NMR Study of a Synthetic DNA Hairpin†

Satoshi Ikuta,‡ Rajagopal Chattopadhyaya,§ Hirataka Ito,§,∥ Richard E. Dickerson,§ and David R. Kearns*,‡

Department of Chemistry, University of California, San Diego, La Jolla, California 92093, and Department of Chemistry, University of California, Los Angeles, Los Angeles, California 90024

Received October 28, 1985; Revised Manuscript Received February 28, 1986

ABSTRACT: The secondary structure of the synthetic oligodeoxyribonucleotide d(CGCGCGTTTTCGCGCG) (I) has been demonstrated to be a unimolecular hairpin structure (hairpin I) over a wide range of oligonucleotide concentrations (2×10^{-5} to 1.6×10^{-3} M) and temperature (0–87 °C). The assignments of the resonances to specific protons were carried out by use of two-dimensional nuclear Overhauser effect and COSY spectra and by comparison with the spectra of the duplex formed by $d(CG)_3$. Comparison of hairpin I and the hairpin of d(ATCCTATTTTTAGGAT) (II) reveals that the exchange of imino protons in stem base pairs with solvent is much slower in I than in II. However, the exchange of thymine imino protons in the loop region is much faster in I than in II even though both hairpins contain four unpaired thymine residues. The secondary structure of hairpin I contains only six G·C base pairs, yet it is more stable than the $d(CG)_8$ duplex containing 16 G·C base pairs at all concentrations of duplex lower than 10^{-3} M. These observations suggest that intramolecular hairpin formation may effectively compete with bimolecular duplex formations when the appropriate intramolecular base pairs can form.

Hairpin structures are present in most naturally occurring RNA molecules [e.g., tRNA (Barrell & Clark, 1974), 5S RNA (Erdmann, 1980), mRNA (Proudfoot & Brownlee, 1974), viroids (Steger et al., 1984), and ribosomal RNA (Glotz & Brimacombe, 1980)], but they are much less frequent in DNA because of competition with formation of completely complementary Watson–Crick duplexes. Nevertheless, under appropriate conditions the formation of looped out regions and hairpin structures in DNA may play an important biochemical role in vivo and in vitro. This is suggested by the fact that palindromic sequences frequently occur in biological control regions (Rosenberg & Court, 1979; Wells et al., 1980; Muller & Fitch, 1982) and the experimental evidence that looped out cruciform structures can be generated by supercoiling of DNA (Lilley, 1980, 1981; Panayotatos & Wells, 1981b). Studies involving the hybridization of short DNA oligonucleotides to the complementary DNA in molecular cloning may also be affected by the tendency of the single-stranded DNA to "hairpin" when the appropriate complementary sequences exist within the same strand. In this regard, we note that even self-complementary DNA oligonucleotides [e.g., d(CGCGTATACGCG) and d(CGCGAATTCGCG)] can form hairpin structures in solution (Marky et al., 1983; Patel et al., 1983; Wemmer et al., 1985) and that the conditions used in hybridization experiments (low concentration) tend to favor hairpin formation. For these and other reasons, there have been many studies of the formation and physical properties of hairpin structures in synthetic DNA fragments with particular emphasis on establishing the base pairing structure, the thermodynamics of hairpin formation, the structure of the hairpin loop, and the effect of loop size on the stability of the hairpin (Lilley, 1980; Panayotatos & Wells, 1981; Mizuuchi et al., 1982; Courey & Wang, 1983; Haasnoot et al., 1983; Germann & Van de Sande, 1985).

In the work presented here, we have used ^1H NMR and other techniques to examine the behavior of the synthetic DNA oligonucleotide d(CGCGCGTTTTCGCGCG) (I). As we shall demonstrate, this molecule only forms a unimolecular hairpin structure (see Figure 1) under a wide variety of conditions. Other physical properties such as thermal stability, imino proton exchange behavior, and loop structure are examined and compared with the behavior previously observed (Haasnoot et al., 1983) for d(ATCCTATTTTATCC), a molecule that also forms a hairpin structure with four thymine residues in the loop, and for the hexamer that comprises the stem of the hairpin [d(CGCGCG)·d(CGCGCG)].

Materials and Methods

Materials. The four DNA molecules d(CGCGCGTTTTCGCGCG) (hairpin I), d(CGCGCGTATACGCGCG), d(CGCGTACGCG), and d(CGCGCG) were synthesized by the modified phosphotriester method (Tan et al., 1983). The positive counterions of the DNA were changed to sodium by applying the DNA onto a DEAE-cellulose column (1 × 3 cm, DE-52, Whatman) followed by elution with 1 M NaCl solution. The sodium form of the DNA was dialyzed against 1 mM NaCl and 1 mM ethylenediaminetetraacetic acid (EDTA). The concentration of the DNA, salt, buffer, and pH are described in figure captions. All NMR measurements were performed with samples in 5-mm NMR tubes containing 0.3 mL of solution. For spectra in D_2O, the samples were dried in the NMR tubes under a stream of N_2 and redissolved with 99.996% D_2O (Stohler Isotopes) 3 times. The pH of each sample was measured by putting the pH meter (Beckman Model 3500) microprobe into the NMR tube.

NMR Measurements. The NMR experiments were performed on a 360-MHz Fourier-transform spectrometer equipped with an Oxford Instruments magnet and Nicolet 1180 computer. The low-field spectra were obtained in H_2O at 360 MHz with a 1331 pulse sequence (Hore, 1983). COSY and NOESY spectra were obtained with 2048 points in t_2 and 128 points in t_1, and 64 scans were accumulated with 4-s delays between acquisitions. Chemical shifts (ppm) were determined

† This work was supported by the National Science Foundation (Grant DMB 83-03374 to D.R.K. and PCM 82-02775 to R.E.D.).

‡ University of California, San Diego.

§ University of California, Los Angeles.

∥ Present address: Asahi Chemical Industry Co., Ltd., Shizuoka, Japan.

0006-2960/86/0425-4840$01.50/0 © 1986 American Chemical Society

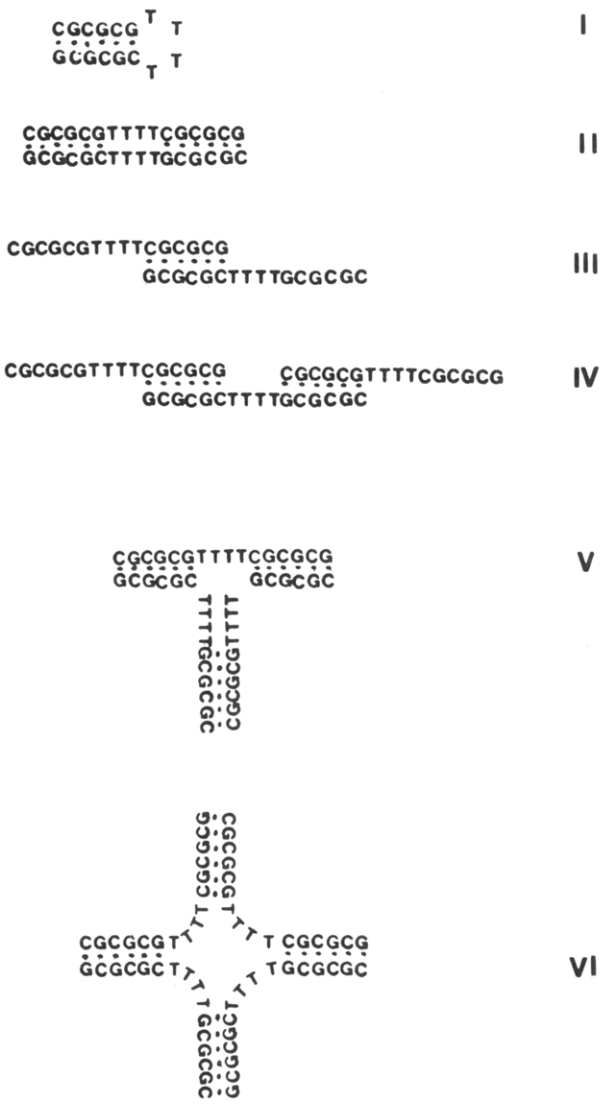

FIGURE 1: Summary of some possible secondary structures for d-(CGCGCGTTTTCGCGCG).

relative to the chemical shift of water (or the residual HOD peak).

Fluorescence Polarization Anisotropy (*FPA*). The anisotropy of the fluorescence polarization of DNA complexes of ethidium was measured with the methods and apparatus described elsewhere (Genest et al., 1985). The excitation was typically at 520 nm, and emission was monitored at 620 nm. The d(CGCGCGTTTTCGCGCG), d(CGCGTACGCG), and $d(CG)_3$ samples were examined in the fluorescence experiments.

Gel Electrophoresis. Prior to their application to 20% polyacrylamide gel, all DNA samples and the gel were equilibrated at 4 °C for at least 24 h. The concentrations of the DNA samples were 10 μg in 20 μL for hairpin I, and others were around 20–25 μg/20 μL. The gels were run at 4 °C in a buffer solution containing 0.1 M NaCl and 10 mM sodium phosphate, pH 7.0, followed by ethidium bromide staining. The concentration of the hairpin I *in the gel* was estimated to be 0.16 mM strand by measuring the volume of the band (0.1 × 0.2 × 0.8 cm) after the gel was run.

Thermal Denaturation. The absorbance vs. temperature profiles of $d(CG)_3$ and d(CGCGCGTTTTCGCGCG) at low concentration (2×10^{-5} M) were measured on a Beckman ACTA CIII spectrophotometer. The temperature was increased at a rate of 1 °C min^{-1} by using a circulating bath

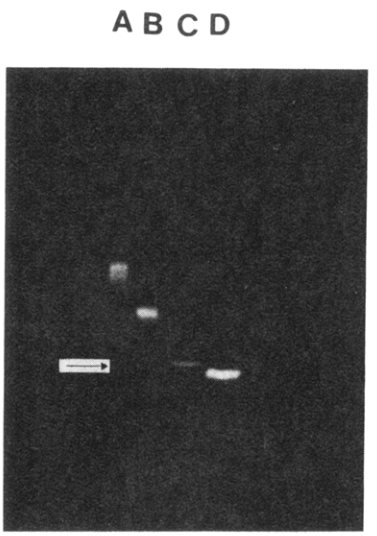

FIGURE 2: The 20% polyacrylamide gel electrophoresis profile of four DNA oligonucleotides 4 °C: (lane A) d(CGCGCGTATACGCGCG), 16mer duplex; (lane B) d(CGCGTACGCG), 10mer duplex; (lane C) d(CGCGCGTTTTCGCGCG), hairpin I; (lane D) $d(CG)_3$, 6mer duplex. A faint band in lane A (indicated by arrow) is probably due to hairpins.

(HAAKE) connected to the spectrophotometer. The temperature in the cuvette was directly monitored by a thermocouple (Bailey Instruments).

RESULTS

Molecular Size of d(CGCGCGTTTTCGCGCG)

We have used three different methods to monitor the size of hairpin I: (a) polyacrylamide gel electrophoresis, (b) FPA measurement, and (c) NMR relaxation methods.

Polyacrylamide Gel Electrophoresis. Figure 2 shows the gel electrophoresis profile of $d(CG)_3$ 6mer, d-(CGCGTACGCG) 10mer, and d(CGCGCGTA-TACGCGCG) 16mer duplexes and hairpin I at 4 °C. Hairpin I shows only one major band with no trace corresponding to a larger structure, such as the bimolecular duplex. Hairpin I migrates as a single band in the gel faster than the 10mer duplex but slower than the 6mer duplex. The apparent molecular size of hairpin I is therefore around that of an eight base pair duplex.

We note that in the gel electrophoresis experiment shown in Figure 2 the d(CGCGCGTATACGCGCG) 16mer exhibits a weak band that migrated faster than the main band corresponding to the 16 base paired duplex. This minor component, which comigrates with hairpin I, is very likely the hairpin form, but additional studies will be required to confirm this.

Fluorescence Polarization Anisotropy (*FPA*). FPA has been used to monitor the size of the DNA and to measure correlation times for molecular rotation about the short and long axes of DNA fragments (Genest et al., 1985). The FPA of ethidium complexes of the 6mer and 10mer duplexes and hairpin I has been measured with the following results: 0.022 for 6mer duplex, 0.0403 for 10mer duplex, and 0.028 for hairpin I. The corresponding rotational correlation times deduced from these experiments are 4.5 ± 0.2 ns for hairpin I, 3.4 ± 0.2 ns for the 6mer duplex, and 6.2 ± 0.2 ns for the 10mer duplex. The rotational correlation time for hairpin I is therefore intermediate between the values obtained for the 6 and 10 base pair duplexes.

NMR Relaxation. The semiselective spin–lattice relaxation rate R_1^{ss} of aromatic H6 protons in cytosine (CH6) was measured for each sample, $d(CG)_3$ 6mer, d(CGCGCGTA-

Table I: Correlation Times Measured by NMR Relaxation Techniques and FPA at 20 °C[a]

	R_1^{ss} (s^{-1})	τ_c (ns)	$\tau_\perp$ (ns)
$d[(CG)_3(TA)_2(CG)_3]$	5.5 ± 0.3	4.2 ± 0.4	
$d[(CG)_2TA(CG)_2]$			6.2
$d[(CG)_3T_4(CG)_3]$	3.5 ± 0.3	2.5 ± 0.4	4.5
$d(CG)_3$	2.7 ± 0.3	1.9 ± 0.4	3.4

[a] R_1^{ss}, semiselective spin–lattice relaxation of H6 protons in cytosine residue; τ_c, effective rotational correlation time deduced from R_1^{ss}; $\tau_\perp$, rotational correlation time obtained by FPA. The following interactions were used with eq 1 to calculate τ_c: CH6–CH5 = 2.40 Å, CH6–H2′ = 2.20 Å, and CH6–H2″ = 2.20 Å.

TACGCGCG) 16mer, and (CGCGCGTTTTCGCGCG) hairpin I, by the inversion-recovery technique (Morris & Freeman, 1978) with a DANTE 180° selective pulse. An example of an R_1^{ss} measurement on hairpin I (1.6 mM strand) at 20 °C is shown in Figure 3A, and a semilogarithmic plot of the recovery of the magnetization is shown in Figure 3B. The values of R_1^{ss} obtained for CH_6 protons were 2.7 ± 0.3 s^{-1} for the 6mer, 3.5 ± 0.3 s^{-1}, for hairpin I, and 5.5 ± 0.3 s^{-1} for the 16mer, respectively. R_1^{ss} of hairpin I measured in an ~10-fold more dilute solution (0.16 mM strand) was 3.5 ± 0.4 s^{-1}. Since these DNAs all have a common $(CG)_3$ stretch in their sequence, we assume the differences in relaxation rates reflect differences in rotational correlation time.

The relationship between R_1^{ss} and τ_c for an isotropic rotor, in the slow motion limit, is (Feigon et al., 1983)

$$R_1^{ss} = 0.1\gamma_H^4\hbar^2\tau_c\sum_j \frac{1}{r_{ij}^6} \quad (1)$$

γ_H = magnetogyric ratio of proton, $\hbar$ = Plank's constant divided by 2π, r_{ij} = internuclear distance between spins i and j, and τ_c = effective molecular correlation time. The measured R_1^{ss} and calculated τ_c for each DNA are listed in Table I. The τ_c obtained by NMR relaxation techniques on hairpin I is between the τ_c for the 6mer and the τ_c for the 16mer, suggesting that hairpin I has a unimolecular hairpin structure rather than a bimolecular or polymeric secondary structure.

NMR Properties

Assignments of Nonexchangeable Protons. NMR spectra of the nonexchangeable protons in $d(CG)_3$ and hairpin I at 20 and 40 °C are presented in Figure 4, along with assignment to proton type. Both hairpin I and $d(CG)_3$ exhibit similar spectra at 20 °C except that the line widths in the hairpin I spectrum are larger and there are four additional well-resolved peaks in the upfield region corresponding to the four thymine methyl protons (T-CH3). On a rise in the temperature to 40 °C, the hairpin I spectrum is much better resolved. Since the proton spectrum of $d(CG)_3$ has been previously assigned (Cheng et al., 1984), most resonances in hairpin I can readily be assigned to proton type just by comparing the spectra shown in Figure 4.

The GH8 protons can readily be distinguished from other aromatic protons since they exchange with D_2O on heating (Schweizer et al., 1964). Comparison of the spectra of the partially exchanged and of the fully protonated molecules proves the GH8 resonances are located only in the cluster around 8.0 ppm. The CH5 and CH6 resonances from six cytosine residues can be unambiguously identified by the COSY (Aue et al., 1976) spectrum, which shows the CH5–CH6 coupling (Figure 5A). The thymine H6 (TH6) resonances were assigned from the NOESY spectra (Jeener et al., 1979; Kumar et al., 1980) shown in Figure 5. Each TH6 has strong cross peaks with its own methyl group (T-CH3) resonating in the region between 1.6 and 1.9 ppm. Significantly,

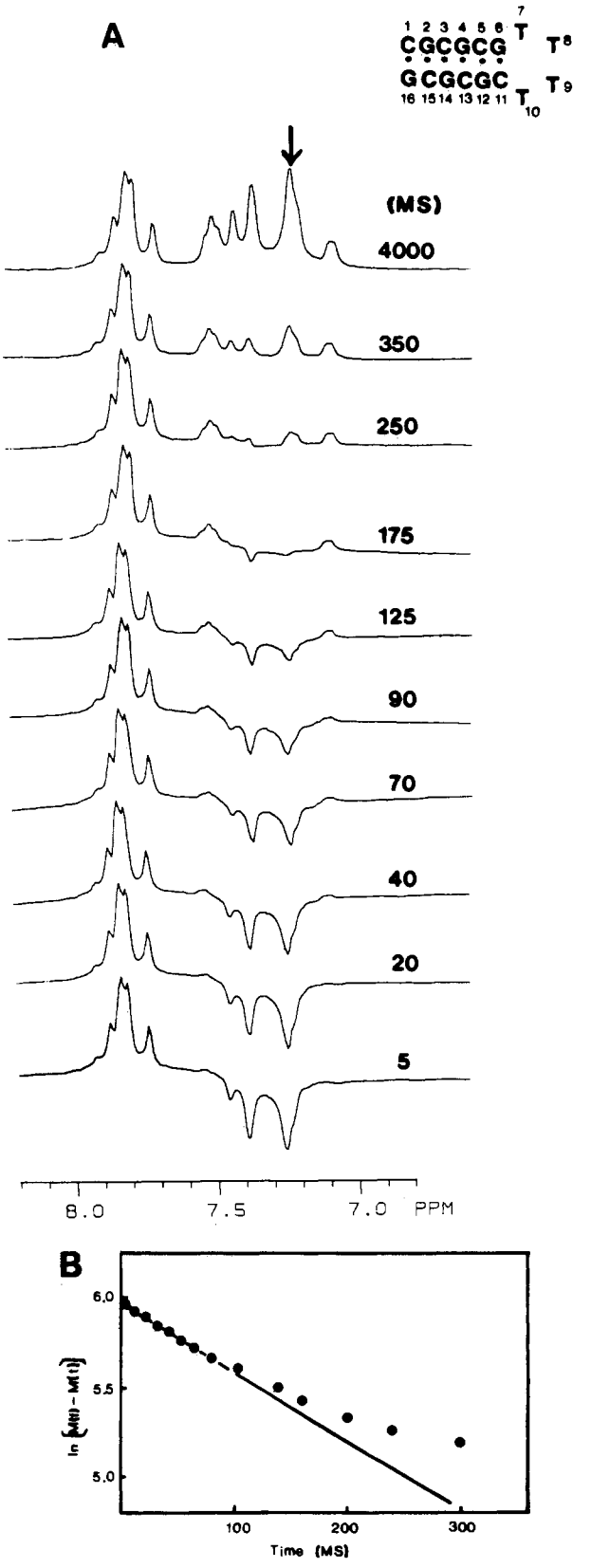

FIGURE 3: (A) Semiselective spin–lattice relaxation of the aromatic protons of hairpin I (1.6 mM strand concentration) in 10 mM sodium phosphate and 0.1 M NaCl, pH 6.8, at 20 °C. Resonances in the aromatic region were selectively inverted by a standard DANTE pulse sequence. Twenty-six hard pulses (2 μs) centered at 7.4 ppm were applied spaced 192 μs apart with a relaxation delay of 8 s and a 90° pulse length of 26 μs. The arrow indicates the H6 proton of cytosine residues (CH6). (B) Semilogarithmic plots of the recovery of the semiselective spin–lattice relaxation in CH6 proton resonance of hairpin I.

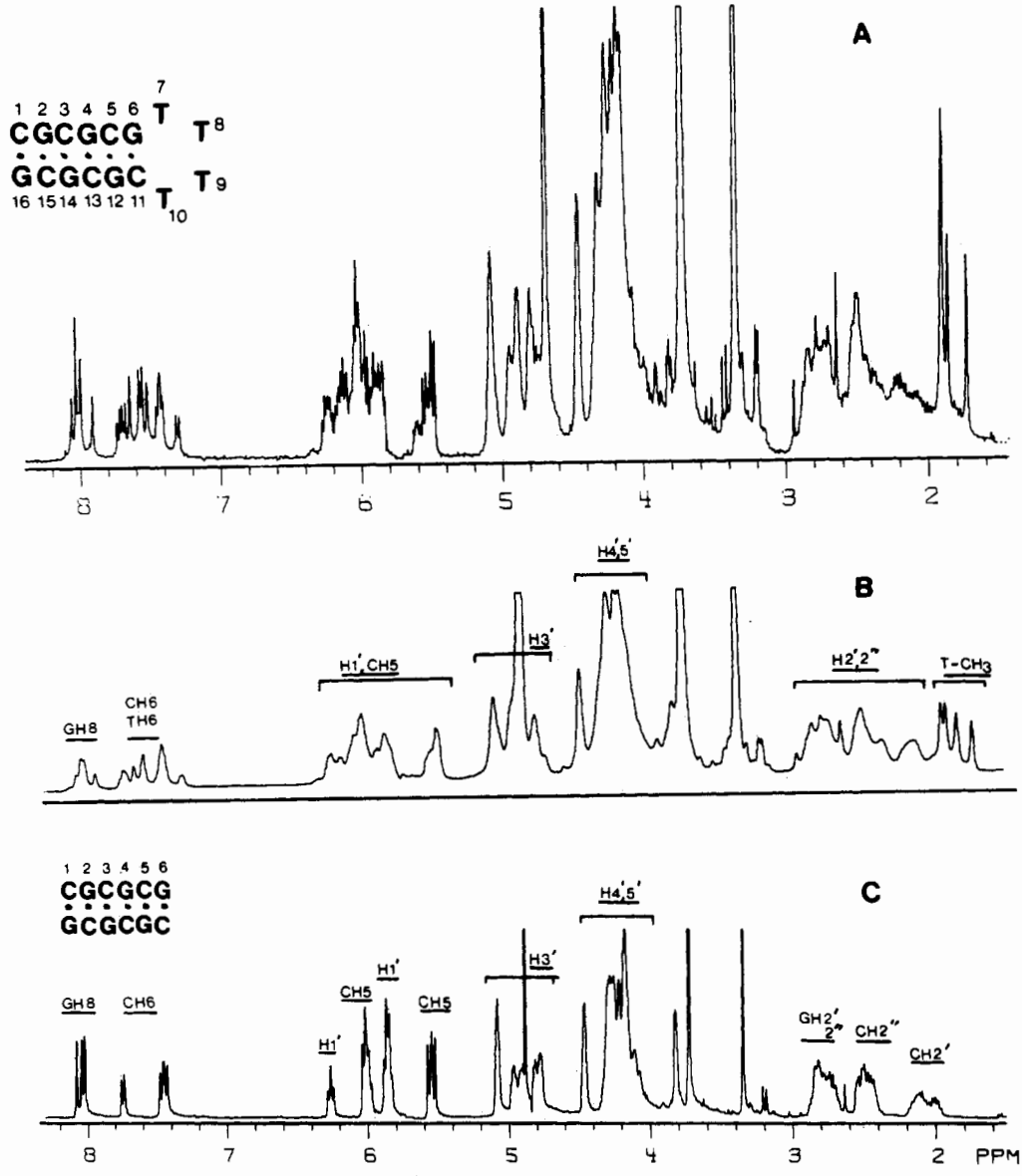

FIGURE 4: The 360-MHz proton NMR spectra of the nonexchangeable protons of two synthetic DNAs in D_2O. Both hairpin I and $d(CG)_3$ have 1.6 mM strand concentration in 10 mM sodium phosphate and 0.1 M NaCl, at pH 6.8. The spectra for hairpin I are obtained at 40 (A) and 20 °C (B) and the spectra for $d(CG)_3$ at 20 °C (C). Assignments to general proton type are given above the spectra.

one T-CH3 resonance (denoted peak b) in the 4 °C spectrum also has a cross peak with one of the GH8 resonances. This permits us to assign these two resonances to T_7 and G_6, respectively. At 22 °C, this cross peak is absent.

Temperature Dependence of Chemical Shifts. Plots of the chemical shifts of the aromatic, CH5, and T-CH3 resonances vs. temperature for the $d(CG)_3$ and hairpin I are shown in Figure 6. The transition midpoint (T_m) of the helix-to-coil transition of DNA can be estimated from chemical shift vs. temperature curves. For $d(CG)_3$ at a strand concentration of 3.2 mM, we obtained $T_m \sim 67$ °C, whereas the T_m of hairpin I was around 87 °C, independent of strand concentration. A plot of $1/T_m$ vs. logarithm DNA concentration for $d(CG)_3$ and for hairpin I is also shown in Figure 6. Over the accessible range of concentrations, the hairpin I is more stable than $d(CG)_3$.

Relaxation Measurements of Internal Motion. We have measured the effective correlation times τ_c for thymine and cytosine residues in the loop and stem region of hairpin I, respectively, by the pre-steady-state NOE method (Clore & Gronenborn, 1984). The cross relaxation rate σ_{ij} between protons i and j can be determined from a measurement of the initial slope of NOE enhancement vs. irradiation time (Wagner & Wuthrich, 1979; Dobson et al., 1982) and is related to τ_c as follows:

$$\sigma_{ij} = \frac{\gamma^4\hbar^2}{10r_{ij}^6}\left(\tau_c - \frac{6\tau_c}{1 + 4\omega^2\tau_c^2}\right)$$

where ω is the spectrometer frequency. Therefore, if the interproton distance r_{ij} is known, the rotational correlation time is readily extracted. We calibrated this method by using the 10 base pair duplex $d(ATATCGATAT)_2$, which contains both

Table II: Cross Relaxation Rates (σ) and Effective Rotational Correlation Time (τ_c) Obtained by Pre-Steady-State NOE Measurement at 18 °C[a]

	$\sigma_{CH6\rightarrow CH5}$ (s^{-1})	τ_c (ns)	$\sigma_{TH6\rightarrow T\text{-}CH_3}$ (s^{-1})	τ_c (ns)
$d(ATATCGATAT)_2$	0.8 ± 0.1	3.2 ± 0.4	0.5 ± 0.1	3.4 ± 0.6
hairpin I	0.5 ± 0.05	2.0 ± 0.2	0.16 ± 0.05	1.3 ± 0.3
$d(ATAT)_2$			~0	~0.5

[a] The following interproton distances (Å) are used: 2.46 for CH6–CH5 and 2.70 for TH6–T-CH3.

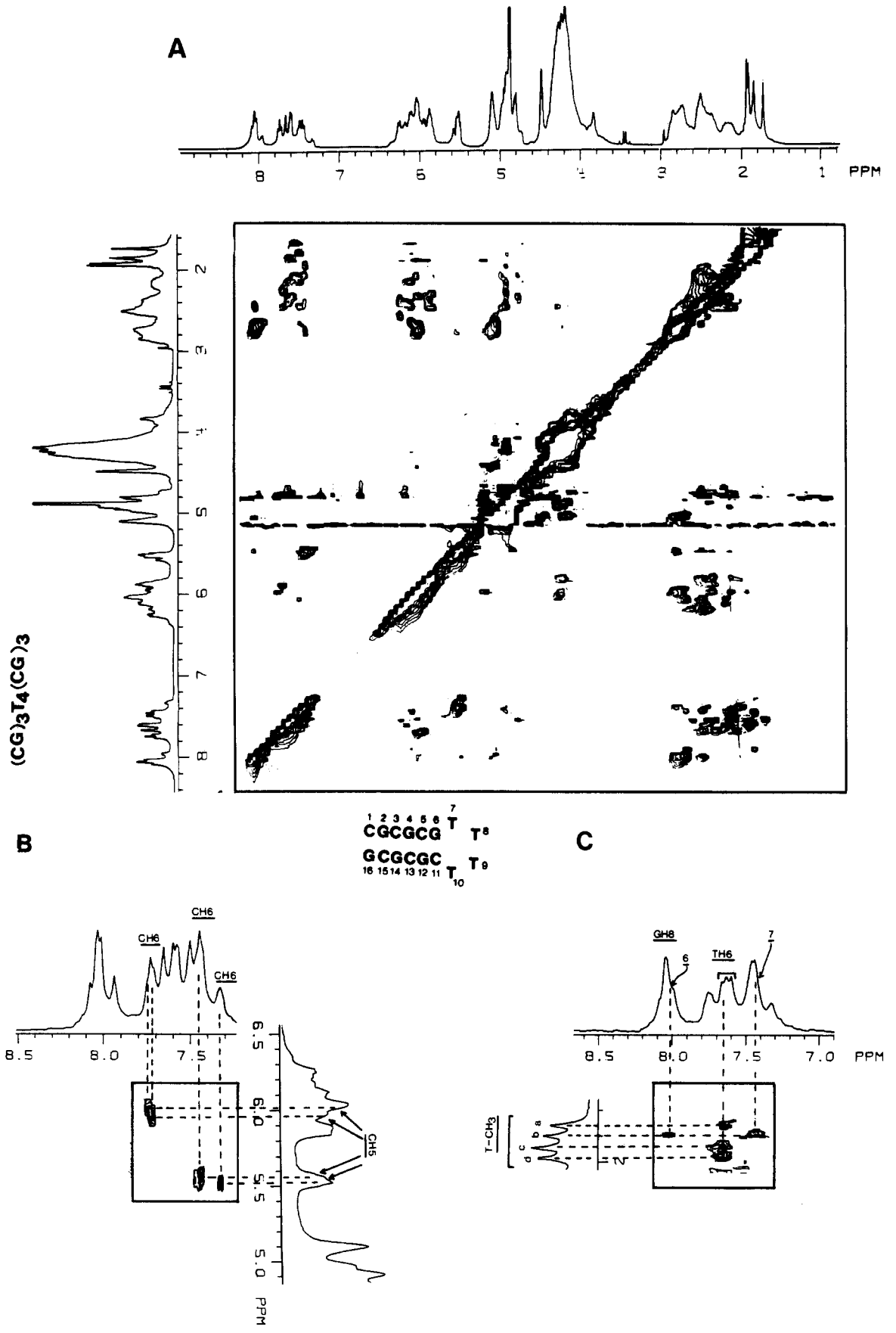

FIGURE 5: (A) NOESY spectrum of hairpin I at 22 °C obtained using a mixing time of 400 ms. (B) Contour plot of the COSY spectrum of hairpin I at 30 °C in D_2O showing the connection between some of the CH6 and CH5 resonances. The DNA concentration and salt for 2D NMR are identical with that in Figure 4. (C) A portion of the NOESY spectrum of hairpin I at 4 °C. The mixing time was 400 ms. The dotted lines connect the cross relaxation between methyl groups of thymine (T-CH3) and TH6. One T-CH3 denoted (b) has cross peaks with TH6 and GH8.

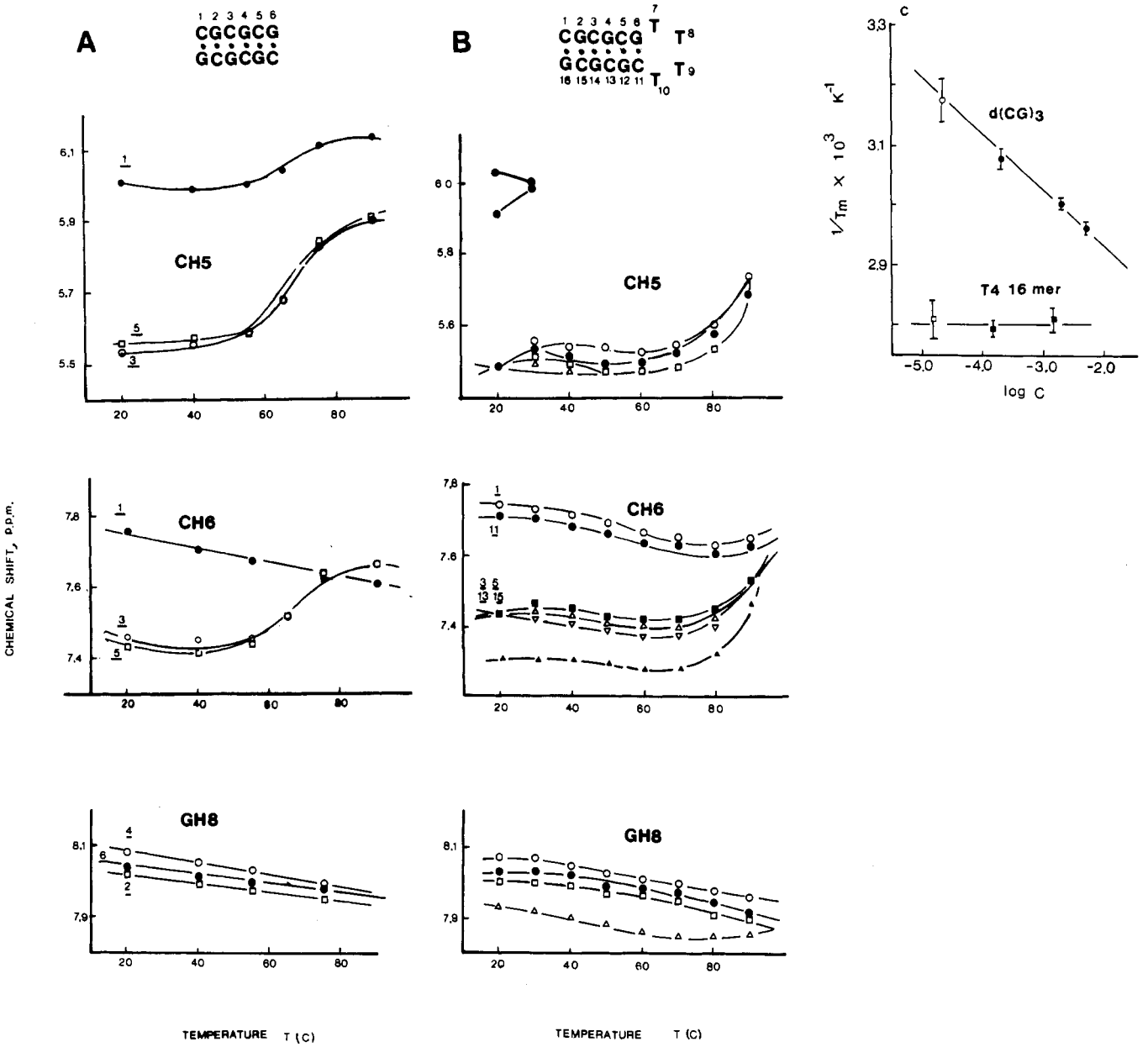

FIGURE 6: (A) Temperature vs. chemical shift of the aromatic (CH6, CH8) resonances of $d(CG)_3$ and hairpin I (B) in D_2O containing 10 mM sodium phosphate and 0.1 M NaCl, pH 6.8. The concentration of DNA is 1.6 and 3.2 mM for hairpin I and $d(CG)_3$, respectively. (C) Plots of $1/T_m$ vs. log C for hairpin I and $d(CG)_3$ in H_2O containing 10 mM sodium phosphate and 0.1 M NaCl, pH 6.8. The closed circle (●) and the square (■) are obtained from the chemical shift vs. temperature curve in NMR, while the open circle (○) and square (□) are from the UV absorption at 260 nm vs. temperature curve.

A·T and G·C base pairs, and these results are summarized in Table II along with data for hairpin I. The cross relaxation rate of hairpin I was 0.5 s^{-1} for CH6 → CH5 interaction and 0.16 s^{-1} for TH6 → T-CH3 interaction. The calculated correlation times for bases in the stem and the loop region were ~2.0 and 1.3 ns, respectively. Therefore, the mobility of the loop region is greatly restricted by the stem and is not independent of the stem.

Imino Proton Resonance. Figure 7A shows the low-field imino proton spectra of hairpin I at various temperatures ranging from 0 to 65 °C at pH 6.8. The collection of the resonances at ~13.2 ppm can readily be assigned to the G-imino protons of a G·C base pair, while the resonances at ~11.0 ppm are due to imino protons of non-base-paired thymine in single-stranded regions (Haasnoot et al., 1983). The low-field spectra obtained at 0.16 mM strand concentration (data not shown) were identical with those obtained at 1.6 mM strand concentration and temperature ranging from 0 to 20 °C (see Figure 7A) under identical conditions (salt, buffer, pH). To test the effect of NaCl on the conformation of hairpin I, the low-field spectra (data not shown) were obtained at temperatures ranging from 0 to 20 °C under high salt (0.5 M NaCl) and very low salt (2.56 mM NaCl, 0.16 mM strand, pH 7.0) conditions. The latter NaCl concentration is just sufficient to neutralize the 16– charged phosphodiester group of hairpin I. The low-field spectra (data not shown) obtained were identical with those in Figure 7A. It should be noted that this concentration (0.16 mM strand) is the same as that in the 20% polyacrylamide gel electrophoresis shown in Figure 2.

Effect of pH and Temperature on Imino Protons. Figure 8 shows the temperature-dependence imino proton spectra of stem and loop bases at pH 7.1 and 5.3. At pH 7.1 (Figure 8A), thymine imino protons as well as guanine imino protons are present at 0 °C, but on a rise in temperature to 10 °C the thymine imino proton resonances become broader and by 20

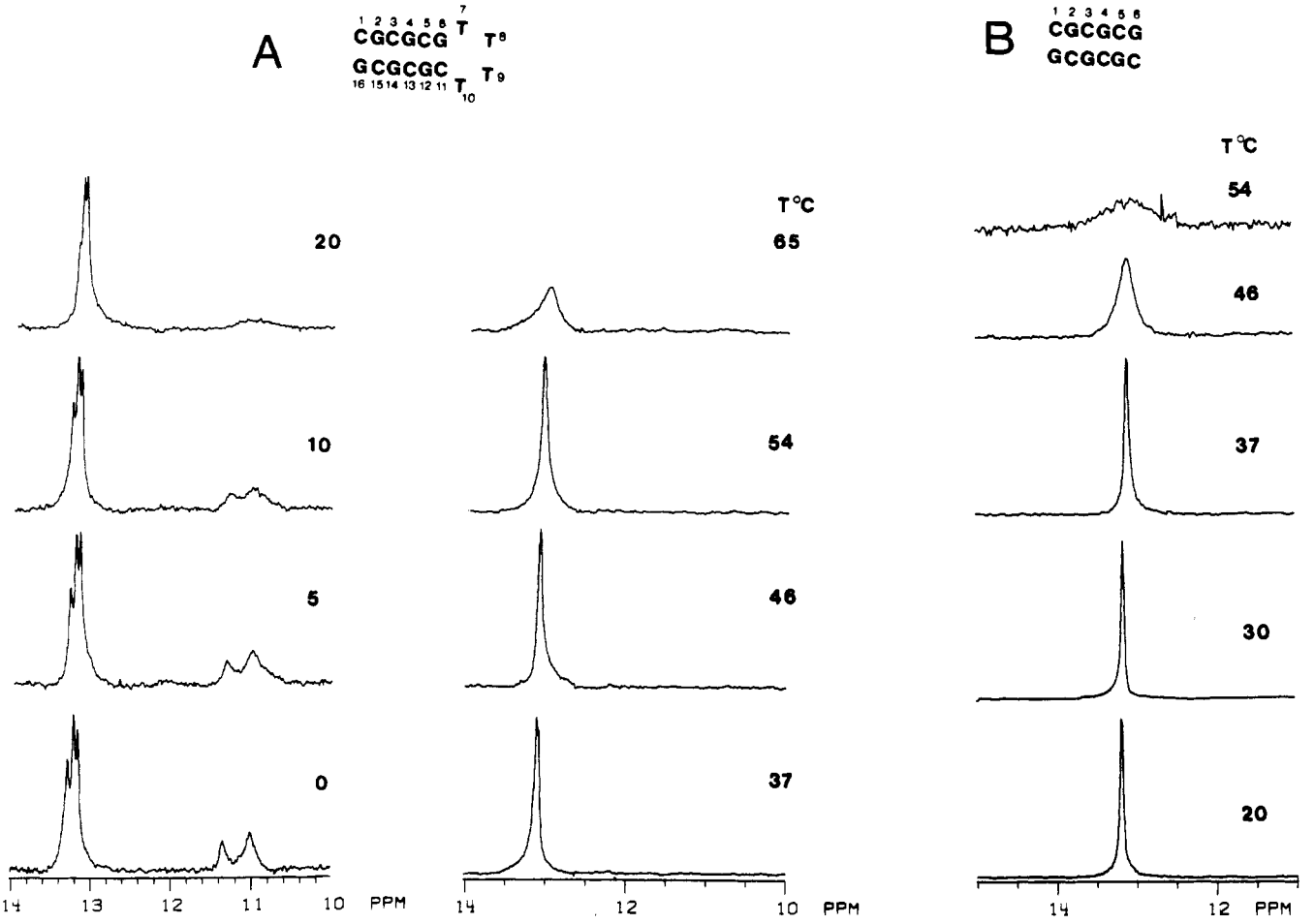

FIGURE 7: Temperature dependence of the 360-MHz low-field spectra (imino resonances) of (A) hairpin I (1.6 mM strand concentration) and (B) $d(CG)_3$ (3.2 mM strand concentration) in H_2O containing 10 mM sodium phosphate and 0.1 M NaCl, pH 6.8.

°C have almost disappeared. The guanine imino resonances, on the other hand, can still be detected at 65 °C, although they too are broadened at this temperature. In contrast, Figure 8B shows that at pH 5.3 thymine imino protons can be observed up to 40 °C even though the line width of the imino protons has increased. These data demonstrate that thymine imino protons in the loop region are sensitive to pH and temperature effects and less sensitive to salt concentration on exchange with solvent protons.

The spectra of the imino proton in the G·C base pairs of hairpin I and $d(CG)_3$ under the same buffer conditions at different temperatures are compared in Figure 7. The hydrogen-bonded imino proton resonances in $d(CG)_3$ are quite broad by 54 °C (T_m = 67 °C), but the imino protons of G·C base pairs in hairpin I are still observable even at 65 °C at pH 6.8 (T_m = 87 °C). The exchange rate of the G imino protons in hairpin I and $d(CG)_3$ deduced from line-width measurements at 54 °C are ~100 and ~800 s^{-1}, respectively. Both hairpin I and $d(CG)_3$ 6mer have a common $(CG)_3$ duplex region, but the imino protons of hairpin I are much more resistant to exchange with the solvent than those in $d(CG)_3$.

DISCUSSION

d(CGCGCGTTTTCGCGCG) Forms a Hairpin Structure in Solution. Figure 1 shows the variety of base pairing structures that might be adopted by this molecule. In most cases, the low-field spectrum is expected to exhibit six resonances from the common $(CG)_3$ stems and four resonances from the protected thymine residues in single-stranded regions. Since the ratio of low-field to protected thymine resonances could be comparable for many structures, the low-field NMR spectra cannot be used to differentiate between the various possibilities. The following experimental observations, however, demonstrate that hairpin I is the major form present in solution over a wide range of experimental conditions.

Polyacrylamide gel electrophoresis demonstrates that hairpin I migrates slightly slower than a 6mer duplex $d(CG)_3$ but considerably faster than a 16mer duplex (Figure 2). This demonstrates that under gel electrophoresis conditions (0.1 NaCl, 0.16 mM strand) the hairpin I migrates like an ~8mer.

Fluorescence polarization anisotropy measurements using ethidium staining show that the rotational correlation time for the hairpin I is intermediate between those obtained for a 10mer and a 6mer (see Table I). The semiselective spin–lattice relaxation rate for the CH6 protons in the hairpin I are intermediate between those observed for the $d(CG)_3$ 6mer and the d(CGCGCGTATACGCGCG) 16mer.

The T_m of hairpin I, estimated from measurements of the temperature dependence of the chemical shifts, is concentration-independent, whereas the T_m of $d(CG)_3$ varies strongly with concentration (see Figure 6). This eliminates all structures for the d(CGCGCGTTTTCGCGCG) except the hairpin, since they would all exhibit melting temperatures that depend on concentration.

Taken together, the above experiments prove that the hairpin structure I in Figure 1 is the correct base pairing structure for hairpin I over the range of concentrations studied here (2×10^{-5} to 1.6×10^{-3} M).

Nature of Loop Structure and Stability. Exchange behavior of the imino protons in the loop and stem regions is expected

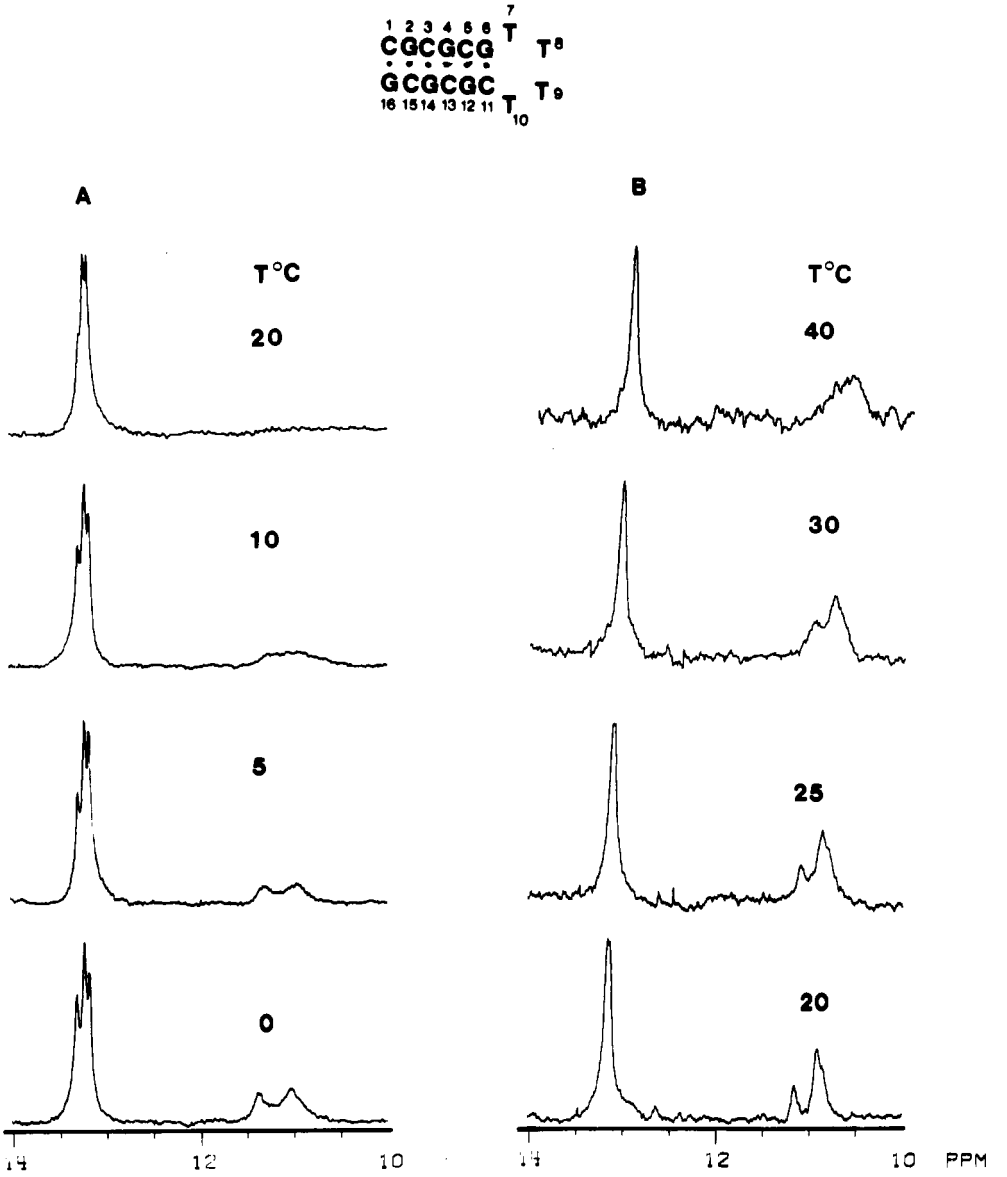

FIGURE 8: Temperature dependence of the 360-MHz low-field spectra (imino resonances) of hairpin I (1.6 mM strand concentration) at (A) pH 7.08 and (B) pH 5.30 in H_2O containing 10 mM sodium phosphate and 0.1 M NaCl.

to be quite different since nonpaired imino protons are more exposed to the solvent. As a consequence, the exchange rate of the nonpaired imino protons is expected to be sensitive to the pH, ionic strength, temperature, and DNA concentration. In particular, comparison of the T-imino proton spectra obtained under different experimental conditions revealed that exchange rates of the nonpaired imino protons are much enhanced on increasing pH from 5.3 to 7.1 but little affected by changes in the concentration of sodium chloride (from 2.6 mM to 0.5 M). These data suggest that while the exchange of the imino protons in the loop region is sensitive to pH and temperature, it is insensitive to the concentration of the salt. It is interesting to compare the hexadecamers d(CGCGCGTTTTCGCGCG), hairpin I, and d(ATCCTATTTTTAGGAT) (hairpin II). Both molecules form hairpin structures with the four T's in the loop, but there appear to be major differences in the resistance of the thymine imino protons to exchange with solvent. Under both comparable and identical experimental conditions of salt and pH, the imino protons from loop thymines in hairpin II are much more stable than in our molecule (hairpin I). Whereas imino resonances are observed up to 40 °C in hairpin II, at which point both stem and loop resonances of imino protons are rapidly exchanging with solvent (hairpin II), in hairpin I they have broadened due to solvent exchange somewhere between 10 and 20 °C. Conversely, the guanine imino protons in hairpin I resist exchange broadening up to 65 °C. The exchange rates of the thymine imino protons inferred from the observed line width at 20 °C are $\sim 600\ s^{-1}$ for hairpin I and $\sim 180\ s^{-1}$ for hairpin II. These observations show that while the stem in hairpin I is more stable than the stem in the hairpin II DNA studied by Haasnoot (Haasnoot et al., 1983), the reverse order of stability is observed for the four consecutive thymine residues in the loop region in the two hairpins. This difference in loop stability, as monitored by the thymine imino protons, might be due to the formation of an A·T pair between the terminal adenine residue of the stem with one of four thymines in the loop region (C. A. G. Haasnoot, private communication). In contrast, the guanine residue of the terminal G·C base pair in hairpin I is not likely to be involved in base pair formation with a thymine in the loop. Differences in loop structure, arising from stacking interactions with the terminal base pairs (G·C in our case, A·T in Haasnoot's molecule), may also contribute to this difference in liability of the thymine imino protons in two molecules.

In the 4 °C 2D NOE experiment presented in Figure 5,

Table III: Thermodynamic Parameter for Helix to Coil Transition in Hairpin I and $d(CG)_3$

	T_m (°C)	$\Delta H_{V.H.}$ (kcal mol^{-1})[a]	ΔS (cal mol^{-1} K^{-1})	$\Delta H_{V.H.}$ (kcal mol^{-1})[b]	ΔS (cal mol^{-1} K^{-1})	$\Delta H_{V.H.}$ (kcal mol^{-1})[c]	ΔS (cal mol^{-1} K^{-1})[c]
$d(CG)_3$	67	50 ± 2	147 ± 6	61 ± 2	179 ± 6	50 ± 2	147 ± 6
hairpin I	87			70 ± 2	195 ± 6	56 ± 2	156 ± 6

[a] The $\Delta H_{V.H.}$ was obtained from the slope of $1/T_m$ vs. log C plot. [b] The $\Delta H_{V.H.}$ was calculated by the equation $\Delta H_{V.H.} = 6RT_m^2(\partial\alpha/\partial t)_{tm}$. [c] The $\Delta H_{V.H.}$ was recalculated after the correction of $(\partial\alpha/\partial t)_{tm}$ to agree with footnote *a*.

there is a strong cross peak between one of the thymine methyl resonances and a GH8 resonance in hairpin I. Simple steric considerations require that these resonances be assigned to T_7 and G_6, and this means that the first thymine of the loop is stacked on the terminal G of the stem. In this way, the stacking pattern of the bases in the stem persists into the loop on the 5' side. The importance of the base at the 5' end of the loop has recently been emphasized by Haasnoot (Haasnoot et al., 1983), who demonstrated by 2D NMR that T_7 is stacked on A_6 in hairpin II.

Evaluation of the effective correlation times for base motion from cross relaxation measurements reveals that the correlation time for the thymine residues in the loop (1.3 ns) was shorter than that for the C residue (2.0 ns) in the stem region. These values, which may be compared with a correlation time of ~0.5 ns for thymine in tetramer duplex $d(ATAT)_2$, indicate that the loop formation has severely restricted the motional freedom of the thymines, but they still remain freer than the C residues in the stem.

Thermal Stability of Hairpin I and $d(CG)_3$. In general, the melting transition midpoint (T_m) of double-stranded DNA is a balance between two opposing reactions, namely, unimolecular dissociation of the helix and bimolecular reassociation of the single strand leading to a T_m that is dependent on DNA concentration. Indeed, $1/T_m$ is linearly related to the logarithm of DNA concentration by the equation (Borer et al., 1974)

$$1/T_m = \frac{2.3R}{\Delta H_{V.H.}} \log C + \frac{\Delta S}{\Delta H_{V.H.}} \quad (2)$$

where ΔS and $\Delta H_{V.H.}$ are entropy and van't Hoff enthalpy changes for helix-to-coil transition, R is the gas constant, and C is the total strand concentration. The values deduced for $d(CG)_3$ are $\Delta H_{V.H.} = 50 \pm 2$ kcal mol^{-1} and $\Delta S = 147 \pm 6$ cal mol^{-1} K^{-1}. In contrast, the van't Hoff enthalpy change ($\Delta H_{V.H.}$) for hairpin I cannot be determined by the $1/T_m$ vs. log C plot since the T_m of hairpin I is independent of concentration (Figure 6). However, $\Delta H_{V.H.}$ for hairpin I can be roughly calculated from the slope of the chemical shift vs. temperature curve according to the expression (Breslauer et al., 1975)

$$\Delta H_{V.H.} = 6RT_m^2(\partial\alpha/\partial t)_{tm} \quad (3)$$

where α represents the fraction of single strand. $\Delta H_{V.H.}$ values for $d(CG)_3$ and hairpin I were calculated to be 61 ± 2 and 70 ± 2 kcal mol^{-1} according to eq 2. However, since $(\partial\alpha/\partial t)_{tm}$ depends on how the base line in the chemical shift vs. temperature curve is determined, the resulting $\Delta H_{V.H.}$ value for $d(CG)_3$ calculated by eq 3 is less accurate than the $\Delta H_{V.H.}$ obtained from the slope of the $1/T_m$ vs. log C plot. Assuming that the slope of the chemical shift vs. temperature curve of CH6 in $d(CG)_3$ and hairpin I is the same at T_m, $\Delta H_{V.H.}$ for hairpin I was recalculated to be 56 ± 2 kcal mol^{-1} after the correction of $(\partial\alpha/\partial t)_{tm}$ to fit 50 ± 2 kcal mol^{-1} of $\Delta H_{V.H.}$ for $d(CG)_3$. The corresponding entropy change ΔS for hairpin I was also calculated to be 156 ± 6 cal mol^{-1} K^{-1}, since $\Delta S = \Delta H_{V.H.}/T_m$. These thermodynamic data are summarized in Table III. It is interesting to note that hairpin I, with only six G·C base pairs, has a T_m (87 °C) that exceeds the T_m of the duplex $d(CG)_8$ with 16 G·C base pairs, except in very concentrated solutions (greater than 1 mM) (see Table IV) (Gotoh & Tagashira, 1981). While the T_m of the hairpin is independent of concentration (see Figure 6), the T_m of the $d(CG)_8$ 16 base pair duplex decreases to 53 °C when the concentration is reduced to the concentrations that are typically used in molecular cloning experiments (typically $\leq 10^{-12}$ M) (Wallace et al., 1981). This points out a potential problem in carrying out hybridization experiments using probe oligonucleotides that are capable of forming hairpin structures, even with secondary structure base pairs that involve non-Watson–Crick base pairs (e.g., G·T, A·G, A·C). As a consequence, hairpin formation might easily compete with duplex formation in dilute solutions typically used in those experiments.

Table IV: Comparison of Calculated T_m for $d(CG)_8$, d(CGCGCGTATACGCGCG), and Hairpin I

T_m (°C)			
$d(CG)_8$	d(CGCGCGTA-TACGCGCG)	hairpin I	concn (M)
88	74	87	10^{-3}
76	62	87	10^{-6}
64	51	87	10^{-9}
53	40	87	10^{-12}

Imino Protons in Stem Region. The exchange of the imino protons in the stem region of hairpin I is reduced by the presence of the loop relative to $d(CG)_3$. This striking difference of the imino proton exchange behavior between $d(CG)_3$ and hairpin I can be ascribed in part (i) to a reduction in the rate at which the base pairs open (k_{op}) by the loop or (ii) to a reduced accessibility of the imino protons near the loop region to solvent and consequent reduction in the overall exchange rate (k_{ex}).

Conclusions

We have examined the secondary structure of synthetic oligodeoxyribonucleotide d(CGCGCGTTTTCGCGCG) by NMR, gel electrophoresis, and fluorescence polarization anisotropy. This 16mer forms only unimolecular hairpin structures, rather than bimolecular duplexes that might be expected to be more stable. Hairpin I is thermally much more stable than $d(CG)_3$. Imino proton exchange from G·C pairs in the stem region in hairpin I is much slower than that of $d(CG)_3$ under identical conditions (salt, pH), due to the presence of the loop in hairpin I. We have also compared the exchange behavior of hairpin I with d(ATCCTATTTTTAGGAT), another DNA that forms a hairpin with a six base pair stem and four thymines in a loop. Under the identical conditions, exchange of thymine imino proton in hairpin I is faster, indicating that the terminal base pairs adjacent to the loop can affect the loop structure/stability. 2D NOE experiments revealed that the first thymine residue in the loop region stacks with the guanosine residue at the end of the stem region as observed by Haasnoot (Haasnoot et al., 1983). These data demonstrate that these two similar hairpin DNAs have

stacking between thymine residues in the loop and the adjacent base at the end of the stem region, but the accessibility of the thymine imino protons in the loop are somewhat different in the two molecules. The motion of the thymine residues in the loop is restricted relative to smaller oligomers but remains freer than C residues in the stem of the hairpin.

ACKNOWLEDGMENTS

We thank T. Härd for excellent technical assistance in the FPA measurements and T. Novet for determining the DNA fragment lengths.

Registry No. d[$(CG)_3(TA)_2(CG)_3$], 94293-94-4; d[$(CG)_2TA(CG)_2$], 102368-21-8; d[$(CG)_3T_4(CG)_3$], 102419-87-4; d$(CG)_3$, 58927-26-7; d$(ATATCGATAG)_2$, 83685-27-2; d$(ATAT)_2$, 68232-81-5; d$(CG)_8$, 102419-86-3; d(ATCCTATTTTTAGGAT), 73202-01-4.

REFERENCES

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Barrell, B. G., & Clark, B. F. C. (1974) in *Handbook of Nucleic Acid Sequences*, pp 1–104, Joynson-Bruvvers, England.

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

Breslauer, K., Sturtevant, J., & Tinoco, I. (1975) *J. Mol. Biol. 99*, 549–565.

Cheng, D. M., Kan, L. S., Frechet, D., Ts'o, P. O. P., Uesugi, S., Shida, T., & Ikehara, M. (1984) *Biopolymers 23*, 775–795.

Clore, G. M., & Gronenborn, A. M. (1984) *FEBS Lett. 172*, 219–225.

Courey, A. J., & Wang, J. C. (1983) *Cell (Cambridge, Mass.) 33*, 817–829.

Dobson, C. M., Olejniczak, E. T., Poulsen, F. M., & Ratcliffe, R. G. (1982) *J. Magn. Reson. 48*, 97–110.

Erdmann, V. A. (1980) *Nucleic Acids Res. 8*, 31–47.

Feigon, J., Denny, W. A., Leupin, W., & Kearns, D. R. (1983) *Biochemistry 22*, 5930–5942.

Genest, D., Mirau, P. A., & Kearns, D. R. (1985) *Nucleic Acids Res. 13*, 2603–2615.

Germann, M. W., & Van de Sande, J. H. (1985) in *Fourth Conversation in Biomolecular Stereodynamics*, pp 42–43, State University of New York, Albany, NY.

Glotz, C., & Brimacombe, R. (1980) *Nucleic Acids Res. 8*, 2377–2396.

Gotoh, O., & Tagashira, Y. (1981) *Biopolymers 20*, 1033–1042.

Haasnoot, C. A. G., de Bruin, S. H., Berendsen, R. G., Janssen, H. G. J. M., Binnendijk, T. J. J., Hilbers, C. W., van der Marel, G. A., & van Boom, J. H. (1983) *Biomol. Stereodyn. 1*, 115–129.

Hore, P. J. (1983) *J. Magn. Reson. 54*, 539–542.

Jeener, J., Meier, B. H., Bachman, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Kumar, A., Ernst, R. R., & Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.

Lilley, D. M. J. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 6468–6472.

Lilley, D. M. J. (1981) *Nucleic Acids Res. 9*, 1271–1289.

Marky, L. A., Blumenfield, K. S., Kozlowski, S., & Breslauer, K. J. (1983) *Biopolymers* 22, 1247–1257.

Mizuuchi, K., Mizuuchi, M., & Gellert, M. (1982) *J. Mol. Biol. 156*, 229–243.

Morris, G. A., & Freeman, R. (1978) *J. Magn. Reson. 29*, 433–462.

Muller, U. R., & Fitch, W. M. (1982) *Nature (London) 298*, 582–585.

Panayotatos, N., & Wells, R. D. (1981) *Nature (London) 289*, 466–470.

Patel, D. J., Kozlowski, S. A., Ikuta, S., Itakura, K., Bhatt, R., & Hare, D. R. (1983) *Cold Spring Harbor Symp. Quant. Biol. 47*, 197–206.

Proudfoot, N. J., & Brownlee, G. G. (1974) *Nature (London) 252*, 359–362.

Rosenberg, M., & Court, D. (1979) *Annu. Rev. Genet. 13*, 319–351.

Schweizer, M. P., Chan, S. I., Helmkamp, G. K., & Ts'o, P. O. P. (1964) *J. Am. Chem. Soc. 86*, 696–700.

Steger, G., Hofmann, H., Fortsch, J., Gross, H. J., Randles, J. W., Sanger, H. L., & Riesner, D. (1984) *J. Biomol. Struct. Dynam. 2*, 543–571.

Tan, Z. K., Ikuta, S., Huang, T., Dugaiczyk, A., & Itakura, K. (1983) *Cold Spring Harbor Symp. Quant. Biol. 47*, 383–391.

Wagner, G., & Wuthrich, K. (1979) *J. Magn. Reson. 33*, 675–680.

Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E., & Itakura, K. (1981) *Nucleic Acids Res. 9*, 879–894.

Wells, R. D., Goodman, T. C., Hillen, W., Horn, G. T., Klein, R. D., Larson, J. E., Muller, U. R., Neuendorf, S. K., Panayotatos, N., & Stirdivant, S. M. (1980) *Prog. Nucleic Acid Res. Mol. Biol. 25*, 167–267.

Wemmer, D. E., Chou, S. H., Hare, D. R., & Reid, B. R. (1985) *Nucleic Acids Res. 13*, 3755–3772.