Biochemistry 28*, 282−293.
50. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
51. Okonogi, T. M., Reese, A. W., Alley, S. C., Hopkins, P. B., and Robinson, P. B. (1999) *Biophys. J. 77*, 3256−3276.
52. Okonogi, T. M., Alley, S. C., Reese, A. W., Hopkins, P. H., and Robinson, B. H. (2000) *Biophys. J. 78*, 2560−2571.
53. Hagermann, P. J. (1988) in *Annual Reviews of Biophysics and Biophysical Chemistry* (Engeleman, D. M., Cantor, C. E., and Pollard, T. D., Eds.) pp 265−285, Biophysics and Biophysical Society, Palo Alto, CA.
54. Hodges-Garcia, Y., and Hagerman, P. J. (1995) *J. Biol. Chem. 270*, 197−201.

BI0102555