Unpaired Cytosine in the Deoxyoligonucleotide Duplex $dCA_3CA_3G \cdot dCT_6G$ Is Outside of the Helix†

Kathleen M. Morden,* Y. Gloria Chu, Francis H. Martin,‡ and Ignacio Tinoco, Jr.

ABSTRACT: The conformation and thermodynamics were determined for the oligonucleotide duplex $dCA_3CA_3G + dCT_6G$ which contains an extra, unpaired cytosine residue. The conformation was investigated by using nuclear magnetic resonance to observe both the aromatic base protons and the base-pairing imino protons. The nuclear Overhauser effects between imino protons on either side of the unpaired cytosine show that the cytosine is extrahelical; the base is not stacked in the helix. The temperature dependence of the chemical shifts of the aromatic base protons supports this conclusion. The observed nuclear Overhauser effects and a large upfield shift in the imino proton resonance of A·T base pair 4, adjacent to the extrahelical cytosine, indicate there is a perturbation in the base-pair structure of this A·T base pair as well as a change in the overlap with adjacent base pairs. Thermodynamic parameters were determined for this duplex as well as for $dCA_6G + dCT_6G$. The extrahelical cytosine causes a decrease in the duplex melting temperature of 15 °C, at a concentration of 200 μM per single strand. The standard free energy for duplex formation at 25 °C is 2.9 kcal mol^{-1} more positive for the duplex with the extra cytosine. This decrease in stability for the perturbed helix is caused by an unfavorable change in enthalpy and a possibly favorable change in the entropy for duplex formation.

The ability to predict the structure of RNA or DNA is based on an understanding of the energetics involved in forming that structure. Although there are data available on the thermodynamics of double-strand formation in RNA (Borer et al., 1974), there are limited data for DNA. Predictions for the thermodynamics of duplex formation in DNA helices must, presently, be done by using the RNA values. Relatively little is known about the thermodynamics for the formation of secondary structures in RNA, such as hairpins (Uhlenbeck et al., 1973; Gralla & Crothers, 1973), mismatched bases, or base bulges (Fink & Crothers, 1972), and even less is known for DNA. Previous investigations have probed the effects of base-pair mismatches in synthetic polynucleotide helices [for a review, see Lomant & Fresco (1974)]. These studies used optical properties to determine if the polynucleotides formed helices containing noncomplementary base oppositions or if the bases, either stacked or not stacked in the helix, had no base opposition on the cross-strand. Evidence for looped-out bases in polynucleotides has been obtained by using circular dichroism (Gray et al., 1980) and photodimerization (Lomant & Fresco, 1973; Evans & Morgan, 1982). The free energy for a bulge defect has been measured for poly(A,A*) + poly(U), where A* is the N-1 oxide of adenine (Fink & Crothes, 1972). The stabilities for base bulges may be valuable in investigating mechanisms of mutagenesis, in particular frame-shift mutagenesis (Streisinger et al., 1966).

Nuclear magnetic resonance (NMR)[1] has been used to study the conformation of a deoxyoligonucleotide containing an extra adenine, and the base was found to be stacked in the helix (Patel et al., 1982). In this paper, we demonstrate the use of nuclear magnetic resonance to study the deoxyoligonucleotide duplex $dCA_3CA_3G + dCT_6G$ in comparison with the parent duplex $dCA_6G + dCT_6G$. By analyzing the temperature dependence of the chemical shifts of aromatic base protons and the nuclear Overhauser effects (NOEs) between imino protons, we have determined that the unpaired cytosine is not stacked in the helix. UV absorption melting curves have been measured and analyzed, for these helices, to obtain the standard free energies, enthalpies, and entropies of duplex formation.

Materials and Methods

All of the oligonucleotides were synthesized by one of two techniques; dCT_6G, dCA_5G, and dCT_5G were synthesized by using diester solution techniques (Khorana, 1968); dCA_3CA_3G and dCA_6G were synthesized by using phosphoramidite solid-support techniques (Beaucage & Caruthers, 1981; Matteucci & Caruthers, 1981). They were desalted several times with Bio-Gel P-2 columns and then lyophilized for storage. Extinction coefficients were estimated from the coefficients for mononucleotides and dinucleotide phosphates (Fasman, 1975). The extinction coefficients per mole of strand at 260 nm and 25 °C for dCA_3CA_3G, dCA_6G, dCA_5G, dCT_5G, and dCT_6G are 98×10^3, 91×10^3, 79×10^3, 58×10^3, and 66×10^3 M^{-1} cm^{-1}, respectively. Samples used for thermodynamic studies were lyophilized from H_2O and then dissolved in a 0.01 M phosphate buffer, pH 7, containing 0.1 mM EDTA and 1 M NaCl. The optical melting curves were obtained by using a procedure described previously (Nelson et al., 1981).

NMR samples were prepared with buffer containing 28 mM phosphate, 0.1 mM EDTA, and 0.2 M NaCl. D_2O samples, containing oligonucleotide and buffer, were lyophilized 3 or 4 times from 99.98% D_2O (Bio-Rad) and finally dissolved in 100% D_2O (Bio-Rad) in a glovebag under a N_2 atmosphere. H_2O samples were prepared following the same procedure but

† From the Department of Chemistry and the Laboratory of Chemical Biodynamics, University of California, Berkeley, Berkeley, California 94720. *Received May 19, 1983.* This work was supported by National Institutes of Health Grant GM 10840 and by the U.S. Department of Energy Office of Energy Research, under Contract 03-82ER60090.000. K.M.M. and Y.G.C. were supported by NIEHS Training Grant ES 07075. We also thank the Stanford Magnetic Resonance Laboratory (supported by NSF Grant GP 26633 and NIH Grant RR 00711) for the use of the HXS 360-MHz facilities.

‡ Present address: Applied Molecular Genetics, Inc., Thousand Oaks, CA 91320.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; TSP, sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$; EDTA, ethylenediaminetetraacetic acid.

0006-2960/83/0422-5557$01.50/0

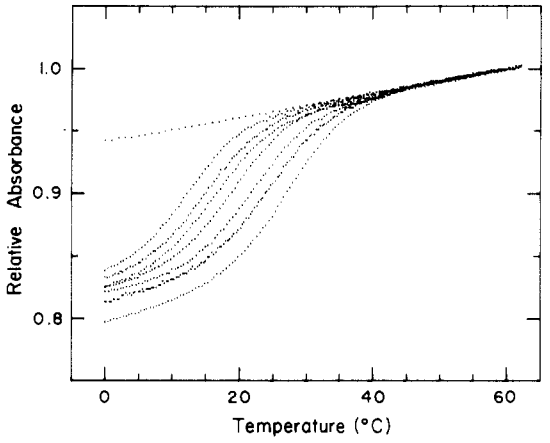

FIGURE 1: Melting curves of $dCA_3CA_3G + dCT_6G$ in 1 M NaCl with concentrations ranging from 0.01 to 0.9 mM in the single strand. The curves are normalized to 1 at 60 °C. The upper curve is the single-strand melting curve as described in the text.

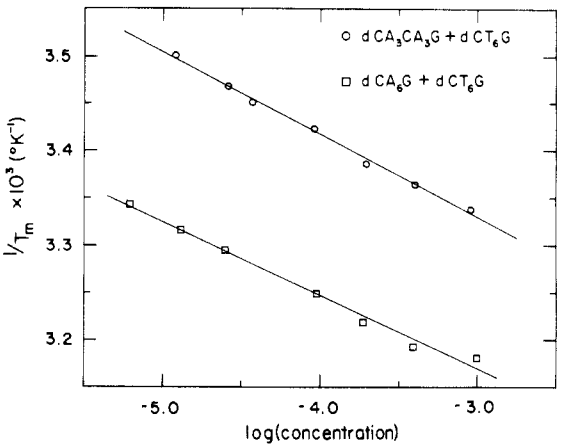

FIGURE 2: $1/T_m$ vs. log C_0 plots for (○) $dCA_3CA_3G + dCT_6G$ and (□) $dCA_6G + dCT_6G$.

were lyophilized from H_2O and prepared outside the glovebag. The water samples contained 10–15% D_2O. All H_2O samples were adjusted to be pH 7 ± 0.1, and D_2O samples were adjusted to pD 7 ± 0.1 (measured as 6.6 on a pH meter). Spectra taken in H_2O were obtained by using a Redfield 21412 pulse sequence (Refield et al., 1975). NOE difference spectra were obtained with an irradiation time of 0.5 s and a repetition rate of ~0.8 s, using a Bruker WM500 instrument at the University of Washington, Seattle. Other NMR spectra were obtained by using the 360-MHz instrument at the Stanford Magnetic Resonance Laboratory or the 500-MHz instrument at the University of California, Davis, NMR facility, as indicated. All spectra were referenced to the internal standard TSP.

Thermodynamic Results

The optical melting curves of $dCA_3CA_3G + dCT_6G$ taken at seven different concentrations are shown in Figure 1. The curves are normalized at 60 °C to an absorbance of 1. The upper curve is the single-strand melting curve, determined from a weighted average of melting curves from the two individual strands. The curves show an increase in hypochromicity as the concentration increases; this has been attributed to aggregation of the double strands (Nelson et al., 1981). The lower base line of the double-strand melt is difficult to determine because this duplex melts at approximately room temperature. A nonlinear least-squares program based on a two-state model was used to fit $\Delta H°$, $\Delta S°$, and the lower base-line slope to the normalized melting curve. The slope for the most concentrated sample was determined to be 1.04 $\times 10^{-3}$/°C. This slope was used to determine low-temperature base lines for the other concentrations. The melting curves were analyzed to determine the melting temperature by using the slope of the lower base line previously discussed and the experimental single-strand melting curve. Figure 2 shows the $1/T_m$ vs. log C_0 plot used to determine $\Delta H°$ and $\Delta S°$, where C_0 is the total concentration of each of the oligomer strands (Nelson et al., 1981). These results are given in Table I. The same procedure was used for $dCA_6G + dCT_6G$ with the exception that the slope of the low-temperature base line for the second most concentrated sample was used to obtain low-temperature base lines for the other curves. This slope was used because it gave base lines that fit the curves better than those determined from the slope of the most concentrated sample. The slope used was 1.55×10^{-3}/°C. Values obtained by using the low-temperature base-line slope of the most concentrated sample were within 3% of the values reported here. The $1/T_m$ vs. log C_0 plot from these data is also shown in Figure 2. Thermodynamic parameters are reported in Table I for $dCA_6G + dCT_6G$ and for $dCA_5G + dCT_5G$ (Nelson et al., 1981).

Table I: Thermodynamic Parameters from the $1/T_m$ vs. log C_0 Plot

	T_m (°C, 200 μM)	$\Delta G°$ (kcal mol^{-1}, 25 °C)	$\Delta H°$, (kcal mol^{-1})	$\Delta S°$ (eu)
$dCA_3CA_3G + dCT_6G$	22 ± 2	−4.9 ± 0.3	−53 ± 2	−161 ± 5
$dCA_6G + dCT_6G$	37 ± 3	−7.8 ± 0.6	−59 ± 3	−172 ± 8
$dCA_5G + dCT_5G$[a]	32	−6.6	−47 ± 2	−136 ± 5

[a] Data from Nelson et al. (1981).

NMR Results and Discussion

Before discussing the thermodynamic parameters, it is helpful to know more about the conformation of the oligonucleotide duplex with the extra base. The conformation has been studied by NMR, both by using the NOE to probe proton–proton distances in the duplex and by investigating the temperature dependence of the chemical shifts of aromatic protons.

Spectra of the base-pairing imino protons from $dCA_5G + dCT_5G$, $dCA_3CA_3G + dCT_6G$, and $dCA_6G + dCT_6G$ are shown in Figure 3. The assignments of the resonances from base pairs 1, 7, and 8 in the $dCA_3CA_3G + dCT_6G$ duplex were made in analogy to the assignments for the $dCA_5G + dCT_5G$ duplex determined by Pardi et al. (1981). The assignments of the other imino resonances were determined by imino–imino NOE, as discussed below.

The nuclear Overhauser effect is a process where spin can be transferred from one nucleus to another through space (Noggle & Schirmer, 1971). The process depends inversely on the sixth power of the distances between the nuclei. The NOE is also affected by several other factors. The intensity is dependent on the correlation time of the molecule, which depends on the length of the duplex. The NOE can also be in competition with other possible relaxation processes, such as exchange with water. We have used this technique to assign resonances and to obtain information about the conformation around the extra cytosine. In a regular helix, the distance between imino protons on neighboring base pairs is less than 4 Å, a reasonable distance to see NOEs. In A·T base pairs, there is also an NOE from the imino proton to the adenine-(H2) on the same base pair, and this NOE has been used to

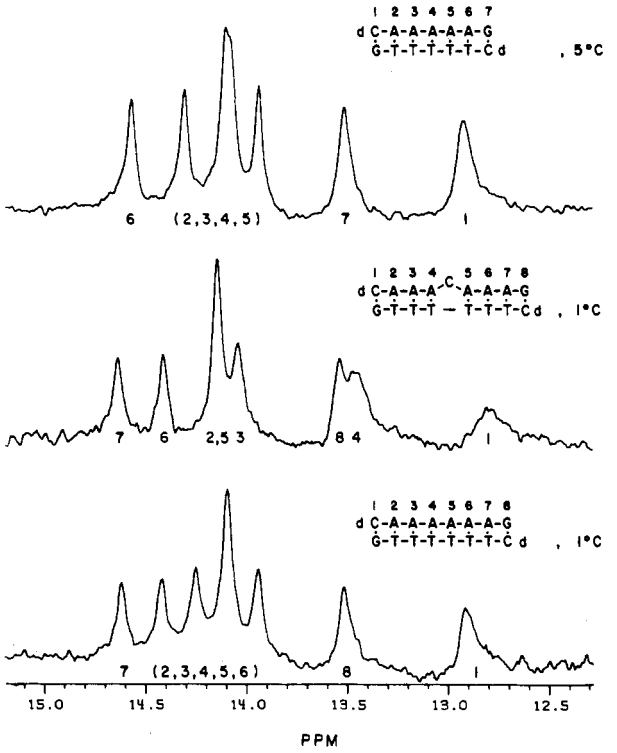

FIGURE 3: 360-MHz ^1H NMR spectra of the base-pairing imino protons from $dCA_5G + dCT_5G$, $dCA_3CA_3G + dCT_6G$, and $dCA_6G + dCT_6G$. Samples were 1 mM per single strand in H_2O. Spectra are the sum of 400–600 accumulations.

assign imino resonances to A·T (or A·U) vs. G·C base pairs (Sanchez et al., 1980; Hare & Reid, 1982). The NOE difference spectra for $dCA_3CA_3G + dCT_6G$ are shown in Figure 4 along with a Redfield spectrum showing all the resonances for reference. All the NOEs are negative; there is a decrease in intensity on irradiation. In spectra b–f, the NOE to a sharp line in the aromatic region, upfield from the imino region, indicates the saturated resonances are due to A·T base pairs. Given the above information, we can now interpret the NOE spectra shown in Figure 4.

The assignment procedure consists of selectively saturating the imino resonances and detecting the NOE. We expect to see an NOE from an imino proton to the imino protons on the neighboring base pairs. In Figure 4b, saturating the imino resonance from base pair 2 should show an NOE to resonances from base pairs 1 and 3. Base pair 1 is one of the terminal G·C base pairs, and an inversion–recovery experiment shows the recovery time of this resonance is less than 0.05 s. Thus, observation of an NOE to this resonance is not likely. As expected, only a peak from base pair 3 is observed in the difference spectrum in Figure 4b. It should be noted that this resonance could also be due to partial saturation as it is close to the resonance for base pairs 2 and 5. Saturation of the resonance from base pair 5 shows an NOE to base pairs 4 and 6, as expected. NOEs from base pair 3 to base pair 4 and possibly base pair 2 are shown in Figure 4c. The peak from base pairs 2 or 5 may again be due to partial saturation. An NOE is observed from the imino resonance for base pair 4 to base pairs 3 and 5 in Figure 4d. In Figure 4e, an NOE is observed to base pairs 5 and 7 when the imino resonance from base pair 6 is saturated. In Figure 4f, an NOE is observed from base pair 7 to base pair 6. No NOE is observed to base pair 8, because this resonance has a rapid relaxation rate. This is also a problem when trying to observe an NOE from the imino resonance of base pair 8, and as shown in Figure 4g, none are observed in the low-field region. From the observation of an imino–imino NOE, described above, we have assigned the low-field resonances as given in Figure 4a.

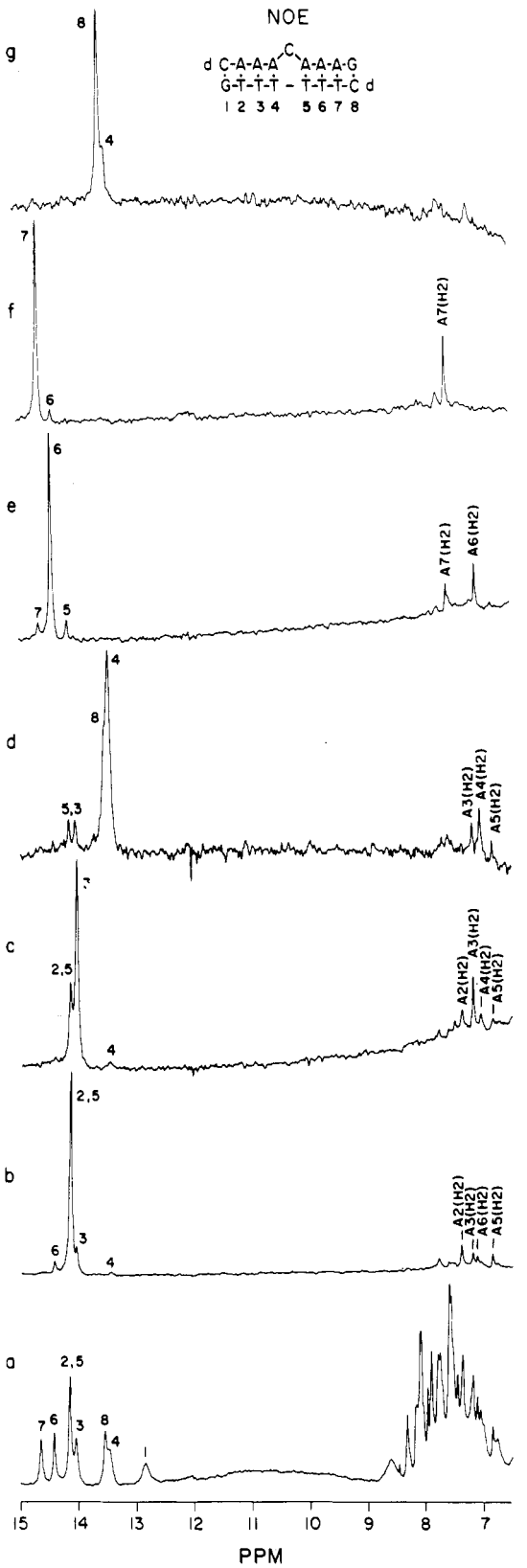

FIGURE 4: 500-MHz ^1H NOE difference spectra from $dCA_3CA_3G + dCT_6G$. Spectra b–g are the difference of 6000 scans irradiating off resonance and 6000 scans irradiating on resonance. Spectrum a was obtained by using a Redfield pulse with no irradiation. The sample was 1 mM per single strand in 90% H_2O.

All the A(H2) resonances can be assigned knowing the assignment of the imino resonance from base pair 7 and the

fact that there is a strong NOE between the imino proton and the A(H2) resonance in an A·T base pair. The assignments for the A(H2) resonances are in good agreement with the assignments for the imino resonances as shown by the following. In Figure 4b, we observe an NOE from the imino proton on base pair 2 to its own H2 proton, A2(H2), and to A3(H2) on the adjacent base pair. In the same spectrum, we observe an NOE from the imino proton in base pair 5 to its own H2 and to the H2 on base pair 6. This general trend is seen in other A·T base pairs in this duplex. Saturation of the imino proton resonance T(H3) gives an NOE to the A(H2) which it is base paired with and to the A(H2) in the 3′ direction from the first A(H2). This is the trend we expect to see, based on distances calculated for a regular DNA helix. Thus, we observe, in Figure 4e, an NOE from the imino proton on base pair 6 to the H2 on that base pair and the H2 on base pair 7. In the spectrum shown in Figure 4f, an NOE is observed only from the imino proton on base pair 7 to its own H2 because base pair 8 is a G·C base pair and does not contain an H2 proton. In Figure 4c, saturation of the imino resonance from base pair 3 gives an NOE to A3(H2) and A4(H2), as expected. We also observe an NOE to A2(H2) and A5(H2) due to partial saturation of imino resonances from base pairs 2 and 5, respectively. The only example that does not follow the trend is shown in Figure 4d. Saturation of the imino resonance from base pair 4, one of the base pairs adjacent to the unpaired C, shows an NOE to A4(H2) and A5(H2), as expected, but in addition, we observe an NOE to the A3(H2), which is the neighboring base pair on the 5′ side. This implies that the distances between the imino proton from base pair 4 and the H2 proton on base pair 3 is less than the analogous distance in other parts of the helix, where a similar NOE is not observed. This decrease in distance could be achieved by a decrease in winding angle between these two base pair (3 and 4) or a translation of one base pair relative to the other.

The observation of an NOE from base pair 4 to base pair 5 characterizes the conformation of the unpaired cytosine. If the extra cytosine were stacked in the helix, the distance between base pairs 4 and 5 would be almost 7 Å, and no imino–imino NOE would be observed. However, as shown in Figure 4d, we do observe an NOE from the imino proton on base pair 4 to the imino proton on base pair 5. Thus, the unpaired cytosine cannot be stacked in the helix. This is corroborated by the observation of an NOE from the imino proton on base pair 4 to the A5(H2), as shown in Figure 4d.

There are several other potential assignment schemes for the imino protons. However, these can be ruled out, either because of inconsistencies in the imino–imino NOEs or because of inconsistencies in the imino–A(H2) NOEs. There is one assignment scheme for the imino protons, other than the one presented, that cannot be ruled out by any obvious inconsistencies. This alternate assignment consists of exchanging the assignments for the imino resonances from base pairs 3 and 4, compared to those shown in Figure 4. Assignment of the A(H2) resonances can be made consistent with this imino assignment by exchanging the A(H2) assignments for base pairs 3 and 4 as well as base pairs 2 and 5. With this alternate assignment scheme, an NOE is observed from the imino protons of base pairs 3, 4, and 5 to their own A(H2) and to the A(H2)'s of the base pairs on both sides of it. This is not consistent with the expected NOE from a regular DNA helix of this sequence, but the assignment cannot be ruled out completely on this basis. There is an NOE observed from the imino proton of base pair 4 to the A5(H2) and from the imino proton of base pair 5 to the A4(H2), so this alternate assignment scheme is also consistent with the extra cytosine being out of the helix.

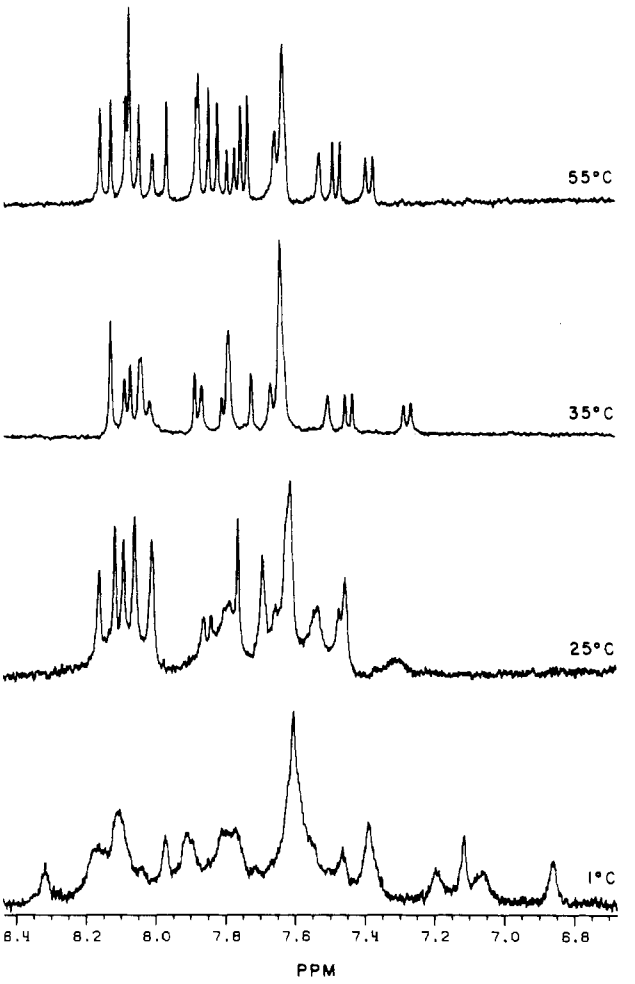

FIGURE 5: 360-MHz ^{1}H NMR spectra of $dCA_3CA_3G + dCT_6G$ as a function of temperature. The sample is 1 mM per single strand in D_2O.

More evidence for the cytosine being extrahelical is obtained by monitoring the aromatic proton resonances and their temperature dependence. When the aromatic bases are placed in a magnetic field, there is a local field induced; thus, a proton directly above or below the base will be shielded from the applied field, and a proton in the plane of the base will be deshielded. For an aromatic proton in a double strand, such ring-current effects will be mainly due to the base pairs above and below the proton and from the base it is base paired to on the cross-strand. The equilibrium between double-strand and single-strand formation changes as a function of the temperature; thus, the environment around a proton changes as the temperature changes. This is observed in the NMR spectrum as a change in the chemical shift and can be seen in Figure 5. At low temperature, the resonances are broad due to exchange between the single-strand and double-strand environment. This broadening makes assignment of the resonances more difficult. Assignment of the cytosine doublets at low temperatures was accomplished by using a double-quantum pulse sequence (spectrum shown at the top of Figure 6). This pulse sequence creates double-quantum coherence and allows observation of only those resonances which have double-quantum transitions (Hore et al., 1982a,b). The region of the spectrum between 6.5 and 8.5 ppm contains only singlets and doublets and becomes simplified to contain only doublets when the double-quantum pulse sequence is used. It is

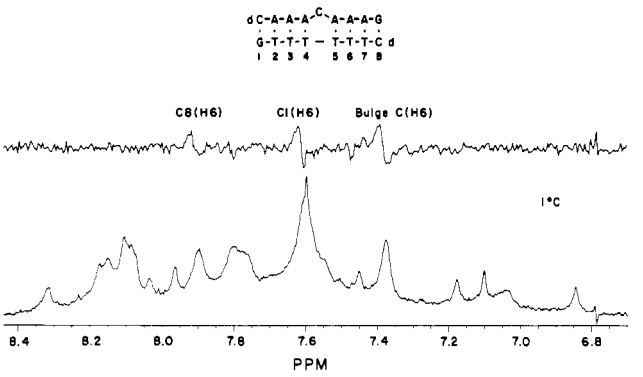

FIGURE 6: 500-MHz ^1H NMR spectra of $dCA_3CA_3G + dCT_6G$. The top spectrum is the sum of 10 400 scans using the double-quantum pulse sequence, as discussed in the text. The bottom spectrum is the single-quantum spectrum of the same sample. The sample was 0.5 mM per single strand in D_2O.

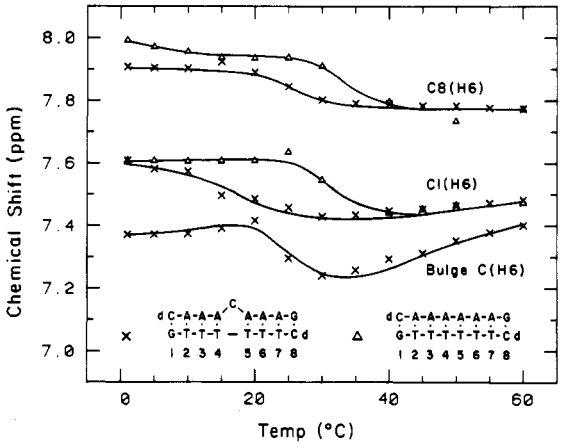

FIGURE 7: Chemical shift vs. temperature curves for the C(H6) protons in (×) $dCA_3CA_3G + dCT_6G$ and (Δ) $dCA_6G + dCT_6G$. The samples were 1 mM per single strand in D_2O.

characteristic of this pulse sequence that the two peaks in the doublet will be 180° out of phase, as seen in the top of Figure 6. The low-temperature chemical shifts of the cytosine doublets were obtained from the spectrum in Figure 6. The high-temperature chemical shifts of the cytosine doublets in $dCA_3CA_3G + dCT_6G$ can be assigned by comparison with $dCA_5G + dCT_5G$ (Pardi et al., 1981) and with $dCA_6G + dCT_6G$. In Figure 7, the chemical shifts as a function of temperature are compared for $dCA_3CA_3G + dCT_6G$ and $dCA_6G + dCT_6G$. The only major difference between the curves for the two terminal cytosines is that the curves for the $dCA_6G + dCT_6G$ resonances shift to higher temperature due to the increased melting temperature of this duplex. The curve at the bottom of Figure 7 can now be assigned to the extra cytosine in $dCA_3CA_3G + dCT_6G$. Because this duplex has non-self-complementary strands, the single-strand chemical shifts as a function of temperature can also be obtained. The chemical shift vs. temperature curves of dCA_3CA_3G, dCT_6G, and the double strand are shown in Figure 8. As duplex formation occurs, with decreasing temperature, both the C1-(H6) and the C8(H6) show an increase in the chemical shift, a downfield shift, due to the deshielding effect of the bases on the cross-strand. This deshielding will not be present for the extra cytosine, regardless of whether or not it is stacked in the helix, because there is no base directly across from it on the cross-strand. Chemical shift calculations done for the extra cytosine, assuming it is in the base stack and all bases are stacked in a B-form geometry around it, predict an upfield shift on going from high to low temperature. If the cytosine

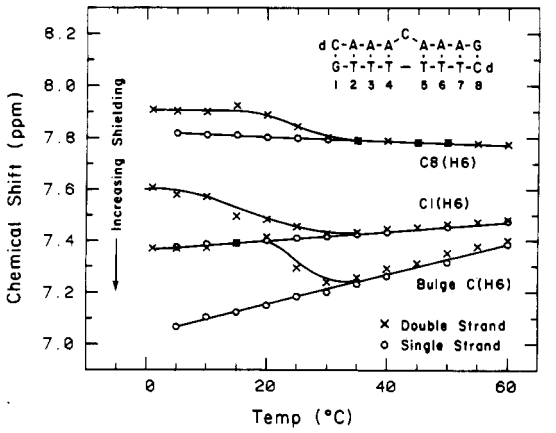

FIGURE 8: Chemical shift vs. temperature curves for the C(H6) protons in $dCA_3CA_3G + dCT_6G$. Both the (×) double-strand and the (O) single-strand curves are shown. The sample was 1 mM per single strand in D_2O.

were stacked in the helix, it would be shielded by the A·T base pairs around it. However, what is observed in Figure 8 is a large (0.3 ppm) increase in the chemical shift for this resonance, relative to the chemical shift in the single strand. We also note that the chemical shift of the extra cytosine at high temperature, where the bases are partially unstacked, is almost identical with the shift at low temperature, where the bases should be stacked. If the unpaired cytosine were extrahelical, then the environment around that base could be similar to the environment in the unstacked or single-strand state. This chemical shift behavior suggests that the extra cytosine is not stacked in the helix.

With the imino proton resonances assigned, the spectra of the imino protons for $dCA_3GA_3G + dCT_6G$ and $dCA_6G + dCT_6G$ can be compared as shown in Figure 3. The largest change in chemical shift is for base pair 4 which has shifted upfield by at least 0.4 ppm. Some of this shift could be due to the increased overlap with A·T base pair 3, as discussed previously. However, most of the ring-current effect due to this A·T base pair is already present in $dCA_6G + dCT_6G$, so 0.4 ppm is a large shift to be explained by only an improved overlap. We propose that some of the upfield shift is due to a weakened or lengthened hydrogen bond in A·T base pair 4. Work done by Wagner et al. (1983) on proteins indicates a strong correlation between H-bond length and chemical shift. Also, the resonance from A·T base pair 4 is broader than the other A·T resonances and may indicate faster exchange with water. This increase in exchange rate would be expected if the hydrogen bond were weaker. Saturation recovery experiments to determine these exchange rates are currently being done in our laboratory.

Discussion of Thermodynamics

We can now discuss the thermodynamic results in view of the result that the cytosine is extrahelical. It should be emphasized that the thermodynamic analysis was done on the basis of a two-state model. Table I compares the thermodynamic parameters for all three helices. The melting temperature, at 200 μM per single strand, decreases by 15 °C when we introduce the perturbation of the extra cytosine. It is decreased 10 °C over the duplex with one less A·T base pair and no perturbation. Another indication of the destabilization is that the $\Delta G°$(25 °C) of double-strand formation is not as energetically favorable for the duplex with the unpaired cytosine: a 2.9 kcal mol^{-1} destabilization from the parent duplex and a 1.7 kcal mol^{-1} destabilization compared to the duplex with one less A·T base pair. In RNA, there is a favorable

enthalpy and unfavorable entropy change on going from a single-strand to a double-strand state (Borer et al., 1974). This trend is observed when comparing $dCA_6G + dCT_6G$ to $dCA_5G + dCT_5G$. In this case, we have one more A·T base pair, which decreases the enthalpy by 12 ± 5 kcal mol^{-1} and decreases the entropy by 36 ± 13 eu. When $dCA_3CA_3G + dCT_6G$ is compared with $dCA_6G + dCT_6G$, there is an unfavorable change in enthalpy of 6 ± 5 kcal mol^{-1}. This change depends on differences in the single strands (dCA_3CA_3G vs. dCA_6G) and in the two double-strand helices. The increase in enthalpy is consistent with a loss of stacking of cytosine in the double strand relative to the single strand and the deformation of the double helix around the unpaired cytosine. There is a small, probably favorable, change in the entropy, 11 ± 13 eu, presumably explained by the freedom of the extra cytosine which is less stacked in the helix.

Several other groups have studied the effects of an extra base on the conformation and thermodynamics (Lomant & Fresco, 1974; Evans & Morgan, 1982; Gray et al., 1980; Fink & Crothers, 1972; Patel et al., 1982). Several of these studies have shown that these bases are extrahelical. Work done by Fink and Crothers on the polyribonucleotide poly(A,A*) + poly(U), where A* is the N-1 oxide of adenine, shows a destabilization in the standard free energy at 25 °C of 2.8 kcal mol^{-1} due to this perturbation. This is in good agreement with the value of 2.9 kcal mol^{-1} obtained in this paper for the deoxyribooligonucleotide. Work done by Patel et al. (1982) on the nearly self-complementary tridecamer dC-G-C-A-G-A-A-T-T-C-G-C-G indicates that the extra adenine is stacked in the helix. There is no difference in enthalpy, within experimental error, for forming the duplex with the extra adenine compared to the completely self-complementary dodecamer dC-G-C-G-A-A-T-T-C-G-C-G. The extra base and the sequence of the duplex around that base apparently determined whether the base remains in the helix or is bulged out. As adenine is the base with the strongest stacking tendency, it is reasonable that an extra adenine remains in the helix but that a cytosine surrounded by adenines is forced out of the helix. Our understanding of the sequence dependence of conformation will become clearer as different sequences containing perturbations are studied.

Conclusions

We have studied the conformation of an extra cytosine in the deoxyribonucleotide duplex $dCA_3CA_3G + dCT_6G$. Using the nuclear Overhauser effect, which depends on the inverse sixth power of the interproton distances, we find that the A·T base pairs on either side of the extra cytosine are stacked on each other. Thus, the cytosine must be extrahelical; that is, it is not stacked with neighboring base pairs in the helix. These experiments also demonstrate that the extrahelical cytosine causes local perturbations in the conformation of these adjacent A·T base pairs. The temperature dependence of the chemical shifts of the cytosine H6 doublets in this duplex also indicates that the unpaired cytosine is extrahelical.

The thermodynamic parameters were determined for this duplex and compared to parameters for $dCA_6G + dCT_6G$. The extrahelical cytosine causes a 2.9 kcal mol^{-1} decrease in the standard free energy at 25 °C and a decrease in the enthalpy of double-strand formation. Comparing the results we obtain for this sequence with results obtained on other sequences (Patel et al., 1982) indicates that stacking interactions may determine whether a base will be stacked or not stacked in the helix.

The effects of an extra base on the conformation and energetics of the duplex have implication on the mechanisms of mutagenesis. Models have been proposed for stacking interactions between frame-shift mutagens and an extrahelical base (Streisinger et al., 1966; Drake & Baltz, 1976). Several studies have demonstrated the possibility of this type of interaction (Lee & Tinoco, 1978; Helfgott & Kallenbach, 1979). The work presented here has shown how $dCA_3CA_3G + dCT_6G$ in solution forms a stable duplex with the unpaired cytosine not stacked on the neighboring base pairs in the helix, and thus this duplex would be ideal for studying the interaction between frame-shift mutagens and extrahelical bases.

Acknowledgments

We are grateful to Dr. Arthur Pardi for useful discussion as well as for generously providing data prior to publication. We thank Dr. Jerry Matson at the University of California, Davis, NMR facility for assistance in obtaining the double-quantum spectra. We also thank Dr. Dave Wemmer for taking the NOE spectra and for giving helpful advice. Support for the Bruker WM 500 NMR instrument at the University of Washington, Seattle, is given by the Murdock Charitable Trust.

Registry No. $dCA_3CA_3G·dCT_6G$, 87261-33-4; $dCA_6G·dCT_6G$, 87261-35-6; $dCA_5G·dCT_5G$, 75579-56-5.

References

Beaucage, S. L., & Caruthers, M. H. (1981) *Tetrahedron Lett. 22*, 1859–1862.

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

Drake, J. W., & Baltz, R. H. (1976) *Annu. Rev. Biochem. 45*, 11–37.

Evans, D. H., & Morgan, A. R. (1982) *J. Mol. Biol. 160*, 117–122.

Fasman, G. D. (1975) *Biochem., Mol. Biol., 3rd Ed. 1*, 589.

Fink, T. R., & Crothers, D. M. (1972) *J. Mol. Biol. 66*, 1–12.

Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 73*, 497–511.

Gray, D. M., Vaughan, M., Ratcliff, R. L., & Hayes, F. N. (1980) *Nucleic Acids Res. 8*, 3695–3707.

Hare, D. R., & Reid, B. R. (1982) *Biochemistry 21*, 1835–1842.

Helfgott, D. C., & Kallenbach, N. R. (1979) *Nucleic Acids Res. 7*, 1011–1017.

Hore, P. J., Zuiderweg, E. R., Nicolay, K., Dijkstra, K., & Kaptein, R. (1982a) *J. Am. Chem. Soc. 104*, 4286–4288.

Hore, P. J., Scheek, R. M., Volbeda, A., & Kaptein, R. (1982b) *J. Magn. Reson. 50*, 328–334.

Khorana, H. G. (1968) *Pure Appl. Chem. 17*, 349–381.

Lee, C.-H., & Tinoco, I., Jr. (1978) *Nature (London) 274*, 609–610.

Lomant, A. J., & Fresco, J. R. (1973) *J. Mol. Biol. 77*, 345–354.

Lomant, A. J., & Fresco, J. R. (1974) *Prog. Nucleic Acid Res. Mol. Biol. 14*, 185–218.

Matteucci, M. D., & Caruthers, M. H. (1981) *J. Am. Chem. Soc. 103*, 3185–3191.

Nelson, J. W., Martin, F. H., & Tinoco, I., Jr. (1981) *Biopolymers 20*, 2509–2531.

Noggle, J. H., & Shirmer, R. E. (1971) *The Nuclear Overhauser Effect: Chemical Applications*, Academic Press, New York.

Pardi, A., Martin, F. H., & Tinoco, I., Jr. (1981) *Biochemistry 20*, 3986–3996.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 445–451.
Redfield, A. G., Kunz, S. D., & Ralph, E. K. (1975) *J. Magn. Reson. 19*, 114–117.
Sanchez, V., Redfield, A. G., Johnston, P. D., & Tropp, J. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5659–5662.
Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77–84.
Uhlenbeck, O. C., Borer, P. N., Dengler, B., & Tinoco, I., Jr. (1973) *J. Mol. Biol. 73*, 497–511.
Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948–5949.

Internal Motions in Ribonucleic Acid Duplexes As Determined by Electron Spin Resonance with Site-Specifically Spin-Labeled Uridines†

Shih-Chung Kao, Carl F. Polnaszek, Charles R. Toppin, and Albert M. Bobst*

ABSTRACT: We have used site-specifically spin-labeled polyribonucleotides to extract motional information for single- and double-stranded polyribonucleotides. The spin-label is attached to either position 4 or 5 of an uridine analogue, and in the latter position the label is linked through various tethers to the base. This strategy together with a motional model has allowed us to separate the motion of the base from the motion of the spin-label in single and double strands. We conclude that the bases in an RNA duplex experience substantial motions with a correlation time in the range of 4 ns.

Internal motions of nucleic acid duplexes have been studied by a variety of techniques (Bolton & James, 1979; Feigon & Kearns, 1979; Klevan et al., 1979; Bobst, 1980; Hogan & Jardetzky, 1980; Robinson et al., 1980a,b; Shido & McGhee, 1980; Bobst et al., 1981a; Hogan et al., 1982; Allison et al., 1982; Hurley et al., 1982; McCain et al., 1982; Wang et al., 1982). Phosphate backbone motion of DNA in nucleosomes occurs with a time constant of 1–30 ns as determined by ^{31}P NMR (Klevan et al., 1979; Shindo & McGhee, 1980), but a rigid DNA in nucleosomes was derived on the basis of ^{1}H NMR line-width measurements (Feigon & Kearns, 1979). More recently substantial DNA motions with a time constant near 30 ns in nucleosomes and histone-free DNA were determined by triplet anisotropy decay of intercalating dyes (Hogan et al., 1982; Wang et al., 1982). Similar mobilities were derived from ESR measurements on spin labeled propidium intercalated in DNA which resulted in correlation times of the order of 30 ns (Robinson et al., 1980a,b; Hurley et al., 1982).

We present results about the motions of single-stranded and double-stranded polyribonucleotides based on site-specifically spin-labeled nucleic acid building blocks. The spin-label, consisting of a six-membered nitroxide radical with tethers of variable lengths, is attached in either position 4 or 5 of the pyrimidine base. In the latter position various tethers are used to attach the nitroxide radical to the base. With this strategy we were able to express the ESR line shapes in terms of geometrical (tilt angle) and two motional parameters ($\tau_{\parallel}$ and $\tau_{\perp}$),[1] thereby allowing us to separate the motion of the tether axis from motions due to torsion and tilting of the base pairs as well as twisting of the bases. We conclude that the motions of the bases in an RNA duplex occur in the range of 4 ns.

Materials and Methods

$(A)_n$ was bought from P-L Biochemicals (minimum s_{20}) and purified through Sephacryl S-200 prior to use. All other materials were commercial products of analytical reagent grade.

Synthesis of $(pps^5U)_2$. A sulfur substituent was introduced into position 5 of uridine 5′-diphosphate according to a published procedure (Ho et al., 1978). The 5-thiolated nucleotide disulfide was isolated as the ammonium salt in overall 20–30% yield and was characterized by its UV spectrum (λ_{max}^{pH7} = 273 nm) (Ho et al., 1978).

Preparation of Spin-Labeled Alkylating Agents. Epoxy-Tempo, the synthesis of which was reported earlier by others (Rauckman et al., 1976), was obtained in small quantities from Molecular Probes, Inc. (Junction City, OR), and in larger quantities from the Josef Stefan Institut in Yugoslavia. α-Iodoacetamido-Tempo was isolated after activation of the

† From the Department of Chemistry, University of Cincinnati, Cincinnati, Ohio 45221 (S.-C.K., C.R.T., and A.M.B.), and the Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455 (C.F.P.). *Received May 9, 1983.* This investigation was supported in part by grants from the U.S. National Science Foundation (No. PCM 7801979) and U.S. Public Health Service (No. GM 27002). Purchase of the Nicolet NTC 300 FT instrument was supported in part by National Science Foundation Grant CHE-81-02974.

[1] Abbreviations: $(A)_n$, poly(adenylic acid); $(U)_n$, poly(uridylic acid); $(pps^5U)_2$, bis[1-(5′-*O*-diphosphono-β-D-ribofuranosyl)uracil-5-yl] disulfide; epoxy-Tempo, 5,5,7,7-tetramethyl-1-oxa-6-azaspiro[2.5]oct-6-yloxy; α-iodoacetamido-Tempo, 4-(α-iodoacetamido)-2,2,6,6-tetramethylpiperidinyl-1-oxy; β-iodopropanamido-Tempo, 4-(β-iodopropanamido)-2,2,6,6-tetramethylpiperidinyl-1-oxy; DTT, dithiothreitol; $(RUMMT,U)_n$, copolymer of RUMMT and uridine; ppRUMMT, (1-oxy-2,2,6,6-tetramethyl-4-hydroxy-4-piperidinyl)methyl 1-(5′-*O*-diphosphono-β-D-ribofuranosyl)uracil-5-yl sulfide; $(RUTT,U)_n$, copolymer of RUTT and uridine; ppRUTT, *N*-[(1-oxy-2,2,6,6-tetramethyl-4-piperidinyl)methyl]carbamoyl 1-(5′-*O*-diphosphono-β-D-ribofuranosyl)-uracil-5-yl sulfide; $(RUMPT,U)_n$, copolymer of RUMPT and uridine; ppRUMPT, *N*-[(1-oxy-2,2,6,6-tetramethyl-4-piperidinyl)ethyl]carbamoyl 1-(5′-*O*-diphosphono-β-D-ribofuranosyl)uracil-5-yl sulfide; $(1s^4U,U)_n$, copolymer of $1s^4U$ and uridine; pp$1s^4U$, *N*-[(1-oxy-2,2,6,6-tetramethyl-4-piperidinyl)methyl]carbamoyl 1-(5′-*O*-diphosphono-β-D-ribofuranosyl)uracil-4-yl sulfide; RNA, ribonucleic acid; T_m^{OD}, melting temperature determined by UV spectroscopy; $\tau_{\parallel}$ and $\tau_{\perp}$, correlation times for rotations about and perpendicular to the principal axis of diffusion, respectively.

0006-2960/83/0422-5563$01.50/0