Volume 13 Number 17 1985 Nucleic Acids Research

Nuclear magnetic resonance and circular dichroism studies of a duplex - single-stranded hairpin loop equilibrium for the oligodeoxyribonucleotide sequence d(CGCGATTCGCG)

Michael F.Summers[1*], R.Andrew Byrd[1], Kathleen A.Gallo[2], Carrie J.Samson[2], Gerald Zon[2] and William Egan[1*]

[1]Biophysics Branch and [2]Molecular Pharmacology Branch, Division of Biochemistry and Biophysics, Food and Drug Administration, 8800 Rockville Pike, Bethesda, MD 20205, USA

Received 6 May 1985; Revised and Accepted 15 August 1985

ABSTRACT.

Nuclear magnetic resonance (NMR) and circular dichroism (CD) studies have been carried out with the oligodeoxyribonucleotide mismatch sequence, d(CGCGATTCGCG), 1. It has been found that 1 exists, in solution, as an equilibrium mixture of slowly interconverting, structured conformational isomers, 1a and 1b. On the basis of the concentration dependence of the 1a ⇌ 1b equilibrium, the ^{1}H NMR spectrum of the imino protons of the nucleotide bases, and the individual CD spectra of 1a and 1b, it is suggested that the two species correspond to a B-type DNA duplex and a single-stranded, hairpin-loop structure; the portion of the single-stranded species not involved in the loop appears to have a B-type DNA structure (on the basis of the CD measurements). To facilitate ^{1}H NMR resonance assignments, the two possible des-methyl thymidine derivatives of 1 were synthesized; the effect of this substitution on the physical chemical properties of 1 was explored. The ^{1}H NMR spectra of 1, as a function of temperature, showed that, under conditions wherein both species were present to a significant extent, the duplex form melted at a lower temperature than the single-stranded, hairpin loop structure.

INTRODUCTION.

The advent of commercially available, automated DNA synthesizers (1) has sparked interest in investigations of the interplay of linear sequence and external environment on the three dimensional structure and dynamics of oligodeoxyribonucleotides. During the past several years, the structural consequences of opposing Watson-Crick mismatches in doubly-stranded or potentially doubly-stranded DNA oligomers, as well as insertions (or, equivalently, deletions) of a particular base, have received considerable attention (2-5). These alterations in linear sequence have not, in general, been observed to markedly alter the overall oligonucleotide geometry relative to that of the parent compound, which usually has a B-type DNA structure (2-5).

In this context, we have had occasion to carry out solution studies (primarily using NMR spectroscopy) on a variant of the dodecamer [d(CGCGAATTCGCG)] studied crystallographically by Dickerson and Drew (6),

namely, the TxT mismatch sequence d(CGCGATTCGCG), 1. We herein report our findings that 1 can exist as an equilibrium mixture of a B-type DNA duplex and a single-stranded, hairpin loop structure. Additionally, we have prepared the two possible des-methyl thymidine (equivalently, 2-deoxyuridine) analogs of 1 and report their effect on the above-mentioned equilibrium as well as their utility in NMR resonance assignments.

EXPERIMENTAL

Reagents

Deuterium oxide ("100" atom % of ^2H) for NMR measurements was obtained from Aldrich Chemical Company (Metuchen, N.J.). Chelex-100 ion-exchange resin was obtained from Bio-Rad Laboratories (Richmond, CA). The phosphoramidite reagents used in the synthesis of the oligomers were obtained from Applied Biosystems (Foster City, CA) and the controlled-pore glass support for the synthesis was obtained from either Applied Biosystems or American Bionuclear (Emeryville, CA). All other analytical reagents, obtained from Sigma Chemical Co. (St. Louis, MO), were of the highest purity, and used without further purification.

Sample Preparation

The oligodeoxyribonucleotides were synthesized as described (7). Following an initial purification using HPLC techniques (7) the oligomers were twice precipitated from aqueous solution by the slow addition of ethanol (to 70 v/v %), and then passed through a column of Chelex-100 ion-exchange resin (prepared as described (8) and then additionally rinsed with a buffer of 1,4-piperazinebis(ethanesulfonic acid) [PIPES]); the oligonucleotide was eluted with a 1 mM solution of PIPES buffer at pH 7.0. The oligonucleotide solutions were then taken to dryness in a Savant refrigerated SpeedVac Concentrator. Concentrations of oligonucleotides were determined spectrophotometrically (at 260 nm, assuming extinction coefficients of 7000 per pyrimidine and 14,000 per purine), using a Hewlett Packard Model 8541 diode-array spectophotometer.

NMR Measurements

^1H and ^{31}P NMR spectral measurements performed using a JEOL GX-400 NMR spectrometer (399.65 MHz for ^1H; 161.7 MHz for ^{31}P) equipped with a variable temperature unit which maintained the temperature to ± 1 ^oC. Samples were prepared in D_2O containing PIPES (0.01M) and disodium ethylenediamine tetraacetate (EDTA) (0.01M). Values of solution pH were measured with a pre-calibrated standard glass electrode on a Radiometer model PHM82 pH-meter;

the "pH" values for D_2O containing solutions correspond to observed readings and were not corrected for possible isotope effects (9).

Proton NMR spectra of the imino protons were obtained in a D_2O/H_2O (1:20) mixture using the 1-3-3-1 water suppresion pulse technique (10). Relevant spectral parameters for the 1-3-3-1 sequence are as follows: 11.25° (1.3 μsec) pulse-width-1, 60 μsec pulse interval, 10 kHz spectral window, and 1 sec sequence repetition delay. Typically, the free-induction-decay (FID) signal was exponentially multiplied prior to Fourier-transformation so as to result in a 1 Hz additional line-broadening in the frequency-domain spectrum. Saturation transfer difference spectra were obtained by alternately adding and subtracting FID signals; prior to data acquisition, a low power level irradiation was applied to the selected resonance on alternate scans. The length of time for which the irradiation was applied was varied in successive experiments in order to determine a pre-exchange lifetime. Proton chemical shifts were measured relative to internal sodium 3-(trimethylsilyl)propionate 2,2,3,3-d_4 (TSP). ^{31}P NMR chemical shifts were measured relative to 1mM trimethylphosphate (aq.); a coaxial NMR tube assembly was utilized to facilitate chemical shift measurements.

Circular Dichroism Measurements

CD spectra were obtained, at ambient temperature (ca. 25 °C), using a JASCO 500A spectrophoto meter. Samples for CD measurements were taken directly from NMR tubes and placed into 0.1 cm pathlength cuvettes. ^1H NMR spectra were recorded before and after CD measurements in order to ensure a known value for the equilibrium distribution between **1a** and **1b**.

RESULTS AND DISCUSSION.

^1H NMR spectra of the thymidine methyl group protons of **1** as a function of temperature are shown in Figure 1. At elevated temperatures (> ca. 75 °C), two, equally intense, methyl group resonances are observed; their relatively low-field position is characteristic of the "melted" state (2, 4, 5), a conclusion that is substantiated by other NMR spectral observations (vide infra). At lower temperatures —0 to 45 °C (only data at temperatures above 25 °C are shown in Figure 1) —four methyl group resonances are observed, consisting of two sets of signals wherein the signal intensities within each set are equal. The two thymidine nucleotide units of **1** are chemically distinct and therefore give rise to two, equally intense, methyl group signals for each species present; barring accidental degeneracies of chemical shifts, and limiting considerations to the NMR chemical shift time-scale, two distinct

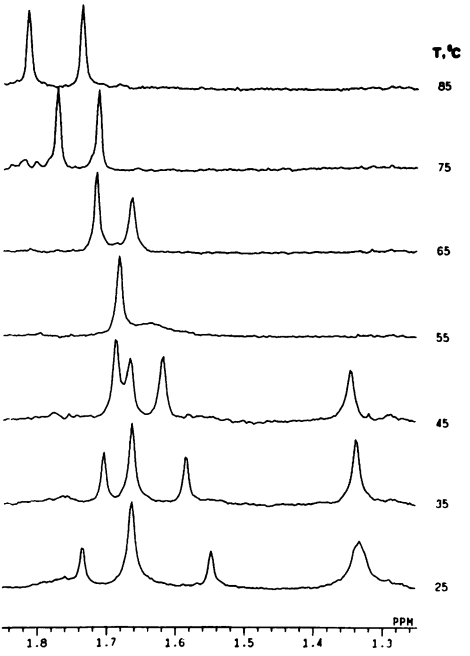


Figure 1. ^{1}H NMR spectra for the region 1.25 - 1.85 ppm, obtained at the indicated temperatures, for a D_2O solution of 1 (1.1 mM) containing PIPES buffer (10 mM), EDTA (1 mM), and TSP; pH = 7.01.

conformational states of 1 are thus present. The two species (hereafter designated 1a and 1b) represented by these signal sets are in thermodynamic equilibrium as evidenced by (1) the reversibility of changes in composition effected by perturbations of the external environment (e.g., changes in ionic strength or temperature) and (2) NMR saturation transfer experiments (11).

Saturation transfer NMR experiments were carried out with 1 at several temperatures. The results of one such experiment are presented in Figure 2, where a transfer of magnetization is observed between one set of corresponding methyl groups. Similar results (not shown) were obtained using the other set of methyl group resonances. The saturation transfer experiment provides pre-exchange lifetimes for species involved in chemical equilibria (11). At 40 °C, the pre-exchange lifetime for 1a was 0.7 ± 0.1 sec. (see Figure 3). (The pre-exchange lifetime was gotten from a least-squares fit of the data to either an exponential or linearized function (with similar results) using the National Institutes of Health's MLAB facilities (12). The saturation transfer experiment also serves to assign

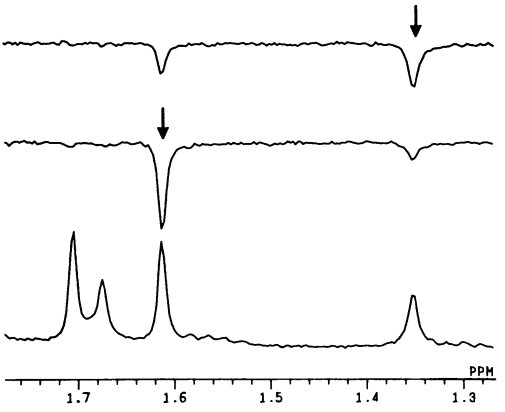
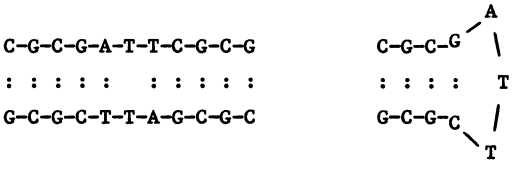


Figure 2. ^{1}H NMR spectra (40 °C) illustrating magnetization transfer between the T6-CH_3 groups of 1a (1.36 ppm) and 1b (1.61 ppm) in D_2O solution containing PIPES, EDTA, and TSP; sample conditions are as given in the legend to Figure 1; pH = 7.00. Arrows indicate irradiation (saturation) frequencies. The normal spectrum is presented in the bottom tracing and the two difference spectra illustrating magnetization transfer from T6-CH_3 of the two equilibrating species are presented in the middle and upper traces.

corresponding methyl groups in the two conformational isomers.

Results, fully in accord with the above findings of a conformational equilibrium for 1, involving two states that are distinct from the melted state, were obtained using either the aromatic resonances of the ^{1}H NMR spectrum or the phosphodiester resonances of the ^{31}P NMR spectrum of 1 (data not shown). These spectra, while corroborating the existence of a conformational equilibrium, did not directly help to establish the nature of the species involved. By contrast, the ^{1}H NMR spectra of the imino protons were more informative. These spectra, which will be discussed in greater detail, showed the presence of non-exchanging, Watson-Crick hydrogen bonded imino protons in both species. The accumulated NMR data thus provides evidence for the presence of two structured conformers of 1. We suggest that the two forms are a B-type DNA duplex (1a) and a single-stranded, hairpin loop structure (1b); the equilibrium is shown schematically below:



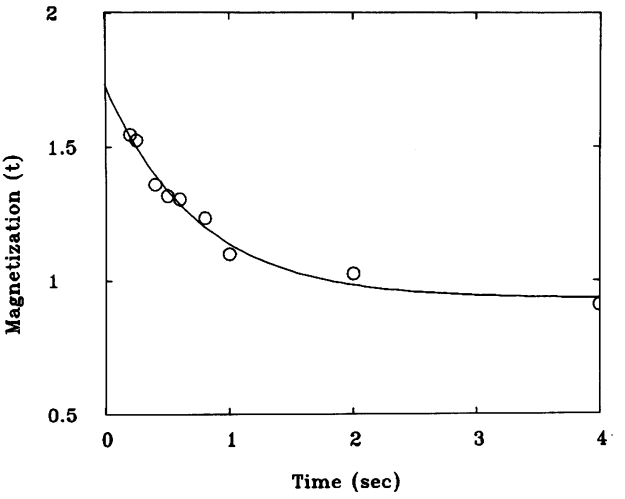


Figure 3. Plot of ^1H magnetization for one of the thymidine methyl groups of 1b (arbitrary units) versus time of irradiation for the corresponding methyl group of 1a (40 °C). Sample conditions are the same as those given in the legend to Figure 1. The solid line represents the least-squares fit to an exponential function (see reference 11) , while the open circles correspond to the data points. The infinity time for the magnmetization transfer experiment corresponded to 12 seconds.

For an equilibrating mixture of monomer and dimer, as suggested above, the concentration of either structure, in solution, should be dependent on the total oligonucleotide concentration (measured as single-stranded species). Provided that there are no intermediates involved in the equilibrium, and that the ionic strength of the solution is kept constant, and assuming ideal behavior, then {1b}, the concentration of hairpin loop structure in solution, will be given by the expression,

$$\{1b\} = [-1 + \sqrt{(1 + 8*K_{eq}*C)}]/[4*K_{eq}] \quad (1)$$

where K_{eq} is the equilibrium constant ([duplex]/[hairpin loop]^2) and C is the total concentration of single strands. A dependence of hairpin loop concentration on total oligonucleotide concentration was observed over the investigated concentration range (0.2 mM -0.9 mM at constant ionic strength); these results, shown in Figure 4, yielded a value for K_{eq} of 0.85 ± 0.15 mM^{-1} (determined by fitting the equilibrium data to equation 1 using MLAB). Under the conditions that were employed in the experiment, we did not observe signals suggestive of a single-stranded, random coil form of 1 (or other

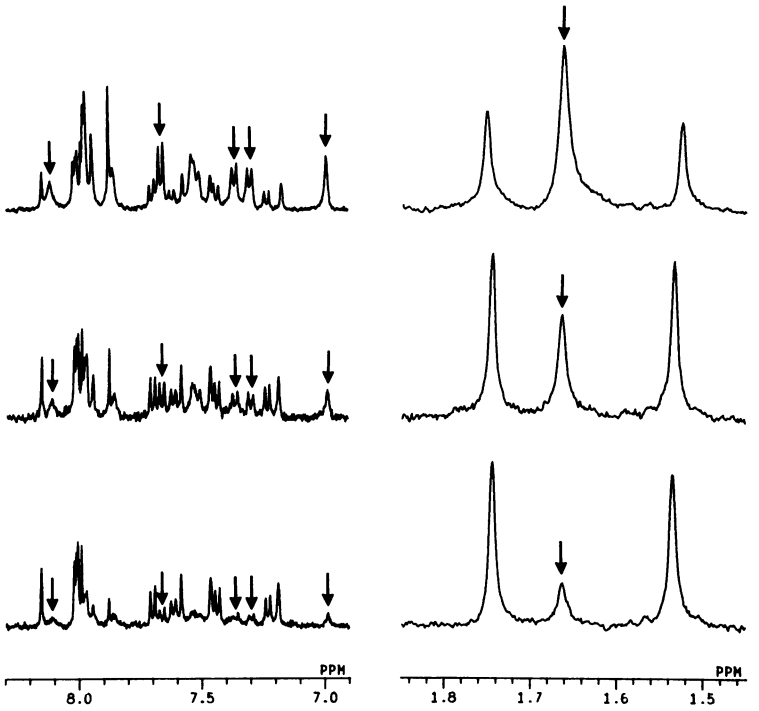


Figure 4. ^1H NMR spectra (6.9 - 8.3 ppm and 1.45 - 1.85 ppm regions) for 1 at 23 °C in D_2O solution as a function of oligonucleotide concentration. Sample conditions are as given in the legend to Figure 1; pH = 7.03. The concentrations of 1 were 0.9 mM (top trace), 0.4 mM (middle trace), and 0.2 mM (bottom trace). Arrows indicate selected signals for the duplex structure.

intermediate species). The resonances that became more pronounced at the lower concentrations are assigned to the monomeric form of 1 (viz., 1b) and the other methyl group resonances, namely, those that became more pronounced at the higher concentrations, are assigned to the duplex form of 1 (viz., 1a).

In order to confirm that only two species were giving rise to the observed methyl group resonances, and to establish absolute thymidine methyl group assignments, the two possible des-methyl thymidine analogs of 1 were synthesized. ^1H NMR spectra of the methyl group region of one of these (wherein thymidine-7 is replaced) is shown in Figure 5. Only two signals, one for each of the two species in solution, were observed. Notably, the chemical shifts for the methyl groups on the remaining thymidines (and, indeed, nearly all of the remaining proton resonances) were unaltered relative to that of the

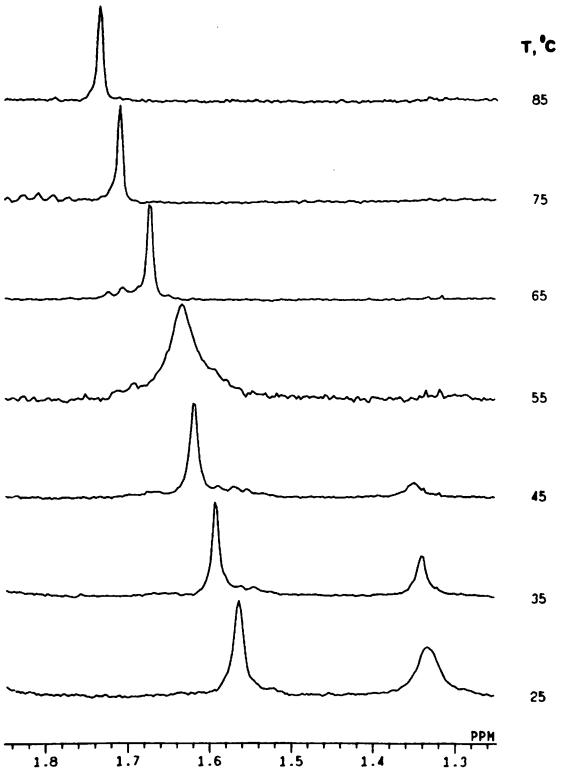


Figure 5. ^1H NMR spectra (1.25 - 1.85 ppm) obtained at the indicated temperatures for a D_2O solution of the des-methyl (T7) analog of 1 under sample conditions as given in the legend to Figure 1; pH = 7.01.

parent compound and, moreover, the equilibrium distribution between the hairpin loop and duplex states was only moderately affected. Replacement of the other thymidine nucleotide (namely, thymidine-6) with its des-methyl congener gave similar results with regard to the observed chemical shifts and the equilibrium distribution of the two species. The chemical shifts for the thymidine methyl resonances of the duplex structure (in the two des-methyl congeners) were nearly the same as those found in the Dickerson dodecamer (13).

The imino proton region of the ^1H NMR spectrum of 1 at two different concentrations is shown in Figure 6. Using the rationale for assigning resonances based on their concentration dependence, six imino proton resonances are assigned to the duplex form of 1 and another set of 6 resonances are assigned to the hairpin-loop form of 1. The relative

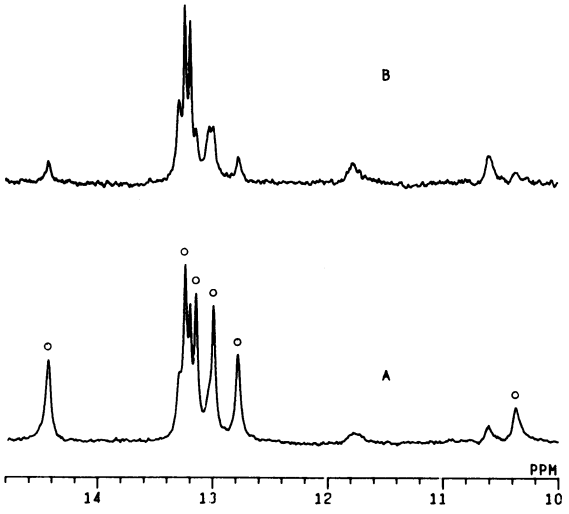


Figure 6. ^{1}H NMR spectra for the N-H protons of la (o) and lb (no symbols over resonances) obtained at 5 °C in a 95 v/v% H_2O/D_2O solution at oligonucleotide concentrations of 0.9 mM (A) and 0.2 mM (B). Other sample conditions are as given in the legend to Figure 1; pH = 7.01.

intensities of the two sets of imino protons were in agreement with the intensities derived from the methyl group protons. At low concentrations, and with no added salt, the monomeric species dominates the equilibrium (Figure 6b). Under these conditions, four non-exchanging Watson-Crick hydrogen bonded imino proton resonances and two, non-base-paired, thymine H-1 resonances are observed. [At very low concentrations (less than 0.1 mM), only the monomeric form is observable in the ^{1}H NMR spectrum.] We interpret these later observations in terms of a hairpin loop structure wherein the base pairs involved in the loop (viz., ATT) are not hydrogen bonded to other nucleotide bases. The terminal base-pair hydrogen bond is exchanging rapidly with the solvent at 25 °C (and is thus not observed) , but is observable at low temperatures (ca. 5 °C and lower).

Under conditions of 1M added salt and at an oligonucleotide concentration of 0.1 mM, the duplex form predominates; under these circumstances, five Watson-Crick hydrogen bonded imino protons are observed (whose chemical shifts are essentially the same as those seen in the Dickerson dodecamer (13)); an additional imino resonance is observed at 10.4 ppm, and corresponds to the ^{1}H proton of the TxT mismatch. Presumably, this imino proton is not involved in a hydrogen bond to another nucleotide; its particular resonance position is

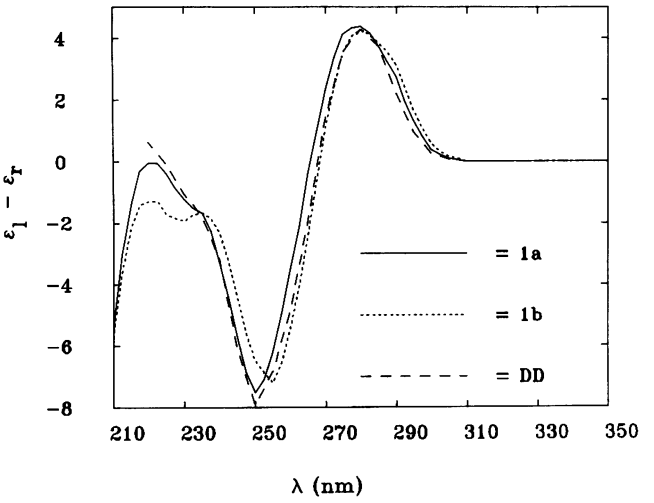


Figure 7. CD spectra for the Dickerson dodecamer (DD), 1a, and 1b (the oligonucleotide concentrations were 0.08 mM for each of the samples); the 1a to 1b equilibrium shift was effected through the addition of NaCl (see Text).

due to shielding by adjacent bases, and indicates that the thymine base is tucked into the helix as opposed to being pushed out into the solvent(14a).

In order to shed more light on the nature of structures 1a and 1b, their CD spectra, as well as that of the Dickerson dodecamer, were obtained (see Figure 7). The spectra were recorded under conditions (vide supra) where, alternately, only 1a or 1b was present (see Material and Methods). The CD spectra for the three species are remarkably similar, at least at wavelengths greater than 240 nm, and, moreover, have the appearance typically found for a B-type DNA structure (15). At intermediate concentrations of added salt, both 1a and 1b are observed in the NMR spectrum; the CD spectra, under these conditions, appear as linear combinations of those found for the basis structures, 1a and 1b. Finally, the CD spectra establish that 1a and 1b are structured conformational isomers.

The assembled data argue for the existence, in solution, of two, slowly interconverting, structured conformers of 1, one of which is a "normal" B-type DNA structure and the other of which is a hairpin-loop structure having, in part (i.e., the four terminal bases from each end) a right-handed B-type DNA structure. Two-dimensional NOE studies with 1 and closely related oligonucleotides that are now in progress (14), have furnished information about the structure of the loop region. In the context of the present report,

these NOE studies serve to establish that it is the monomer (1b) that is dominant at low salt/low concentration. Moreover, the NOE build-up rates directly evidence that the low salt/low concentration species is the smaller molecule, i.e., has a shorter rotational correlation time.

Hairpin loop structures have been previously observed in DNA oligomers by Haasnoot et al. (15) and Patel et al. (16) using proton NMR spectroscopy and by Marky et al. (17) using chiroptic and calorimetric methods. Haasnoot et al. (14) observed that, for their sequences [d(ATCCTA(T)$_n$TAGGAT), where n = 2, 3, 4, 5], the duplex forms melted at a higher temperature than the hairpin-loop structures and the smallest loop structure was, most likely, composed of four base pairs. In contrast, we find that the hairpin loop structure melts at a higher temperature than that of the duplex structure, and, moreover, the loop structure is composed of only three nucleotides. Marky et al. (17) found that, in the Dickerson dodecamer, the hairpin loop structure melted after the duplex (as we have observed) and that the loop comprised four nucleotide bases. Patel et al. (15) observed a hairpin loop form for the Dickerson dodecamer; melting studies were, however, not reported. Undoubtedly, the number of nucleotides in a loop structure, as well as the loop's stability relative to a duplex structure, will depend on a number of factors, including the nature of and position of the constituent nucleotides and the parity of the linear sequence (even or odd number of nucleotides).

ACKNOWLEDGEMENTS.

We would like to thank: Professor L. Marzilli (Emory University) for suggesting the use of the Chelex-100 ion-exchange resin in our purification scheme (for removal of residual triethylammonium acetate); Dr. Peter Roller (National Cancer Institute, National Institutes of Health) for use of the CD spectrometer; the Sigma Xi Foundation for a Grant-in-Aid of Research (to M.F.S.).

*To whom correspondence should be addressed

REFERENCES

1. "Chemical and Enzymic Synthesis of Gene Fragments: A Laboratory Manual," Gassen, H.G. & Lang, A., Eds., Verlag Chemie, Deerfield Beach, FL, 1982.
2. Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984) Biochemistry 23, 3218 - 3226.
3. Keepers, J. W., Schmidt, P., James, T. L., & Kollman, P. A. (1984) Biopolymers 23, 2901 - 2929.
4. Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., & Breslauer, K. J. (1982) Biochemistry 21, 445 - 451.

5. Tibanyenda, N., de Bruin, S. H., Haasnoot, C. A. G., van der Marel, G. A., vanBoom, J. H., & Hilbers, C. W. (1984) Eur. J. Biochem. 139, 19 - 27.
6. Dickerson, R. E. & Drew, H. P. (1981) J. Mol. Biol. 149 761 - 768; Drew, H. R. & Dickerson, R. E. (1981) J. Mol. Biol. 151 535 - 556.
7. Stec, W. J., Zon, G., Egan, W., Byrd, R. A., Phillips, L. R., & Gallo, K. A. (1985) J. Org. Chem. in press.
8. Egan, W., Schneerson, R., Werner, K. E., & Zon, G. (1982) J. Amer. Chem. Soc. 104 2898 - 2910.
9. Lumry, R., Smith, E. L., & Glantz, R. R. (1951) J. Amer. Chem. Soc. 73 4330 - 4340.
10. Hore, P. J. (1983) J. Mag. Reson. 54 539 - 542.
11. Forsen, S. & Hoffman, R. A. (1963) J. Chem. Phys. 39 2892 - 2901.
12. Knott, G. D. (1979) Comput. Progr. in Biomed. 10 271 - 280.
13. Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C. Rice, J. A., Itakura, K., & Breslauer, K. J. (1982) Biochemistry 21 428 - 436.
14. Summers, M. F., Bax, A., Byrd, R. A., Pastor, R. W., Brooks, B. R., Zon, G., & Egan, W. (1985), unpublished results.
15. (a)Cornelis, A. G., Haasnoot, J. H. J., den Hartog, J. F., de Rooij, M., van Boom, J. H., & Cornelis, A. (1979) Nature 281 235 -236;(b) Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, J. F. M., van Boom, J. H., & Altona, C. (1980) Nucl. Acids Res. 8 169 -181; Haasnoot, C. A. G., de Bruin, S. H., Berendsen, R. G., Janssen, H. G. J. M., Binnedijk, T. J. J., Hilber, C. W., van der Marel, G. A., & van Boom, J. H. (1983) J. Biomol. Struct. Dynam. 1 115 -129.
16. 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.
17. Marky, L. A., Blumenfeld, K. S., Kozlowski, S. A., & Breslauer, K. J. (1983) Biopolymers 22 1247 - 1257.