Deoxycytidine methylation does not affect DNA·RNA hybrid formation or B–A transitions of (dG)_n·(dC)_n sequences

G.Terrance Walker

Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, CA 94720, USA

Received November 9, 1987; Revised and Accepted February 5, 1988

ABSTRACT

Optical thermal denaturation and circular dichroism (CD) experiments were performed with the following non-selfcomplementary duplex DNA, RNA and DNA·RNA hybrids: (I) dGAG₃C₃G₃CTC·dGAGC₃G₃C₃TC, (II) dGAG₃m⁵C₃G₃m⁵CTC·dGAGm⁵C₃G₃m⁵C₃TC, (III) rGAG₃C₃G₃CUC·rGAGC₃G₃C₃UC, (IV) dGAG₃C₃G₃CTC·rGAGC₃G₃C₃UC, (V) rGAG₃C₃G₃CUC·dGAGC₃G₃C₃TC, (VI) dGAG₃m⁵C₃G₃m⁵CTC·rGAGC₃G₃C₃UC, (VII) rGAG₃C₃G₃CUC·dGAGm⁵C₃G₃m⁵C₃TC. Duplex stabilities (ΔG° at 60°C) increase in the order: I < IV < II = V = VI < VII < III. Large enthalpic stabilization is associated with intrastrand stacking of guanosine (rG) residues. CD spectroscopy indicates B-form conformations for the unmethylated and methylated DNA (I,II), A-form geometry for the RNA (III), and DNA·RNA hybrid (IV - VII) conformations resembling but not identical to A-RNA. C5-methyldeoxycytidine does not significantly influence DNA conformation, DNA·RNA hybrid formation, or the ability of DNA to adopt an A-type conformation in trifluoroethanol solutions.

INTRODUCTION

Even though DNA·RNA hybrids are an integral part of transcription, reverse transcription and DNA replication, few structural and thermodynamic investigations of hybrids have been performed. In addition to providing insight into such biological functions, physical studies of hybrids provide an ideal system for investigating the propensity of different DNA sequences to adopt A-form structures. For example, dA_n·rU_n hybrids are extremely unstable, presumably due to reluctance of deoxyadenosine residues to adopt an A-form conformation (1).

The DNA sequence dG_n·dC_n exhibits greater disposition than random sequence DNA for adopting an A-form conformation. Several DNA duplexes with dG_n·dC_n sequences (e.g., d(↑CCGG), d(G₂C₂G₂C₂), d(G₂m⁵CCG₂C₂), d(C₄G₄), d(G₄C₄)) have crystallized in A-form conformations (2). However, it is apparent that B- to A-form transitions of dG_n·dC_n sequences are often associated with dehydration and/or intermolecular packing forces as observed through a combination of crystal and solution studies (3,4). In order to adequately understand B-A transitions in DNA, it is necessary to differentiate between the intrinsic A-form stability of various sequences in dilute solutions and the tendency of certain sequences to undergo aggregation or dehydration, both of which may induce B- to A-form conversions.

The effect of deoxycytidine methylation upon B-A transitions of dG_n·dC_n sequences is not

adequately established. Nishimura et al. (5) determined that C5-methylation markedly stabilized the A-form of poly(dG)·poly(dC), and Frederick et al. (2) suggested that methylation may stabilize A-DNA. In contrast, Sarma et al. (6) reported methylation causes an A to B transition in both strands of poly(dG)·poly(dC). The extent of methylation in eukaryotic DNA (7) lends considerable importance to the issue.

The present study was designed to investigate the structural and thermodynamic properties associated with the base sequence $(G)_n \cdot (C)_n$ in DNA, RNA and DNA·RNA hybrids. The effect of methylation at the C5 position of deoxycytidine also was of interest in this context. Unfortunately, such a study of $(G)_n \cdot (C)_n$ sequences is hampered because long runs of $(G)_n \cdot (C)_n$ tend to aggregate (8-11) and are susceptible to base pair slippage. The duplex oligonucleotide sequence 5'GAG₃C₃G₃CTC·5'GAGC₃G₃C₃TC (with U substituted for T in RNA) was chosen as representative of $(G)_n \cdot (C)_n$ sequences and to ensure formation of specific duplexes. DNA (unmethylated and methylated), RNA and all possible DNA·RNA hybrid combinations were investigated.

Structures of the DNA, RNA and DNA·RNA hybrids were investigated by circular dichroism (CD) spectroscopy which is sensitive to both primary sequence and secondary structure (12,13). Thermodynamic properties of the duplexes were determined from optical thermal denaturation experiments. The effect of deoxycytidine methylation upon B-A transitions was assessed in the presence of trifluoroethanol which is known to produce a B to A conversion in DNA (14). The results provide information about the effect of deoxycytidine methylation upon B-A transitions of $(G)_n \cdot (C)_n$ sequences and extend general knowledge of DNA·RNA hybrids.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification. DNA oligomers were synthesized using an Applied Biosystems 381A instrument. RNA oligomers with 5'-triphosphates were synthesized using T7 RNA polymerase and DNA templates (15). T7 RNA polymerase was isolated from the overproducing *E. coli* strain BL21 containing the plasmid pAR1219 according to the method of Davanloo et al. (16). Oligonucleotides were purified by preparative 20% acrylamide gel electrophoresis under denaturing conditions (7 M urea). Oligonucleotides were eluted from gel electrophoresis bands and purified on Sep-Pak (Millipore). RNA 5'-triphosphates were removed for the thermodynamic experiments using calf intestinal phosphatase (Boehringer-Mannheim) prior to Sep-Pak purification. Presence of 5'-triphosphates did not affect CD spectra. The integrity of the oligonucleotides was checked by gel electrophoresis after completion of all experiments.

Circular Dichroism Spectroscopy. CD spectra were recorded on a JASCO J500C spectropolarimeter using 1 cm pathlength cuvettes. Instrument conditions were 1 mdeg/cm, 2 nm/min, 16 sec time constant, 1nm spectral bandwidth. Molar ellipticity, $[\Theta]$, was calculated from

$$[\Theta] = 100 \cdot (\theta / Cl)$$

where θ is ellipticity (degrees), C is the molar base pair concentration, and l is the pathlength (centimeters). Base pair concentrations (12-25 μM) were determined based upon extinction coefficients calculated from published dimer and monomer extinction coefficients (17) using the nearest-neighbor approximation (18). The amount of single-stranded excess in the non-selfcomplementary duplex samples is estimated to be less than 5%.

Thermodynamic Analysis. Helix to coil transitions were monitored at 260 nm using a Gilford 250 uv/vis spectrometer with a Gilford 2527 thermoprogrammer at a heating rate of 1 deg/min. A 1 cm pathlength cuvette was used. Absorbance at 260 nm versus temperature data were converted to fraction single-stranded nucleic acid versus temperature and then smoothed by an individual point least-square fit to a second order polynomial (19). Van't Hoff enthalpy values were calculated from the slope of $\ln(K)$ versus $1/T$ or from the derivative of the smoothed fraction single-stranded nucleic acid versus temperature data (20). The two methods agreed within 5%. Linear van't Hoff plots validated use of a two-state model. Entropy values were calculated from the slope of ΔG versus temperature plots. The high melting temperature of the duplexes precluded performance of thermal denaturation experiments over a range of oligonucleotide concentrations. Consequently, thermodynamic values could not be obtained from the dependence of melting temperature on oligonucleotide concentration.

RESULTS

Thermodynamics. Thermodynamic values for duplex formation of the DNA, RNA and DNA-RNA hybrids in 50 mM NaCl, 6 mM Na_2HPO_4 , 2 mM NaH_2PO_4 , 0.1 mM Na_2EDTA , pH 7 were determined from optical thermal denaturation experiments (Figure 1). Values of van't Hoff enthalpy (ΔH°), entropy (ΔS°), free energy (ΔG°), and melting temperature (T_m) are reported in Table I. Hypochromicity measurements verified duplexes as the dominant form as opposed to

Figure 1: Optical thermal denaturation experiments with the duplex DNA, RNA and DNA-RNA hybrids in 50 mM NaCl, 6 mM Na_2HPO_4 , 2 mM NaH_2PO_4 , 0.1 mM Na_2EDTA , pH 7 at nucleotide concentrations of 50 μM . Fraction of single-stranded nucleic acid is plotted versus temperature. Transition profile labels correspond to the sequences in Table I. (A) Unmethylated DNA, RNA and the corresponding hybrids. (B) Methylated DNA, RNA and the corresponding hybrids.

Table I. Thermodynamic Values of Duplex Formation

duplex number	sequence	ΔH° (kcal/mol)	ΔS° (eu)	$\Delta G^\circ(60^\circ\text{C})$ (kcal/mol)	T_m ($^\circ\text{C}$)
I	5'dGAG ₃ C ₃ G ₃ CTC dCTC ₃ G ₃ C ₃ GAG5'	-98	-263	-10.4	65
II	5'dGAG ₃ m ⁵ C ₃ G ₃ m ⁵ CTC dCTm ⁵ C ₃ G ₃ m ⁵ C ₃ GAG5'	-101	-264	-13.0	73
III	5'rGAG ₃ C ₃ G ₃ CUC rCUC ₃ G ₃ C ₃ GAG5'	-125	-323	-17.4	84
IV	5'dGAG ₃ C ₃ G ₃ CTC rCUC ₃ G ₃ C ₃ GAG5'	-104	-276	-12.1	70
V	5'rGAG ₃ C ₃ G ₃ CUC dCTC ₃ G ₃ C ₃ GAG5'	-114	-303	-13.1	72
VI	5'dGAG ₃ m ⁵ C ₃ G ₃ m ⁵ CTC rCUC ₃ G ₃ C ₃ GAG5'	-109	-287	-13.4	74
VII	5'rGAG ₃ C ₃ G ₃ CUC dCTm ⁵ C ₃ G ₃ m ⁵ C ₃ GAG5'	-122	-321	-15.1	77

Buffer was 50 mM NaCl, 6 mM Na₂HPO₄, 2 mM NaH₂PO₄, 0.1 mM Na₂EDTA, pH 7. Nucleotide concentrations were 50 μM . Reproducibility of each value is $\pm 3\%$.

higher order structures such as triple helices.

The RNA duplex (III) is more stable (ΔG° at 60 $^\circ\text{C}$) than the two DNA duplexes (I,II) due to a larger enthalpic component. The methylated DNA duplex (II) is more stable than the unmethylated DNA duplex (I) due to a more favorable enthalpic term. DNA-RNA hybrid stabilities are intermediate between RNA and DNA, closer to the respective DNA duplexes. The methylated DNA-RNA hybrids are slightly more stable than the respective unmethylated DNA-RNA hybrids (compare IV with VI, and V with VII). Large enthalpic stabilization is associated with guanosine (rG) residues as indicated by ΔH° values for the RNA (III) and DNA-RNA hybrids V and VII.

The individual oligonucleotide strands form imperfect selfcomplementary duplexes due to the (G)_n-(C)_n commonality of the sequences. Existence of imperfect duplexes with 10 Watson-Crick base pairs, two mismatches and dangling ends was confirmed by nuclease S₁ digestion experiments for solutions containing only one strand (data not shown). Although these imperfect duplexes do display thermal denaturation profiles, melting temperatures are significantly lower than those of the non-selfcomplementary duplexes in Table I. The duplexes in Table I display

Figure 2: CD spectra of the duplex DNA, RNA and DNA·RNA hybrids in 50 mM NaCl, 6 mM Na_2HPO_4 , 2 mM NaH_2PO_4 , 0.1 mM Na_2EDTA , pH 7 at -25°C . Spectrum labels correspond to the sequences in Table I. Molar ellipticity values are in terms of base pairs. (A) Unmethylated DNA, RNA and the corresponding hybrids. (B) Methylated DNA, RNA and the corresponding hybrids.

single phase thermal denaturation profiles (Figure 1). Nuclease S_1 digestion of the duplexes in Table I produced no evidence of imperfect duplex or hairpin formation.

Circular Dichroism Spectroscopy. CD spectroscopy is a powerful tool for determining A- or B-form structures of nucleic acids. A hallmark of A-form RNA is a CD spectrum with large positive intensity at ~ 270 nm, slight negative intensity at ~ 240 nm and large negative intensity at 210 nm. In contrast, B-form DNA of identical sequence exhibits a CD spectrum with less positive intensity at ~ 270 nm, greater negative intensity at ~ 240 nm and near zero intensity at 210 nm (12,13).

CD spectra of the unmethylated DNA (I), RNA (III) and the corresponding DNA·RNA hybrids (IV,V) are shown in Figure 2A. The RNA spectrum exhibits a small negative band at 305 nm, a large positive band at 268 nm and a large negative band at 210 nm. In comparison to the RNA, the DNA spectrum lacks the negative band at 305 nm, exhibits a less intense positive band centered at 266 nm, and approaches zero at 210 nm. Observed DNA and RNA spectra closely agree with respective B-DNA and A-RNA spectra calculated according to the nearest neighbor method of Gray et al. (12,13), which is based upon an empirical set of polynucleotide spectra. Spectra of the DNA·RNA hybrids (IV,V) more closely resemble the spectrum of RNA than DNA. Hybrid spectra display negative bands at 305 nm of less intensity than the RNA spectrum. The hybrid spectra exhibit positive bands at 267 nm that are intermediate intensity between the DNA and RNA spectra. The spectrum of the deoxyguanosine (dG) rich hybrid (IV) displays greater positive intensity at 267 nm than the guanosine (rG) rich hybrid (V). The DNA·RNA hybrid and

Figure 3: CD spectra of the (A) unmethylated and (B) methylated DNA duplexes in trifluoroethanol solutions. The solutions contain 2 mM NaCl, 0.1 mM Na₂EDTA and the indicated amount (v/v) of 2,2,2-trifluoroethanol at ~25°C. Spectrum labels correspond to the sequences in Table I. Molar ellipticity values are in terms of base pairs.

RNA spectra are identical over 250–220 nm. The negative intensities of the hybrid CD bands at 215 nm are intermediate between the DNA and RNA spectra with hybrid IV exhibiting greater intensity than hybrid V.

CD spectra of the methylated DNA (II), RNA (III) and the corresponding DNA·RNA hybrids (VI,VII) are shown in Figure 2B. The methylated DNA spectrum (Figure 2B) resembles the unmethylated DNA spectrum (Figure 2A) but with reduced positive intensity at 262 nm. The methylated DNA·RNA hybrids exhibit spectra nearly identical to the respective unmethylated hybrids (compare IV with VI, and V with VII).

B to A Transitions in Trifluoroethanol. DNA is known to adopt an A-form conformation in concentrated trifluoroethanol (TFE) solutions as detected by CD (14). CD spectra were recorded for the unmethylated and methylated DNA in 60, 65 and 70% TFE, conditions under which the oligonucleotides remain double-stranded (Figure 3). Both DNA duplexes exhibit spectra with increasing positive and negative ellipticity at 270 and 210 nm, respectively, with increasing TFE concentration. The spectrum of the unmethylated DNA in 70% TFE (Figure 3) closely resembles the RNA spectrum (Figure 2) apparently due to a TFE-induced B- to A-DNA transition. The methylated DNA in 70% TFE exhibits a spectrum similar to the RNA spectrum but with reduced positive intensity at 270 nm. Methylation of deoxycytidine does not significantly affect the trifluoroethanol concentration required to induce an A-type conformation in this DNA sequence as judged by CD data at 210 nm. The range of TFE concentration necessary to induce B to A transitions in these DNA duplexes is approximately the same range required for other DNA oligomers and polymers (14).

DISCUSSION

The RNA duplex is more stable than the unmethylated and methylated DNA duplexes due to a larger enthalpic term. Large enthalpic stabilization of the RNA and DNA·RNA hybrids V and VII indicates extensive intrastrand guanosine (rG) stacking in an A-form conformation. In contrast to the present results for these $(G)_n \cdot (C)_n$ duplexes, the DNA duplex $dCA_5G \cdot dCT_5G$ is more stable than the RNA duplex $rCA_5G \cdot rCU_5G$ despite larger enthalpic stabilization of the RNA (1). Sequence is a strong determinant of the relative stabilities of DNA, RNA and DNA·RNA hybrids.

Earlier literature reports supported A-form conformations for DNA·RNA hybrids (21-24). However more recent reports demonstrate that not all hybrid sequences are predominantly A-form, especially $(rA)_n \cdot (dT)_n$ sequences (25-30). Gray & Ratliff (31) recorded CD spectra of poly(dA-dC)·poly(dG-dT), poly(rA-rC)·poly(rG-rU) and the corresponding hybrids, and concluded the conformations of the hybrid polynucleotides were intermediate between A-RNA and B-DNA. Nuclear magnetic resonance investigations have also demonstrated that deoxynucleotides can retain significant B-form characteristics in DNA·RNA hybrids (32,33). In the present study, CD spectra of the DNA·RNA hybrids resemble the RNA spectrum, although significant differences at 305, 265 and 215 nm indicate deviation of DNA·RNA hybrid conformation from a strict RNA type A-form structure. Thermodynamic properties of the DNA·RNA hybrids are intermediate between DNA and RNA. Differences in thermodynamic properties between hybrids of identical base sequence provide evidence for sugar (ribose or deoxyribose) dependent DNA·RNA hybrid conformations distinct from either pure B- or A-form. The CD and thermodynamic results support predominantly A-form DNA·RNA hybrid conformations, resembling but not identical to A-RNA.

Differences in CD spectra between unmethylated and methylated DNA may indicate different DNA conformations, although a variance in transition dipole moments for cytosine and 5-methylcytosine must be considered. However, nearly identical CD spectra for the corresponding unmethylated and methylated DNA·RNA hybrids suggest any difference in transition dipole moments is minor. Comparable ΔH° values for the unmethylated and methylated DNA duplexes support similar base stacking. Differences in CD spectra of unmethylated and methylated DNA may reflect dielectric contributions from different salt/hydration structures associated with two DNA helices of similar but not identical B-type conformations. Combined structural and thermodynamic results support B-form conformations for the unmethylated and methylated DNA.

Previous studies of similar DNA oligonucleotides demonstrated the propensity of $d(G)_n \cdot d(C)_n$ sequences to adopt A-form conformations under dehydrated crystal conditions (2). In contrast, present CD results and trifluoroethanol titrations indicate B-form conformations in dilute solutions for both the unmethylated and methylated $d(G)_n \cdot d(C)_n$ DNA duplexes. A- to B-form transitions of $d(G)_n \cdot d(C)_n$ sequences apparently involve dehydration and/or intermolecular packing, although

the variety of crystal lattices observed for A-DNA argues against intermolecular packing as a major force (2).

Methylation of deoxycytidine did not significantly affect either the stabilities or conformations of the DNA-RNA hybrids or strongly influence B-A transitions of the DNA duplexes in trifluoroethanol solutions. However, methylation may influence dehydration of the helix (2) and play a significant role in experiments performed under conditions of reduced water activity (5,6).

ACKNOWLEDGEMENTS

The author was supported by EHS Postdoctoral Training Grant 5 T32 ES07075 -10. The investigation was performed in the laboratory of Prof. I. Tinoco, Jr., whose funding was sponsored by Department of Energy grant #DE-FG03-86ER-60406 and USPH grant #GM 10840. The author would like to thank Prof. Tinoco for his support and helpful discussion. The *E. coli* strain BL21 containing the plasmid pAR1219 was a generous gift of Dr. F. W. Studier.

REFERENCES

1. Martin, F. H. and Tinoco, I., Jr. (1980) *Nucleic Acids Res.* **10**, 2295-2299.
2. Frederick, C. A., Saal, D., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., and Rich, A. (1987) *Biopolymers* **26**, S145-S160.
3. Nishimura, Y., Torigoe, C., and Masamichi, T. (1985) *Biopolymers* **24**, 1841-1844.
4. Benevides, J. M., Wang, A. H. -J., Rich, A., Kyogoku, Y., van der Marel, G. A., van Boom, J. H., and Thomas, G. J., Jr. (1986) *Biochemistry* **25**, 41-50.
5. Nishimura, Y., Torigoe, C., and Tsuboi, M. (1986) *Nucleic Acid Res.* **14**, 2737-2748.
6. Sarma, M. H., Gupta, G., and Sarma, R. H. (1986) *Biochemistry* **25**, 3659-3665.
7. Adams, R. L. P., and Burdon, R. H. (1985) *Molecular Biology of DNA Methylation*, Springer-Verlag, New York.
8. Morgan, A. R. and Wells, R. D. (1968) *J. Mol. Biol.* **37**, 63-80.
9. Zimmerman, S. B., Cohen, G. H., and Davies, D. R. (1975) *J. Mol. Biol.* **92**, 181-192.
10. Johnson, D. and Morgan, A. R. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1637-1641.
11. Lee, J. S., Woodsworth, M. L., Latimer, L. J. P., and Morgan, A. R. (1984) *Nucleic Acids Res.* **12**, 6603-6614.
12. Gray, D. M., Liu, J. -J., Ratliff, R. L., and Allen, F. S. (1981) *Biopolymers* **20**, 1337-1382.
13. Gray, D. M., Hamilton, F. D., and Vaughan, M. R. (1978) *Biopolymers* **17**, 85-106.
14. Minchenkova, L. E., Schyolkina, A. K., Chernov, B. K., and Ivanov, V. I. (1986) *J. Biomol. Struct. Dyn.* **4**, 463-476.
15. Milligan, J. F., Groebe, D. R., Witherell, G. W., Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15**, 8783-8798.
16. Davanloo, P., Rosenberg, A. H., Dunn, J. J., and Studier, F. W. (1984) *Proc. Natl. Acad. Sci.* **81**, 2035-2039.
17. Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 197, CRC Press, Cleveland, OH.
18. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 589, CRC Press, Cleveland, OH.
19. Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill Book Co.
20. Gralla, J. and Crothers, D. M. (1973) *J. Mol. Biol.* **78**, 301-319.
21. Milman, G., Langridge, R., Chamberlin, M. J. (1967) *Proc. Natl. Acad. Sci. U.S.A.* **57**, 1804-1810.
22. O'Brien, E. J., and MacEwan, A. W. (1970) *J. Mol. Biol.* **48**, 243-261.
23. Selsing, E., Wells, R. D., Early, T. A., and Kearns, D. R. (1978) *Nature* **275**, 249-250.

24. Pardi, A., Martin, F. H., and Tinoco, I., Jr. (1981) *Biochemistry* **20**, 3986-3996.
25. Gupta, G., Sarma, M. H., and Sarma, R. H. (1985) *J. Mol. Biol.* **186**, 463-469.
26. Shindo, H., and Matsumoto, U. (1984) *J. Biol. Chem.* **259**, 8682-8684.
27. Zimmerman, S. B., and Pfeiffer, B. H. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 78-82.
28. Krueger, W. C., Li, L. H., Moscovitz, A., Prairie, M. D., Petzold, G., and Swenson, D. H. (1985) *Biopolymers* **24**, 1549-1572.
29. Katahira, M., Nishimura, Y., Tsuboi, M., Sato, T., Mitsui, Y., Iitaka, Y. (1986) *Biochim. Biophys. Acta* **867**, 256-267.
30. Steely, H. T., Jr., Gray, D. M., and Ratliff, R. L. (1987) *Nucleic Acid Res.* **14**, 10071-10090.
31. Gray, D. M., and Ratliff, R. L. (1975) *Biopolymers* **14**, 487-498.
32. Mellema, J. -R., Haasnoot, C. A. G., van der Marel, G. A., Wille, G., van Boeckel, C. A. A., van Boom, J. H., and Altona, C. (1983) *Nucleic Acids Res.* **11**, 5717-5738.
33. Reid, D. G., Salisbury, S. A., Brown, T., Williams, D. H., Vasseur, J. -J., Rayner, B., and Imbach, J. -L. (1983) *Eur. J. Biochem.* **135**, 307-314.