

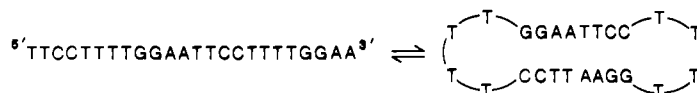
A Dumbbell-Shaped, Double-Hairpin Structure of DNA: A Thermodynamic Investigation[†]

Dorothy Erie,[‡] Navin Sinha,[§] Wilma Olson,[‡] Roger Jones,[‡] and Kenneth Breslauer^{*:‡}

Department of Chemistry and Waksman Institute of Microbiology, Rutgers—The State University of New Jersey, Piscataway, New Jersey 08854

Received February 17, 1987; Revised Manuscript Received June 1, 1987

ABSTRACT: We report the first calorimetric and spectroscopic investigation on a member of a new class of nucleic acid secondary structures in which *both* ends of a duplex core are closed by single-stranded loops. Such structures can be formed intramolecularly from appropriately designed base sequences. We have synthesized the 24-mer sequence shown, and we present calorimetric, spectroscopic, and electrophoretic



evidence that it adopts a dumbbell-shaped, double-hairpin structure. Our data allow us to reach the following conclusions: (1) The phosphodiester gap in the center of the core duplex of the dumbbell does not reduce the transition enthalpy relative to that measured for the corresponding octameric duplex d(GGAATTC)₂. (2) Incorporation of a 5'-phosphate group into the gap decreases the thermal stability of the dumbbell relative to its unphosphorylated sequence. On the basis of the salt dependence of this effect, we propose that the phosphorylation-induced decrease in thermal stability is electrostatic in origin. From the changes in the transition enthalpy and entropy, we suggest that the phosphorylation-induced decrease in thermal stability of the double hairpin arises from electrostatically induced based unstacking at the nick. (3) The thymine residues in the loop behave both electrostatically and enthalpically like denatured single strands. Published nuclear magnetic resonance studies reveal partial stacking of thymine residues in the loops of linear hairpin structures. If this feature persists in the double-hairpin structure, then the spatial overlap of thymine residues in the loops does not necessarily produce a favorable enthalpic contribution. (4) When both ends of the nicked octameric core duplex are constrained by loops of only four thymine residues, the dumbbell structure may adopt conformations in which the 5' and 3' ends at the nick are twisted relative to the helical axis and therefore are not in phase. Such conformations would account for the observed resistance of the double-hairpin structure to ligation, since the 3'OH and 5'P would no longer be collinear.

A large number of laboratories have investigated the effect of sequence, conformation, and drug binding on the order-disorder transitions of specifically designed and synthesized oligonucleotides (Bloomfield et al., 1974, and references cited therein; Cantor & Schimmel, 1980, and references cited therein). These studies have provided a great deal of information about the molecular forces that dictate and control the structural stability, conformational preferences, and drug-binding capabilities of nucleic acid sequences. Nevertheless, many important questions concerning sequence-dependent conformational preferences remain unanswered due to deficiencies in the oligomeric and polymeric model systems studied to date. For convenience and clarity, we list some of these limitations: (1) In typical model oligomeric compounds, the ends frequently represent a significant fraction of the structure, thereby exerting a disproportionate effect on the physical properties under investigation. (2) The bimolecular melting transitions of most dimeric helices (e.g., double helices formed by the association of two complementary strands) do not re-

semble the local monomolecular melting within a natural nucleic acid polymer. (3) The melting of most oligomeric duplexes yields "linear" single strands since the ends are "open", while the local melting within a nucleic acid polymer yields an interior "bubble" or circle since the ends of the melting domain are constrained. [The pioneering studies of Baldwin and co-workers represent a notable exception (Scheffler et al., 1970; Elson et al., 1970; Baldwin, 1971).]

To avoid some of the shortcomings listed above, we have synthesized oligomeric DNA sequences which are designed to form double-hairpin structures possessing closed or constrained ends that undergo monomolecular order-disorder transitions. Optimally, the constraints of chain closure in this new class of model oligomeric duplexes should not affect the ability of the "core duplex" to undergo an order-disorder transition. As illustrated schematically in Figure 1, a sequence in which both ends can fold back on themselves to form a dumbbell-shaped molecule should satisfy these requirements. Once ligated, the final melted states of these dumbbell structures—covalently closed, single-stranded circles—will more closely resemble the final states formed by the local domain melting of natural DNA since the ends are constrained and their melting process is monomolecular. In fact, others have proposed the formation of somewhat larger dumbbell

[†]This work was supported by NIH Grants GM34469, GM233509, GM31483, GM20861, and GM24391, American Cancer Society Grant CA248B, and the Busch Memorial Fund.

[‡]Department of Chemistry.

[§]Waksman Institute of Microbiology.

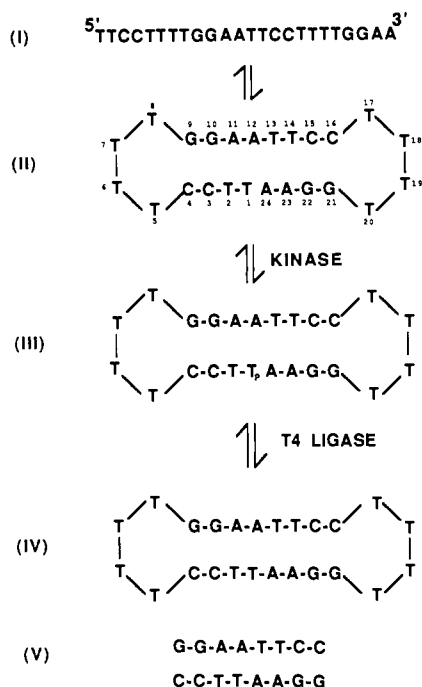


FIGURE 1: Schematic of the oligomeric system under investigation: (I) primary sequence; (II) gapped, unkinased dumbbell; (III) nicked, kinased dumbbell; (IV) ligated, closed circular dumbbell; (V) core duplex.

structures as possible forms of a 66 base pair operator sequence (Betz & Sadler, 1981).

In light of the considerations noted above, we have designed and synthesized the 24-mer sequence d-(TTCCTTTTGAATTCCTTTTGGAA). We had three reasons for selecting this sequence as our first target: (1) The sequence possesses both 5' and 3' domains that can loop back on themselves to form a dumbbell-shaped, double-hairpin structure (see Figure 1). (2) As illustrated in Figure 1, the expected structure will have a gap between the 5' and 3' ends. Phosphorylation of this "gapped-dumbbell" using T4 polynucleotide kinase would produce a nicked double-hairpin structure with a 5'-phosphate in the *interior* of the helix. If this phosphorylated and nicked dumbbell could be ligated with T4 DNA ligase, it would produce the desired single-stranded, closed circular DNA. (3) The octameric core duplex formed in this double-hairpin structure, d(GGAATTCC)₂, provides a well-defined reference state for interpreting experimental data. We and others (Patel & Canuel, 1979; Patel et al., 1982; Connolly & Eckstein, 1984; Breslauer et al., 1986; Breslauer and Marky, unpublished data) previously have characterized its structural and thermodynamic properties.

MATERIALS AND METHODS

Synthesis. The 24-mer sequence d-(TTCCTTTTGAATTCCTTTTGGAA) was synthesized on an Applied Biosystems Model 380A automated DNA synthesizer, employing the phosphoramidite method (Carruthers et al., 1982). The protecting groups were removed by overnight treatment with concentrated ammonium hydroxide at 65 °C. The 24-mer was purified first by gel electrophoresis in 20% acrylamide under denaturing conditions (7 M urea) and then by HPLC.¹ Electrophoresis was carried out at a constant

voltage of 700 V in Tris-borate-EDTA gel buffer (1× TBE). The gels were UV-shadowed, and the band that traveled the shortest distance was cut out and eluted overnight in 0.5 M ammonium acetate at room temperature. The eluent was desalted by using a LIDEX MINISEP C18 cartridge; the oligomer was eluted in 30% acetonitrile-water. The isolated 24-mer was analyzed by HPLC (Waters C₁₈ Nova-PAK). There was one major peak followed by a few minor peaks. Consequently, the oligomer was further purified by HPLC, as described previously (Gaffney et al., 1984). The final product was a single peak on HPLC and a single electrophoretic band on a 20% polyacrylamide, 7 M urea gel (after labeling with [γ -³²P]rATP).

Phosphorylation. Solutions of 5.0–50 nmol of 24-mer ([oligomer] = 5–50 μ M) in 70 mM Tris-HCl (pH 7.4), 10 mM MgCl₂, 300 μ M rATP/[γ -³²P]rATP, and 1 mM dithiothreitol (DTT) were incubated with 5 units/nmol of oligomer of T4 polynucleotide kinase (Bethesda Research Laboratories; Bethesda, MD) at 37 °C. After 30 min, a twofold excess of kinase and rATP was added. After an additional 90 min, the reaction was terminated by heating the mixture for 10 min at 65 °C. The extent of the phosphorylation reaction was monitored by chromatographing aliquots of the reaction mixture on a Sephadex G-50 column and comparing the amount of labeled oligomer to free [γ -³²P]rATP. Phosphorylation was found to be 100% for all samples kinased.

Ligation. After the above reaction mixture was cooled, rATP and DTT were added such that their final concentrations were 1 and 5 mM, respectively, followed by a large excess of T4 DNA ligase. The same enzyme stock was used to ligate similar small molecules under the same reaction conditions. This result confirms that the T4 DNA ligase was active. The ligation reactions were carried out over a range of temperatures (0–37 °C), polyethylene glycol (PEG) concentrations (0–15%), and incubation periods (1–48 h). PEG has been shown to enhance the ligation of small molecules (Pheiffer & Zimmerman, 1983; Zimmerman & Pheiffer, 1983). The extent of ligation was assayed by removing aliquots of the reaction mixture and measuring the resistance of the molecule to degradation by exonucleases (T4 polymerase, λ exonuclease, Exo III) or to dephosphorylation by alkaline phosphatase.

Exonuclease Assays. All exonuclease assays were performed at 37 °C for 30 min. The conditions for the T4 polymerase and Exo III assays were 33 mM Tris-OAc, pH 7.8, 10 mM MgOAc₂, 66 mM KOAc, 100 μ g/mL nuclease-free bovine serum albumin, and 0.5 mM DTT. The conditions for the T4 polymerase assays were 1 mM DTT, 3 mM MgOAc₂, and 67 mM glycine, pH 9.4. The alkaline phosphatase assays were performed in 10 mM Tris-OAc at 65 °C for 30 min. All of the above reactions were terminated by the addition of 0.2 M EDTA such that its final concentration was 20 mM. After the reactions were terminated, the reaction mixtures were chromatographed on a Sephadex G-50 column, and the percent of oligomer that still contained label was measured.

Circular Dichroism (CD) Spectroscopy. CD spectra were recorded with a Model 60DS AVIV spectropolarimeter (Lakewood, NJ) equipped with a thermoelectrically controlled cell holder. All solutions contained 10 mM phosphate buffer (pH 7.0), 1 mM EDTA, and no added NaCl.

Ultraviolet (UV) Spectroscopy. Absorbance versus temperature profiles were measured at 260 nm on a Perkin-Elmer 575 spectrophotometer interfaced to a Tektronix 4051 computer. Melting profiles were obtained by increasing the temperature from 0 to 100 °C at a constant rate of 1.0 °C/min by using a programmable, thermoelectrically controlled cell

¹ Abbreviations: HPLC, high-performance liquid chromatography; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; ATP, adenosine 5'-triphosphate; DSC, differential scanning calorimetry; CD, circular dichroism; NMR, nuclear magnetic resonance.

holder. Oligomer concentrations ranged from 0.6 to 50 μM . These concentrations were determined spectroscopically by using extinction coefficients calculated by the nearest-neighbor method (Cantor et al., 1970). The melting studies also were performed over a range of salt concentrations from 0.002 to 1.0 M. All solutions contained 0.1 mM EDTA and 1.0 mM sodium phosphate buffer (pH 7.0).

The van't Hoff transition enthalpies (ΔH_{vH}) were determined by converting the experimental absorbance versus temperature profiles into α versus temperature curves, where α is the fraction of single strands. This conversion is dependent on the assignment of base lines and the assumption that the fractional change in absorbance at any temperature is directly proportional to the extent of reaction at that temperature. Sloping base lines were used. The van't Hoff transition enthalpies as well as the corresponding entropies can be obtained from these curves by using protocols described previously (Marky et al., 1983; Breslauer, 1986, and references cited therein; Marky & Breslauer, 1987).

Analysis of the Salt-Dependent Melting Data. UV melting profiles were determined over a broad range of sodium ion concentrations (0.002–1.0 M). The corresponding thermodynamic data were obtained by evaluating the experimental curves according to the methods described above. Since double-stranded DNA has a greater charge density than single-stranded DNA, denaturation is expected to release thermodynamically bound counterions. This counterion dependence is included in the following expressions for any order-disorder transition and its corresponding equilibrium constant:



$$K_{\text{eq}} = [\text{DNA}(\text{denatured})][\text{M}^+]^{\Delta n} / [\text{DNA}(\text{native})] \quad (1)$$

In this formulation, M^+ is any monovalent cation and Δn the change in the number of thermodynamically bound counterions per double-stranded region undergoing denaturation. The value of Δn can be determined experimentally by evaluating how the equilibrium constant, K_{eq} , varies with salt concentration, since $(\partial \log K_{\text{eq}} / \partial \log [\text{M}^+])_T = \Delta n$. More frequently, however, the salt dependence of the melting temperature, T_m , is measured since the melting temperature represents the most convenient and experimentally accessible parameter. The value of $dT_m / d \log [\text{M}^+]$ is proportional to the change in the thermodynamic degree of ion dissociation, Δi , that accompanies denaturation, where $\Delta i = i_{\text{denatured}} - i_{\text{native}}$. The specific relationship between $dT_m / d \log [\text{M}^+]$ and Δi is given by

$$dT_m / d \log [\text{M}^+] = (2.3RT_m^2 / \Delta H_{\text{obsd}}^{\circ}) \Delta i \quad (2)$$

where T_m is the melting temperature, $[\text{M}^+]$ the counterion concentration, $\Delta H_{\text{obsd}}^{\circ}$ the enthalpy change per nucleotide, and Δi the change in ion dissociation per phosphate. [This equation can be derived from $(\partial \log K_{\text{eq}} / \partial \log [\text{M}^+])_T$ by use of the chain rule for partial derivatives, as previously shown (Record et al., 1978).] Δn and Δi differ by a multiplicative factor that is equal to the number of phosphates involved in denaturation. In this work, eq 2 is used to calculate Δi from the salt-dependent T_m values and the enthalpy data.

Calorimetry. The change in heat capacity for the thermally induced, order-disorder transition of the 24-mer sequence was measured as a function of temperature on a Microcal 2 differential scanning calorimeter. The calorimetric experiments were carried out in 0.1 mM EDTA, 1.0 mM sodium phosphate buffer (pH 7), and no added NaCl. The oligomer concentration was 49 μM in single strands. The details of the experimental procedure and methods of data analysis have been described in previous studies (Breslauer et al., 1975; Marky

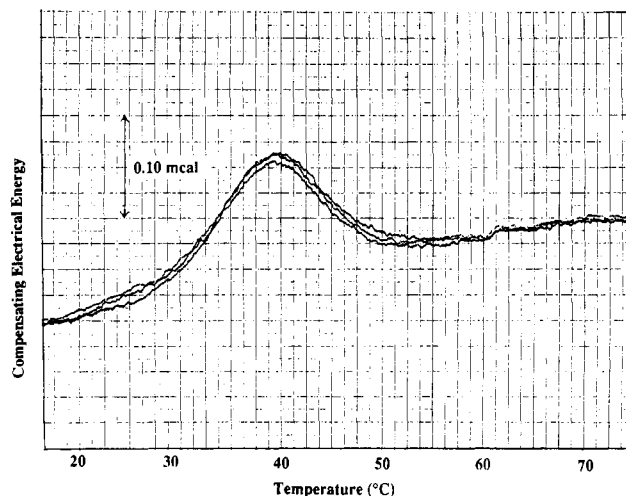


FIGURE 2: Experimental calorimetric melting profiles for the thermally induced, order-disorder transition of the structure formed by the 24-mer sequence d(TTCCT₄GGAATTCCT₄GGAA).

et al., 1983; Breslauer, 1986).


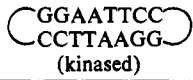
Gel Electrophoresis. Analysis of the size of the native form of the 24-mer sequences was performed by using polyacrylamide gel electrophoresis under nondenaturing conditions. The gel contained 20% acrylamide, 0.75% bis(acrylamide), and 1 \times TBE. The gel was run at a constant voltage of 300 V at room temperature until the bromophenol blue marker band was two-thirds down the plate. The gel was stained with ethidium bromide and was visualized with a short-wavelength UV transilluminator.

RESULTS

Calorimetry. The calorimetric heat capacity curves obtained for the thermally induced, order-disorder transition of the structure formed by the unphosphorylated 24-mer sequence in 0.1 mM EDTA, 1.0 mM sodium phosphate buffer (pH 7.0), and no added NaCl are shown in Figure 2. Corresponding calorimetric studies were not conducted on the phosphorylated 24-mer sequence since only small amounts were kinased with labeled rATP; hence, the phosphorylated sequence was not produced in sufficient quantities for calorimetric measurements. Two important features of the calorimetric measurements on the unphosphorylated 24-mer should be noted. First, the initial calorimetric scan and the rescans are nearly superimposable, indicating that the transition is reversible. Second, the calorimetric experiments were performed on a solution in which the oligomer concentration was 49.0 μM in single strands (which corresponds to a 260-nm absorbance of approximately 10 for a 1-mL solution in an optical cell with a 1-cm path length). This is a considerable decrease in the amount of material normally required for DSC studies. This measurement is possible due to instrumental improvements that have resulted in enhanced sensitivity. Significantly, this reduction in the quantity of material required for DSC studies allows for a better correlation between optical and calorimetric melting studies.

We determined the calorimetric transition enthalpy (ΔH_{cal}) for the unphosphorylated 24-mer structure by integrating the area under each experimental curve. We also calculated the van't Hoff transition enthalpy ($\Delta H_{\text{cal}}^{\text{vH}}$) from the shape of the calorimetric transition curve using previously described methods (Breslauer et al., 1975; Marky et al., 1981; Breslauer, 1986; Marky & Breslauer, 1987). The resulting calorimetric transition enthalpy, obtained by averaging six runs, appears in Table I. Note that the average value equals 54 kcal/mol

Table I: Thermodynamic Data^a

structure	<i>f</i> (concn)	ΔH° (kcal/mol)	ΔS° (eu)	$\partial T_m / \partial \log [\text{Na}^+]$	Δi (phosphate ⁻¹)	Δi^b (phosphate ⁻¹)
GGAATTCC CCTTAAGG (core duplex)	yes	58		12.3	0.096	0.096
 (unkinased)	no	54	173	11.5	0.057	0.088
 (kinased)	no	45	143	13.4	0.057	0.084

^a Thermodynamic data for the unkinased and kinased 24-mer sequences and their corresponding core duplex d(GGAATTCC)₂. The enthalpy data for the core duplex and the unkinased sequence were determined by using calorimetry, and data for the kinased sequence were determined from spectroscopy. ^b Δi was calculated by assuming that the loop phosphates behave electrostatically as single-stranded phosphates.

Table II: Salt-Dependent Thermodynamic Data^a

[Na ⁺] (mM)	<i>T_m</i> (°C)		ΔH_{vH} (kcal/mol)		ΔS (cal/mol·K)		Δi (phosphate ⁻¹)		Δi^b (phosphate ⁻¹)	
	U	K	U	K	U	K	U	K	U	K
2.44	37.1	31.6	51.5	42.2	166.0	138.8	0.058	0.056	0.090	0.083
8.40	42.9	38.2	55.0	44.8	174.0	143.9	0.060	0.057	0.092	0.085
33.6	49.4	46.3	57.4	47.2	178.0	147.7	0.060	0.057	0.092	0.085
100.0	55.3	53.9	53.3	49.4	162.3	151.0	0.054	0.057	0.083	0.085
300.0	61.2	60.5	55.5	49.7	166.0	149.0	0.054	0.055	0.083	0.082

^a Complete thermodynamic profiles of the unkinased (U) and kinased (K) 24-mer sequences at the indicated salt concentrations. ^b Δi was calculated by assuming that the loop phosphates behave electrostatically as single-stranded phosphates.

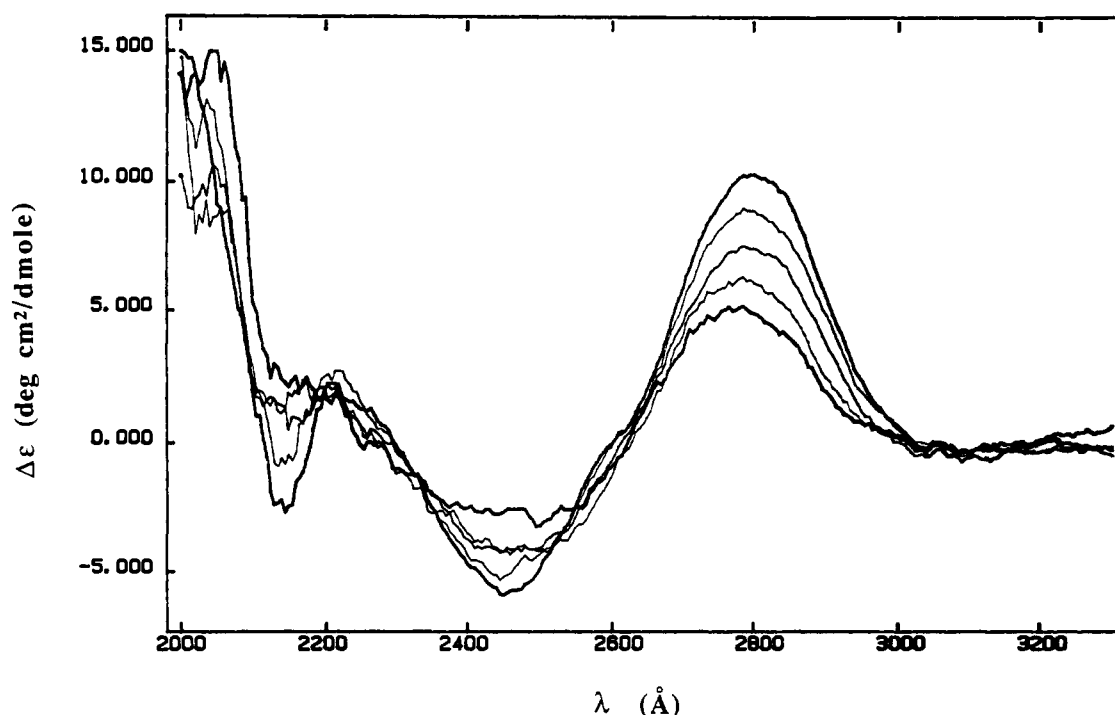


FIGURE 3: Family of temperature-dependent CD spectra for the unphosphorylated sequence at $T = 15, 40, 46, 50,$ and 80°C . The total oligonucleotide concentration is $2.5\ \mu\text{M}$ in single strands. The 15 and 80°C spectra exhibit positive bands with the smallest and largest amplitudes at $2800\ \text{\AA}$, respectively.

of strands, which is identical with the corresponding van't Hoff transition enthalpy. We will consider the implications of this equality under Discussion.

CD Spectroscopy. Figure 3 shows the family of normalized temperature-dependent CD spectra obtained for the unphosphorylated 24-mer sequence. We observed similar behavior for the phosphorylated sequence. As will be discussed later, the overall shape of the wavelength-dependent CD spectra of both the unphosphorylated and phosphorylated (not shown) molecules supports the formation of a B-like structure at low temperatures.

UV Melting Studies. Figure 4 shows the differential melting curves obtained for the unphosphorylated 24-mer at different strand concentrations. Limited quantities of the ³²P-labeled phosphorylated sequence preclude a similar concentration-dependent study. Figure 5 shows how the melting temperatures, T_m , of the unphosphorylated (unkinased) and the phosphorylated (kinased) sequences vary with sodium ion concentration. Table II lists the thermodynamic data for both the unphosphorylated and the phosphorylated sequences at all oligomer and salt concentrations. We calculated the van't Hoff transition enthalpies from the shapes of the optical transition

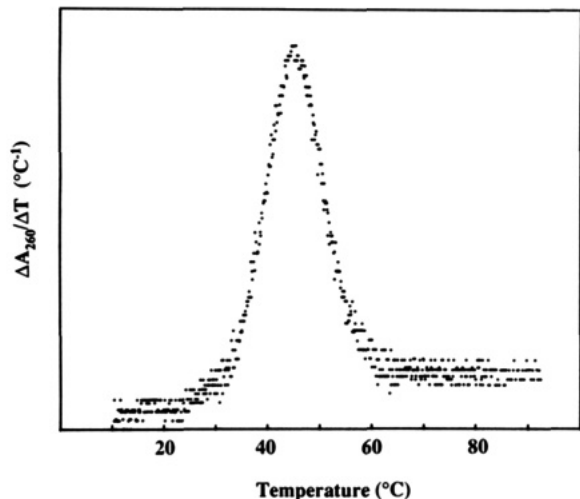


FIGURE 4: Differential melting curves for the unkinased sequence at oligomer strand concentrations of 2.56 and 15.6 μM .

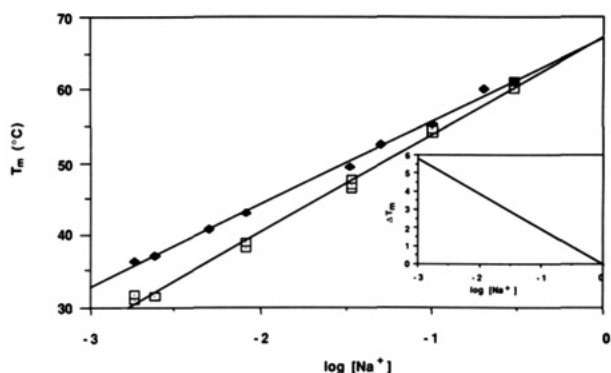


FIGURE 5: Plots of the melting temperature, T_m , versus $\log [\text{Na}^+]$ for the unkinased (\blacklozenge) and kinased (\square) forms of the 24-mer sequence. The T_m difference between the unkinased and kinased forms of the 24-mer (ΔT_m) is plotted versus the $\log [\text{Na}^+]$ in the insert.

curves. Inspection and comparison of the data listed in Table II and shown in Figures 4 and 5 lead to the following observations:

(1) For the unkinased oligomer, the T_m 's are independent of oligomer concentration (Figure 4). (Recall that the kinased oligomer was not produced in sufficient quantities to permit a corresponding concentration-dependent study.)

(2) For both oligomers, the T_m 's are a linear function of $\log [\text{Na}^+]$ (Figure 5). However, the specific values of $dT_m/d \log [\text{Na}^+]$ are different for the two oligomers.

(3) The thermal stability (T_m) of the structure formed by the 24-mer sequence is decreased by 5'-phosphorylation (Table II and Figure 5).

(4) The magnitude of the phosphorylation-induced decrease in T_m is found to decrease with increasing salt concentration. In fact, as shown in Figure 5 and emphasized by the insert, the T_m 's of the unphosphorylated and phosphorylated sequences extrapolate to equal values at 1 M $[\text{Na}^+]$.

(5) The magnitude of the change in ion dissociation, Δi , that accompanies the transition of the structure formed by the 24-mer sequence is not altered by 5'-phosphorylation.

Under Discussion, we will interpret these five observations in terms of structural changes that may result from the incorporation of a 5'-phosphate into the interior of our proposed dumbbell.

Gel Electrophoresis. Figure 6 is a photograph of the 20% polyacrylamide gel after it has been stained with ethidium bromide. Lanes 1 and 2 are the markers produced by the hexamer reference duplex $d(\text{CGCGCG})_2$ and the nine base

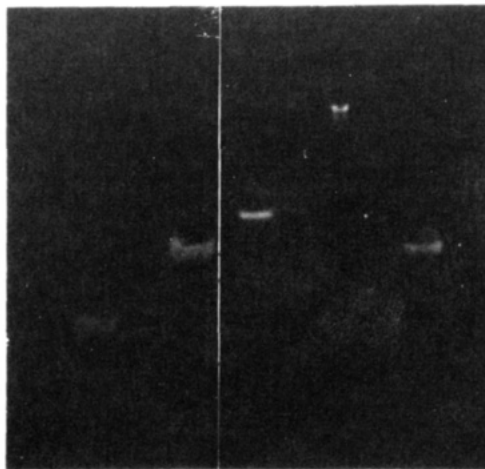


FIGURE 6: Photograph of an ethidium-stained 20% polyacrylamide gel of the kinased 24-mer sequence and marker sequences under nondenaturing conditions: (lane 1) hexamer duplex $d(\text{CGCGCG})_2$; (lane 2) 9 base pair sequence $d(\text{GGTTGTTGG})\text{-}d(\text{CCAACAACC})$; (lane 3) phosphorylated 24-mer sequence; (lane 4) 16 base pair sequence $d(\text{GGAAAAAAAAAAGG})\text{-}d(\text{CCTTTTTTTTTTTTCC})$; (lane 5) nonamer duplex $d(\text{GGTTCTTGG})\text{-}d(\text{CCAAGAACC})$. The origin of the gel is at the top of the figure. The lanes are numbered from left to right.

pair duplex $d(\text{GGTTGTTGG})\text{-}d(\text{CCAACAACC})$, respectively. Lane 3 is the phosphorylated 24-mer sequence. The phosphorylated and unphosphorylated 24-mer sequences have been found to comigrate in several previous experiments. Lanes 5 and 6 are markers produced by the 16-mer duplex $d(\text{GGAAAAAAAAAAGG})\text{-}d(\text{CCTTTTTTTTTTTTCC})$ and the nine base pair duplex $d(\text{GGTTCTTGG})\text{-}d(\text{CCAAGAACC})$, respectively. Significantly, the mobilities of both 24-mer sequences are similar to each other, slightly less than those of the nine base pair reference sequences, and significantly greater than that of the 16-mer marker duplex. We conclude that the structures assumed by the native forms of the 24-mer sequences are similar in size to a nine base pair duplex.

Ligation Reactions. The phosphorylated 24-mer was found to be resistant to ligation under all reaction conditions attempted. That is, after ligation, the phosphorylated 24-mer was still susceptible to digestion by exonuclease.

DISCUSSION

In the sections that follow, we describe how our data support the formation of a dumbbell-shaped, double-hairpin structure for both the unkinased (unphosphorylated) sequence and the kinased (phosphorylated) sequence. We also propose a structural basis for the quantitatively different melting behaviors exhibited by the 24-mer sequence before and after 5'-phosphorylation.

Evidence for Formation of Dumbbell-Shaped, Double-Hairpin Structures. The experimental data detailed above provide strong circumstantial evidence that the double-hairpin structures illustrated in Figure 1 are formed by both the unkinased and kinased 24-mer sequences. Our supporting experimental evidence is outlined below.

(1) The CD spectra shown in Figure 3 are consistent with the formation of ordered, B-like helical regions. Such ordered structures would be present in the central cores of the proposed double-hairpin structures.

(2) The temperature-dependent CD spectra in Figure 3 are characteristic of the melting of hairpin helices such as those present in the proposed dumbbell-shaped molecules. Unlike the melting of an isolated double helix where the positive CD

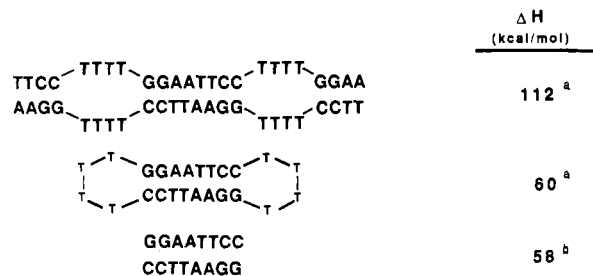


FIGURE 7: Comparison between several structures and their measured or calculated transition enthalpies, ΔH . The first two structures could be formed by our 24-mer sequence through either a bimolecular or a monomolecular process, respectively. The third structure is the isolated octamer duplex that corresponds to the core duplex in our proposed dumbbell. ^aThe transition enthalpy was calculated by using the nearest-neighbor transition enthalpy values (Breslauer et al., 1986) and assuming that the internal and hairpin loops make zero contribution to the transition enthalpy. ^bThe transition enthalpy was measured calorimetrically by using DSC.

band undergoes very little change, the positive CD bands in Figure 3 are found to change significantly with increasing temperature. Similar behavior is observed when single hairpin helices are melted (unpublished results).

(3) As illustrated in Figure 4, the melting temperature, T_m , is independent of oligomer concentration. This behavior is consistent with a monomolecular process, such as the intramolecular folding event that we believe forms the proposed dumbbell-shaped, double-hairpin structure. Alternatively, if the imperfect dimer duplex shown in Figure 7 were formed from two strands of the 24-mer sequence, a concentration-dependent equilibrium would be expected.

(4) The calorimetric enthalpy that we have measured for the melting of the ordered structure formed by the unkinased 24-mer sequence is nearly identical with the enthalpy change that we have measured for the isolated core duplex (54 kcal/mol of duplex versus 58 kcal/mol of duplex) (Patel et al., 1982; Breslauer, 1986; Breslauer et al., 1986). Although the calorimetric studies on these two systems were conducted at different salt concentrations, the comparison is justified since our van't Hoff data reveal no significant salt dependence of the transition enthalpy of the proposed dumbbell (see Table II). It is known from other studies that four thymines in a loop contribute little to the transition enthalpy of an adjacent stem duplex (Marky et al., 1985, and unpublished results). Consequently, the near identity of the ΔH values noted above is consistent with the formation of the proposed dumbbell-shaped, double-hairpin structure. If the imperfect dimer duplex shown at the top of Figure 7 were formed, a substantially larger transition enthalpy would be expected.

(5) Under non-denaturing conditions, both the unphosphorylated and phosphorylated sequences (in their native forms) are found to migrate similarly to 9 and 12 base pair duplexes in a 20% polyacrylamide gel (Figure 6). This result is consistent with the formation of the dumbbell-shaped double hairpin since the dimensions of the dumbbell structure would be similar to those of a 12 base pair duplex. If the imperfect bimolecular duplex shown in Figure 7 were formed, it would exhibit a significantly lower electrophoretic mobility, probably similar to that of a 24 base pair duplex.

The experimental evidence listed above argues that our unphosphorylated and phosphorylated 24-mer sequences intramolecularly form the dumbbell-shaped, double-hairpin structures illustrated in Figure 1. Consequently, in the sections that follow we take the liberty of interpreting our thermodynamic data in terms of these proposed structures. Ultimately, when sufficient quantities of material are synthesized, we will

use NMR to examine in detail the solution structures formed by these molecules.

The Gap in the Core Duplex of the Dumbbell Does Not Reduce Its Transition Enthalpy. In addition to providing evidence for the formation of the double-hairpin structure, the similarity of ΔH_{cal}^o for the gapped-dumbbell and the isolated octameric core duplex (54 kcal versus 58 kcal, see Table I) suggests an interesting correlation between structure and energetics. Specifically, in the isolated octameric reference duplex, covalent phosphodiester bonds connect all the nucleosides on the same strand. In contrast, in the unphosphorylated dumbbell-shaped molecule, the structure contains a *gap* between the central 5' T and 3' A residues on the "lower strand". Despite this structural difference, we measure similar transition enthalpies for these two species. In the absence of fortuitous compensations, this result suggests that the enthalpic potential of essentially all interactions (e.g., base pairing, base stacking, etc.) in a dumbbell-shaped molecule can be expressed fully even if one of the strands contains a gap.

Incorporating a 5'-Phosphate Group into the Gap Decreases the Thermal Stability of the Dumbbell Structure. As noted above, we have compared the CD and UV melting profiles of the structures formed by the unphosphorylated and phosphorylated 24-mer sequences. This comparison has allowed us to define quantitatively the consequences of 5'-phosphorylation on the thermal and salt-dependent stability of the dumbbell-shaped structures. Significantly, we observe a reduction in thermal stability (T_m) for the dumbbell-shaped structure upon 5'-phosphorylation (see Figure 5 and Table II).

The Phosphorylation-Induced Decrease in Thermal Stability Is Electrostatic in Origin. We propose that the phosphorylation-induced decrease in thermal stability for the dumbbell structure results from an increase in electrostatic repulsions in the double-hairpin molecule caused by incorporation of the 5'-phosphate group into the gap. This electrostatic interpretation agrees with the greater salt dependence of T_m that we measure for the phosphorylated structure ($\partial T_m / \partial \log [Na^+] = 13.4$) relative to that of its unphosphorylated counterpart ($\partial T_m / \partial \log [Na^+] = 11.5$). We report these data in Table I. Furthermore, the ΔT_m we observe at low salt concentrations between the phosphorylated and unphosphorylated structures approaches zero at high salt concentrations, as expected for an electrostatically induced differential stability (see insert in Figure 5). In fact, according to the insert in Figure 5, ΔT_m equals zero at 1 M sodium ion concentration. At this concentration, the equilibrium constant does not depend on Δn , the thermally induced change in the degree of ion dissociation. Equation 1 reflects mathematically this behavior, since K_{eq} is independent of Δn when $[M^+]$ is unity.

Phosphorylation-Induced Structural Changes. Clearly, an increase in electrostatic repulsions is generated by the addition of a phosphate group (5'-phosphorylation) to the 24-mer sequence. It is tempting to propose that this increase in electrostatic repulsions can be relieved by "fraying" of the bases at the central nick since they are not covalently constrained in this nicked dumbbell. A computer-generated model for one conformation of a frayed dumbbell is illustrated in Figure 8A. The model was generated by rotating about the O3'-P phosphodiester linkage between T(1) and T(2). The phosphate distances, however, are not significantly increased by such a disturbance of the 5' and 3' termini. Consequently, fraying at the nick may not be the best mechanism for relieving the phosphorylation-induced repulsions.

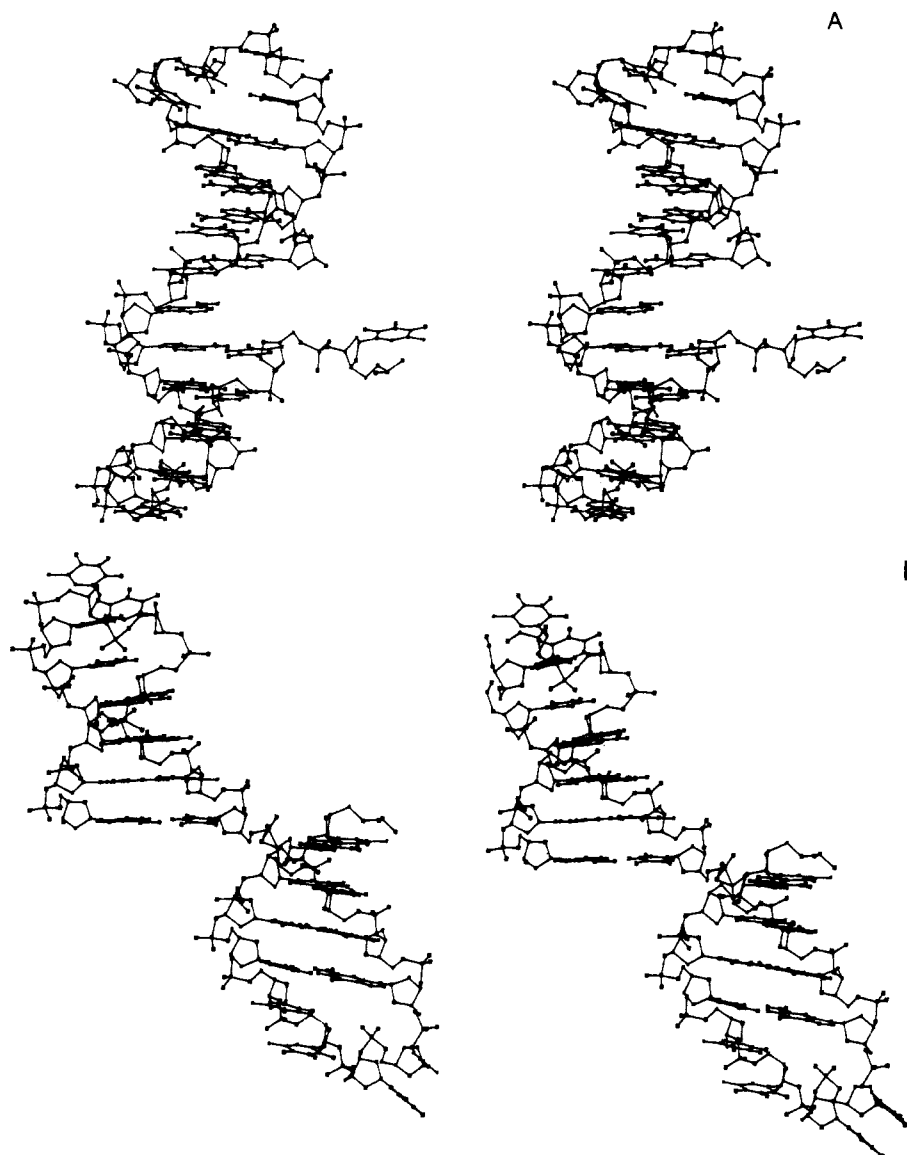


FIGURE 8: (A) Computer-generated stereo molecular model of the proposed frayed dumbbell. The model of the dumbbell was created on an Evans and Sutherland picture system using INSIGHT (Molecular Modeling System User Guide Version 1.08, BIOSYM Technologies), and the stereo images were generated by using CHEM-X (Davies, 1986). (B) Computer-generated stereo molecular model of the proposed twisted dumbbell produced with the same procedures as in (A).

Alternatively, we propose that the increased electrostatic repulsions caused by 5'-phosphorylation of our dumbbell structure can be relieved more effectively by rotation of the two hairpin halves of the dumbbell, to form a "twisted dumbbell" (Figure 8B). This structural picture differs from the more conventional frayed dumbbell structure in that the two halves of the central duplex no longer form a continuous, collinear, core duplex. Significantly, computer modeling reveals that the phosphate distances at the nick of the twisted dumbbell are considerably greater than in the frayed dumbbell, thereby reducing electrostatic repulsions. We generate this twisted dumbbell simply by rotation about the central O3'-P phosphodiester linkage between the bases A(12) and T(13) without disturbing the remaining interactions. For convenience of illustration, we have shown in Figure 8B the structure of the twisted dumbbell that results from a 90° rotation around the phosphodiester bond opposite the nick. However, the interpretations we propose for our experimental data also support any family of rotamers in which the 5' and 3' helical ends are not collinear. We explain below how the salt dependence and the relative values of the ΔH_{vH} and ΔS data listed in Tables I and II agree with the formation of an

electrostatically induced twisted dumbbell.

To discern thermodynamic differences that are electrostatic in origin, it is best to compare the ΔH_{vH} and ΔS values at low salt concentrations where screening is minimal and, thus, where electrostatic effects are most significant. According to the *low* salt data listed in Table II ($\approx 2.4\text{--}33.6$ mM in Na^+), the transition enthalpy of the phosphorylated derivative is approximately 10 kcal less endothermic than that of its unphosphorylated counterpart. The magnitude of this reduced endothermicity for the 5'-phosphorylated dumbbell-shaped structure is qualitatively consistent with the disruption of approximately one base pair stack. In the twisted dumbbell model, some, if not all, of the central AT/TA stacking interaction at the nick is disrupted. Significantly, at high salt concentrations (100 mM and above), the differences in the ΔH_{vH} transition values for the phosphorylated and unphosphorylated structures are reduced. Such a salt dependence is consistent with a twisted dumbbell at low salt concentrations that can relax to an "aligned dumbbell" at high salt concentrations where electrostatic repulsions are screened. It is recognized, however, that the enthalpic differences noted fall within the uncertainty of the van't Hoff analysis.

The relative values and the salt dependence of the ΔS data may be interpreted in a similar fashion to that used for the enthalpy data. At low salt concentrations, the ΔS for melting the phosphorylated derivative is approximately 30 cal/K·mol less than the ΔS for melting its unphosphorylated counterpart. This less favorable entropy change for melting the phosphorylated derivative is consistent with the twisted dumbbell since free rotation about the nick would yield a less ordered initial state, thereby reducing the change in entropy upon denaturation. The entropy data also are consistent with the twisted dumbbell being less hydrated, possibly due to disruption of the minor groove water spine. At high salt concentrations, the ΔS values for the two molecules are found to converge, as also is observed for the ΔH_{vH} term. This salt-dependent behavior of ΔS is consistent with our suggestion that the entropic differences are caused by the proposed electrostatically induced twisting at low salt concentrations.

We recognize that the structural interpretation of the enthalpy and entropy data presented above assumes that the high-temperature, final melted states of the phosphorylated and unphosphorylated sequences exist in similar thermodynamic states.

Orientation of the 5' and 3' Ends and Resistance to Ligation. We have shown above that 5'-phosphorylation at the gap of the dumbbell induces a destabilization that appears to be electrostatic in origin. To explain this observation on a microscopic level, we have proposed a structural picture in which the phosphorylation-induced increase in electrostatic repulsions induces the formation of a twisted dumbbell at low salt concentrations. We have shown how this structural interpretation is consistent with our T_m , ΔH , and ΔS data as well as their salt dependencies. However, we have yet to explain one additional observation: namely, all our efforts to close the 5'-3' nick in the 24-mer dumbbell using T4 DNA ligase have failed. At first, we found this observation surprising since at least one of the rotamers of the twisted dumbbell appeared to have the 5' and 3' ends juxtaposed in a manner that would allow ligation. Maybe we have yet to identify the appropriate conditions for ligation, if ligation is possible. However, we have attempted to induce ligation over a broad range of temperatures and solution conditions. Thus, assuming that the observed resistance to ligation is an inherent property of this 24-mer molecule, we describe below a possible structural explanation.

Shore and Baldwin have shown that the ligation reaction which forms closed, circular, double-stranded DNA strongly depends on the DNA chain length (Shore & Baldwin, 1983). They explain this observation by proposing that successful intramolecular ligation requires the proper phasing of the 5' and 3' ends. (Proper phasing normally does not influence bimolecular ligation reactions since they generally are not subject to the topological constraints of intramolecular ligation reactions.) We suggest that the resistance to ligation of the nicked 24-mer dumbbell structure studied here reflects improper phasing between the two terminal residues at the nick. One can generate the nicked 24-mer dumbbell structure conceptually by a two-step process—folding two hairpins of identical sequence, d(pTTCCCTTTGGAA), and then forming a phosphodiester bond between the 5'P of one hairpin and the 3'OH of the other. These two steps can produce a nicked-dumbbell structure in which the 3' and 5' ends are out of phase if the terminal loops distort the core duplex. Specifically, a loop of only four residues may constrain the conformation of each hairpin half (Haasnoot et al., 1986), thereby preventing the formation of a regular B-helix. However, a larger loop

may provide enough flexibility to achieve proper spatial orientation (phasing) at the nick. If the terminal loops indeed are distorting the helical region of this molecule, then proper phasing of the terminal residues may in fact be crucial not only for *double-stranded* intramolecular ligation reactions (Shore & Baldwin, 1983) but also for topologically constrained, intramolecular ligation reactions that result in the formation of covalently closed, *single-stranded* circles. To investigate this possibility, we are performing ligation experiments on the corresponding 26-mer dumbbell structure that contains the same core duplex but has five rather than four thymine residues in the loop. We recognize that elongation of the core duplex also may alleviate this phasing problem.

The interpretations presented above provide a structural basis for the salt dependence and the relative values of the thermodynamic data we have obtained for the phosphorylated and unphosphorylated 24-mer sequences as well as the resistance to ligation of the nicked dumbbell. Additional structural studies should allow us to test the validity of our microscopic interpretations of the macroscopic data.

The Thymine Residues in the Loop Electrostatically and Enthalpically Behave as Denatured Single Strands. Inspection of the data listed in Table I reveals that our phosphorylated and unphosphorylated dumbbell-shaped molecules exhibit $\partial T_m / \partial \log [\text{Na}^+]$ values (11.5 and 13.4) similar to the value of 12.3 that we measure for the isolated octameric duplex that corresponds to the central core. This similarity suggests that the thymine residues in the loop do not contribute significantly to the electrostatic character of the duplex state. However, inspection of eq 2 reveals that $\partial T_m / \partial \log [\text{Na}^+]$ depends not only on the denaturation-induced change in the number of thermodynamically bound counterions but also on the ΔH° of the transition. Thus, to avoid erroneous conclusions caused by differences in ΔH° , one should compare the values of Δi rather than $\partial T_m / \partial \log [\text{Na}^+]$. We provide this comparison of Δi data in the last two columns of Tables I and II.

The Δi data listed in the next-to-last columns of Tables I and II were calculated by assuming the thymine residues in the loop to behave electrostatically as if they were in the duplex state. This assumption is consistent with the treatment of Record and Lohman (1978). On the basis of this assumption (which implies that the native dumbbell structure will behave electrostatically like a dodecamer duplex), a Δi value of 0.057 is calculated. Surprisingly, this value is only half the magnitude of that calculated for a dodecamer duplex based on salt-dependent melting studies (Elson et al., 1970; Baldwin, 1971; unpublished data). This disparity is an indication that the above assumption may not be valid; that is, the phosphates in the loop may not be behaving electrostatically as if they were in the duplex state.

Alternatively, it can be assumed that the phosphates in the loop behave electrostatically as denatured coils rather than native duplexes. On the basis of this assumption, the initial state would be better modeled by an octamer rather than a dodecamer duplex. Employing this assumption, we calculate the Δi values listed in the final columns of Tables I and II. The suggestion that the double-hairpin structures behave electrostatically like octamer duplexes is in fact supported by the close agreement of the Δi values for the dumbbell structures with the corresponding value for the octameric core duplex. In other words, the phosphates in the loops of the native dumbbell structures electrostatically behave as if they were in denatured coil states. This conclusion is contrary to assumptions made in the original work on oligoelectrolyte theory (Record & Lohman, 1978) when few experimental data

were available. Future work will be required to evaluate whether this inconsistency is unique to our dumbbell or if it is a general feature of looped regions.

Our results may be a reflection of greater phosphate distances in the loop compared with the duplex regions. This electrostatic behavior is consistent with our calorimetric data. It should be recalled that similar transition enthalpies for the unphosphorylated dumbbell and the isolated octameric core duplex were measured. In the absence of fortuitous compensations, this similarity in transition enthalpies is indicative of the thymines in the loops of the initial dumbbell state being enthalpically similar to the thymines in the final coiled state. In other words, the bases in the loop are thermodynamically unstacked.

In summary, the salt dependencies of our dumbbell-shaped molecules in conjunction with their transition enthalpies suggest that the secondary structure of the hairpin loops more closely resembles the structure of the coil state rather than the structure of the double-helical state. We now are using computer modeling in conjunction with energy minimization and Monte Carlo calculations to evaluate the structural implications of these results. In the sections that follow, we interpret our enthalpic and electrostatic data in the context of available NMR structural data.

Loop Enthalpy and Loop Structure. On the basis of our enthalpy data, it is tempting to predict that the thymine residues are unstacked in both loops of each dumbbell structure. To evaluate this structural prediction, NMR data on dumbbell-shaped molecules similar to ours would be useful. However, until these experiments are performed, NMR data are available only on hairpin structures possessing a single loop of four thymine residues (Hare & Reid, 1986; Haasnoot et al., 1986; Hilbers et al., 1985; Summers et al., 1985; Ikuta et al., 1986). If the unpaired regions of our double-hairpin structures are assumed to be structurally similar to the unpaired region present in single-hairpin structures, then the available NMR data can be used to evaluate our prediction that the thymine residues in both loops of our dumbbell structures are unstacked. Obviously, the conclusions reached in the discussion that follows are dependent on the validity of this assumption.

NMR studies on single-hairpin loops reveal some degree of stacking between four thymine residues in a loop (Hare & Reid, 1986; Haasnoot et al., 1986; Hilbers et al., 1985). However, our enthalpy data suggest that such loop interactions do not significantly contribute to the transition enthalpy measured for the double-hairpin structure. [Recall, we measure similar transition enthalpies for the unphosphorylated dumbbell structure and its isolated core duplex (see Table I).] Taken together, these two results suggest that structural stacking of thymine residues in a loop does not necessarily result in a favorable enthalpy contribution. This conclusion may simply reflect the fact that the stacking of thymine residues is an intrinsically low-enthalpy event. Alternatively, our results may reflect an unusual effect of loop conformation on base stacking. To distinguish between these two possibilities, we will be investigating double-hairpin molecules with other sequences in the loops.

Loop Electrostatics and Loop Structure. The available proton NMR data provide information only on the distances and relative positions of the bases rather than the phosphate groups in the loop. Our salt-dependent data, however, reveal that the loop phosphates electrostatically behave as single strands rather than as duplex structures. Therefore, we conclude that a loop of four thymine residues has a charge density

similar to that of a single strand. Significantly, we have reached this conclusion in the absence of structural data concerning phosphate distances in the loop. Once structural data become available on dumbbells, we can combine it with our enthalpic and electrostatic results to produce a more complete picture of hairpin loop structures.

Comparison with the Results of Wemmer and Benight. As part of an independent effort, Wemmer and Benight have reported an optical and NMR study on a dumbbell-shaped molecule (Wemmer & Benight, 1985). However, their study and ours differ in several respects. They produced their dumbbell by ligating two hairpin structures. This approach requires synthesizing two different sequences, each which forms hairpin structures with sticky ends. In contrast, our approach requires synthesizing only a single sequence that is designed to fold intramolecularly into a dumbbell-shaped structure with a gap between the 5' and 3' ends. Thus, we also can study the influence of a structural gap within a dumbbell and the consequences of inserting a phosphate group into the gap prior to ligation. Furthermore, in the work of Wemmer and Benight, the melting of the isolated core duplex is complicated by competing monomolecular (hairpin) and bimolecular (duplex) transitions. These complicating effects are due, in part, to the fact that the core duplex in their investigation is twice as large as the core duplex studied in this work. In contrast, the thermodynamic reference state for the core duplex of our system melts in a monophasic, two-state manner, thereby simplifying the thermodynamic analysis (Patel et al., 1982; Breslauer, 1986; Breslauer et al., 1986). In fact, the melting of the isolated core duplex of our dumbbell structure has been studied previously by both calorimetric (Patel et al., 1982; Breslauer, 1986) and NMR techniques (Patel & Canuel, 1979; Patel et al., 1982; Connolly & Eckstein, 1984). As a consequence of these previous studies, we are able to isolate and to evaluate the influence of the two terminal looped regions by factoring out the influence of the core duplex.

The Wemmer and Benight study and the work reported here produced different information since each investigation employed substantially different experimental methods. Ultimately, we hope to combine the NMR approach of Wemmer and Benight with the calorimetric and spectroscopic techniques reported here to investigate a common dumbbell structure. Such a coordinated effort should yield important correlations between macroscopic thermodynamic data and microscopic structural data for this intriguing new class of nucleic acid secondary structures.

ADDED IN PROOF

We have successfully ligated the corresponding 26-mer dumbbell with five T residues in each loop.

ACKNOWLEDGMENTS

We thank Dr. Luis A. Marky for his assistance with the calorimetry experiments.

REFERENCES

- Baldwin, R. L. (1971) *Acc. Chem. Res.* 4, 265-272.
- Betz, J. L., & Sadler, J. R. (1981) *Gene* 13, 1-12.
- Bloomfield, V. A., Crothers, D. M., & Tinoco, I., Jr. (1974) *Physical Chemistry of Nucleic Acids*, Chapters 6 and 7, pp 293-476, Harper & Row, New York.
- Breslauer, K. J. (1986) in *Thermodynamic Data for Biochemistry and Biotechnology* (Hinz, H.-J., Ed.) Chapter 15, pp 402-427, Springer-Verlag, New York.
- Breslauer, K. J., Sturtevant, J. M., & Tinoco, I., Jr. (1975) *J. Mol. Biol.* 99, 549-565.

- Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 3746-3750.
- Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Vol. III, Chapters 15, 16, and 22-24, pp 849-938 and 1109-1326, Freeman, San Francisco.
- Cantor, C. R., Warshaw, M. M., & Shapiro, H. (1970) *Biopolymers* 9, 1059-1077.
- Carruthers, M. H., Beaucage, S. L., Becker, C., Efcavitch, W., Fisher, E. F., Gallupi, G., Goldman, R., deHaseth, P., Martin, P., Matteucci, M., & Stabinsky, Y. (1982) in *Genetic Engineering* (Stelow, J., & Hollaender, A., Eds.) Vol. 4, pp 1-18, Plenum, New York.
- Connolly, B. A., & Eckstein, F. (1984) *Biochemistry* 23, 5523-5527.
- Davies, E. K. (1986) CHEM-X Program Suite, Chemical Design Ltd., Oxford, England.
- Elson, E. L., Scheffler, I. E., & Baldwin, R. L. (1970) *J. Mol. Biol.* 54, 401-415.
- Gaffney, B. L., Marky, L. A., & Jones, R. A. (1984) *Biochemistry* 23, 5686-5691.
- Haasnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Boom, J. H., Singh, U. C., Pattabiraman, N., & Kollman, P. A. (1986) *J. Biomol. Struct. Dyn.* 3, 843-857.
- Hare, D. R., & Reid, B. R. (1986) *Biochemistry* 25, 5341-5350.
- Hilbers, C. W., Haasnoot, C. A. G., de Bruin, S. H., Joordens, J. J. M., van der Marel, G. A., & van Boom, J. H. (1985) *Biochimie* 67, 685-695.
- Ikuta, S., Rajagopal, C., Hirataka, I., Dickerson, R. E., & Kearns, D. R. (1986) *Biochemistry* 25, 4840-4849.
- Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers* 26, 1601-1620.
- Marky, L. A., Canuel, L., Jones, R. A., & Breslauer, K. J. (1981) *Biophys. Chem.* 13, 141-149.
- Marky, L. A., Blumenfeld, K. S., Kozlowski, S., & Breslauer, K. J. (1983) *Biopolymers* 22, 1247-1257.
- Marky, L. A., Kallenbach, N. R., Seeman, N. C., Haasnoot, C. A. G., & Breslauer, K. J. (1985) *Book of Abstracts, Fourth Conversation in Biomolecular Stereodynamics, June 4-8, 1985* (Sarma, R. H., Ed.) p 74, Institute of Biomolecular Stereodynamics, Albany, NY.
- Patel, D. J., & Canuel, L. L. (1984) *Eur. J. Biochem.* 96, 267-276.
- Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry* 21, 451-455.
- Pheiffer, B. H., & Zimmerman, S. B. (1983) *Nucleic Acids Res.* 11, 7853-7871.
- Record, M. T., Jr., & Lohman, T. M. (1978) *Biopolymers* 17, 159-166.
- Record M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys.* 11, 103-178.
- Scheffler, I. E., Elson, E. L., & Baldwin, R. L. (1970) *J. Mol. Biol.* 48, 145-171.
- Shore, D., & Baldwin, R. L. (1983) *J. Mol. Biol.* 170, 957-981.
- Summers, M. F., Byrd, R. A., Gallo, K. A., Samson, C. J., Zon, G., & Egan, W. (1985) *Nucleic Acids Res.* 13, 6375-6386.
- Wemmer, D. E., & Benight, A. S. (1985) *Nucleic Acids Res.* 13, 8611-8620.
- Zimmerman, S. B., & Pheiffer, B. H. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 5858-5856.

Poly(2-amino-8-methyldeoxyadenylic acid): Contrasting Effects in Deoxy- and Ribopolynucleotides of 2-Amino and 8-Methyl Substituents

Eiko Nakagawa Kanaya,[†] Frank B. Howard, Joe Frazier, and H. Todd Miles*

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Received March 17, 1987; Revised Manuscript Received June 18, 1987

ABSTRACT: Poly(2-amino-8-methyldeoxyadenylic acid) interacts readily with pyrimidine polynucleotides to form double helices only slightly less stable than those in which the purine polymer lacks the 8-Me group. In the ribo series, by contrast, complexes formed with poly(2-amino-8-methyladenylic acid) are very strongly destabilized by the 8-Me group, despite a larger stabilizing effect of the 2-NH₂ group in the ribo series. These results are interpreted in terms of a smaller steric interference of the 8-Me group with 2'-CH₂ than with 2'-CHOH, leading to a smaller population of syn structures in the deoxy chain and a consequent lower interference with homopolymer duplex formation. UV, circular dichroism (CD), and IR spectra of the new polymer and its complexes are reported and related to structural and energetic characteristics of the molecules. Since direct synthesis of 2-amino-8-methyldeoxyadenosine was not feasible, the corresponding riboside was prepared, the 3'- and 5'-positions were protected with a disilyloxy group, and a 2'-[(imidazol-1-yl)thiocarbonyl] group was introduced. Reduction with tributyltin hydride followed by deprotection gave the nucleoside, which was then converted to the triphosphate by standard methods. The homopolymer was prepared with terminal deoxynucleotidyl transferase.

SSecondary and tertiary structures of nucleic acids are determined by a number of energetic and conformational factors. These are frequently not well resolved, and predictions are

often difficult and unreliable. In order to help separate and analyze some of these factors, we have employed chemical modifications of nucleotides and polynucleotides to perturb their properties in defined and controllable ways. Thus, substitution of 2-NH₂ in poly(A) permits three interbase hydrogen bonds to be formed to U and T polymers. In the ribo

[†]Present address: Mitsubishi-Kasei Institute of Life Sciences, 11 Minamiooya, Machida, Tokyo 194, Japan.