Volume 13 Number 22 1985 Nucleic Acids Research

Anomalous hairpin formation in an oligodeoxyribonucleotide

J.G.Nadeau and P.T.Gilham

Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

Received 8 July 1985; Revised and Accepted 21 October 1985

ABSTRACT

An accurate method for deriving molar absorptivity-temperature profiles applied to a set of single-stranded oligodeoxyribonucleotides shows that the undecamer CGAGTTTGACGp exists in a hairpin conformation involving Watson-Crick base pairing between the two terminal CG dinucleotides. The hairpin, which has a transition midpoint of 40 °C in 0.115 M Na^+, is unusually stable in comparison with previously reported hairpins. A non-linear least squares analysis of the undecamer's profile in terms of a two-state equilibrium model indicates that the hairpin-to-coil transition occurs with an enthalpy change about twice that expected if only combinations of Watson-Crick base-paired stacking interactions are considered. The analogous hairpin structure (containing an identical CG/CG stem) assignable to the complementary strand CGTCAAACTCGp does not form above 0 °C. Measurements on the two undecamers indicate that variation in non Watson-Crick interactions within the loops of two similar hairpins can produce a difference in stability of at least 2.2 kcal/mol (25 °C, 0.115 M Na^+), roughly equal to the amount contributed to a double helix by a 5'-CG-3'/5'-CG-3' base-paired stacking interaction.

INTRODUCTION

The nucleic acid hairpin structure, consisting of a base-paired double-helical stem that is closed at one of its ends by a loop of contiguous unpaired nucleotide residues, constitutes a major structural element in ordered forms of single-stranded RNA and DNA. In addition, hairpins serve as structures with particular biological significance, such as the anticodon regions in tRNAs and as components of mechanisms for the control of gene expression at the levels of both transcription and translation. A complete set of rules governing the formation and stability of hairpins has yet to be defined, although several studies on oligonucleotides (1-5) have demonstrated that loop size and stem length and sequence are key factors in determining stability. One aspect of hairpin structure that has received little attention is the influence of loop sequence on stability, even though some reports on RNA hairpins have indicated that this effect is probably significant (6,7). During a preliminary study of duplexes formed from several pairs of comple-

© IRL Press Limited, Oxford, England.

mentary oligodeoxyribonucleotides, we noticed that one of the single strands, CGAGTTTGACGp, displayed an unusual absorptivity-temperature profile. Further investigation has led to the conclusion that this oligonucleotide adopts a hairpin conformation that is surprisingly stable in view of the molecule's limited capacity for forming intramolecular W-C base pairs.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification. Oligonucleotides were prepared in 10-20 mg quantities by the solution phase phosphotriester strategy with block condensation procedures previously reported (8,9). The fully protected oligomers were purified by silica gel column chromatography and the blocking groups were removed by successive treatments with tetramethylguanidinium pyridine-2-aldoximate, ammonia, and acetic acid as described (9). The oligonucleotides were separated from salts and the various blocking groups by chromatography on a column (1 X 75 cm) of Sephadex G-10 gel filtration beads with EtOH-H_2O (1:4, v/v) as eluting solvent. Each oligomer was then purified in 5 mg portions by ion-exchange HPLC on a column (1 X 25 cm) of microparticulate silica coated with cross-linked polyethyleneimine (10) with various linear gradients of $(NH_4)_2SO_4$ in aqueous 0.05 M KH_2PO_4 (pH 6.0) containing MeCN (15%, v/v). Appropriate fractions were combined, concentrated and dialyzed against water to remove the bulk of the salt. Remaining quantities of salt were then removed with a Sephadex G-10 column (2.5 X 60 cm) using the aqueous ethanol solvent system described above. The amount of salt in each oligomer preparation at this stage was shown to be negligible by Nessler assay for NH_4^+ (11). Oligonucleotides were converted to their sodium forms with a short column of Dowex 50W-X2(Na^+) ion-exchange resin. A small quantity of each molecule prepared in this way was analyzed for purity by HPLC; in no case did the high resolution system show UV-absorbing impurities of more than 1%.

Molar Absorptivities. The method requires the treatment of the oligomer (ca. 1 A_{260} unit) in 1 ml of 0.01 M 2-(N-morpholino)ethanesulfonate (MES) buffer, pH 6.5, with spleen phosphodiesterase (0.2 unit; Pharmacia, Milwaukee, WI) at 25 °C. The absorbance at 260 nm is monitored until it is constant, and the hyperchromicity is calculated by dividing the final A_{260} by the initial value after making allowance for changes in volume and absorbance caused by the addition of the enzyme. At this point, a small quantity of the hydrolysate is analyzed by the HPLC method described above to ensure that complete conversion to monomers has taken place. Some oligomers (other than those used in this work) require higher temperatures to undergo complete degradation; in

these cases, however, the initial and final absorbance measurements are also made at 25 °C. The molar absorptivity is obtained by dividing the hyperchromicity into the sum of the ε_{260} values of the nucleotides corresponding to all the positions in the chain. For oligomers possessing 3' terminal phosphate groups, the ε values for the 3' nucleotides are used; the same values are also employed for a chain with a free 3' hydroxyl group except for the 3' terminal position where the ε value of the corresponding nucleoside is substituted. Some batches of the commercial enzyme contain a contaminating deoxyadenosine deaminase activity, and, if used in the analysis of oligonucleotides terminating in a dA residue that contains a free 3' hydroxyl group, an allowance for the dA to dI absorbancy change must be made in the calculation of the hyperchromicity. For the monomer absorptivities, all four deoxyribonucleosides were first purified by multiple recrystallizations from water or ethanol, and then measured in MES buffer; the ε_{260} values (M^{-1} cm^{-1}) are averages of 3 determinations (± standard deviation): dA, 14,970 ± 200; dC, 7,350 ± 100; dG, 11,890 ± 100; dT, 8,860 ± 200. The corresponding value for each 3'-nucleotide was obtained by observing the change in A_{260} resulting from its conversion in MES buffer to the parent nucleoside with alkaline phosphatase. Averages of 3 determinations gave: dAp, 15,220 ± 200; dCp, 7,500 ± 100; dGp, 11,890 ± 100; dTp, 9,090 ± 200. To minimize round-off errors in the calculation of oligonucleotide absorptivities, the number of significant figures included in each monomer value exceeds that dictated by the standard deviation. Molar absorptivities of oligonucleotides in the standard Tm buffer (0.1 M NaCl-10 mM sodium phosphate-0.1 mM EDTA, pH 7.0) were obtained from the ε_{260} values determined in MES buffer, by comparing the absorbances of equal amounts of each oligomer in the two buffer systems. The difference in absorptivity in the MES and Tm buffers was less than 1% in each case.

Base Composition and Sequence Analysis. The solution of hydrolyzed oligomer resulting from an absorptivity determination is treated with 20 µl of conc. NH_4OH and applied to a column (20 X 0.4 cm) of Dowex 1 ion-exchange resin (AG 1-X2, -400 mesh, Bio-Rad Laboratories). Separation is effected with 200 ml of 10% EtOH in water (v/v) containing a linear gradient of 0.05 M NaCl-0.005 M NH_4OH to 0.5 M NaCl-0.05 M NH_4OH at 20 ml/h. Elution volumes (ml) in this system are: dC, 9; dA, 12; dT, 13; dG, 56; dCp, 54; dAp, 72; dTp, 87; dGp, 106. Quantitation is carried out by monitoring the UV absorbance of the eluate. The sequence analysis method used is a new procedure employing a simple two-dimensional separation on a single polyethyleneimine cellulose thin layer sheet (12).

Absorbance-Temperature Profiles. Absorbance measurements were carried out on solutions of oligomers in the standard Tm buffer (defined above), the low salt Tm buffer (10 mM sodium phosphate-0.1 mM EDTA, pH 7.0), or the high salt Tm buffer (1 M sodium chloride-10 mM sodium phosphate-0.1 mM EDTA, pH 7.0). A Perkin-Elmer Lambda-5 spectrophotometer equipped with a thermoelectric cell holder and digital temperature controller and programmer was used to make readings over the range 0-85 °C. For each profile, the temperature was raised at a rate of not more than 0.5°/min, and the absorbance (260 nm) and temperature were recorded every minute; each transition's reversibility was confirmed by measurements during return to the starting temperature. For each oligonucleotide, at least one of these cooled samples was analyzed by HPLC to ensure that no structural damage had resulted from exposure to high temperature. Absorbance values were corrected for volume expansion prior to data analysis. The solution used for the measurement of each duplex profile was prepared by mixing the two appropriate oligomers dissolved in the standard Tm buffer to give equimolar proportions of the strands, based on the calculated molar absorptivities, with a final total strand concentration of 10-15 μM.

RESULTS

Oligonucleotides. The oligodeoxyribonucleotides examined in this study are listed in Table I, along with their hyperchromicities and molar absorptivities. Each of these oligomers gave the expected nucleotide ratios upon base composition analysis; in the case of X-11 and Y-11, the structures were also confirmed by sequence analysis. Each hyperchromicity value listed is the average (± standard deviation) of at least three measurements. The standard deviations for the absorptivities contain contributions from uncertainties in both the oligonucleotide hyperchromicities and mononucleotide absorptivities.

TABLE I: Oligodeoxyribonucleotides

Number	Sequence	Hyperchromicity[a]	ε_{260} at 25 °C[b]
X-11	CGAGTTTGACGp	1.307 ± .004	91,200 ± 700
X-11a	CGAGT	1.264 ± .002	43,800 ± 300
X-11b	TTGACGp	1.190 ± .001	54,500 ± 400
X-9	GAGTTTGACp	1.246 ± .002	80,900 ± 400
X-9a	GAGTp	1.266 ± .001	38,000 ± 300
X-9b	TTGACp	1.182 ± .001	44,800 ± 400
Y-11	CGTCAAACTCGp	1.267 ± .002	93,100 ± 800
Y-9	GTCAAACTCp	1.268 ± .001	77,600 ± 700

[a]Measured at 25°C with oligomers dissolved in 0.01 M MES buffer, pH 6.5.
[b]Absorptivity values for oligomers dissolved in 0.1 M NaCl-10 mM sodium phosphate-0.1 mM EDTA, pH 7.0.

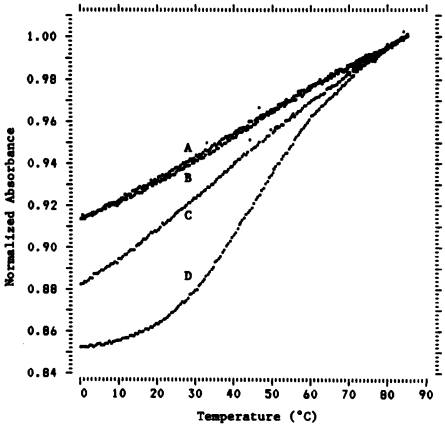

FIGURE 1: Normalized 260 nm absorbance-temperature plots for Y-9 (A), Y-11 (B), X-9 (C), and X-11 (D). Readings were carried out on samples dissolved in 0.1 M NaCl-10 mM sodium phosphate-0.1 mM EDTA, pH 7.0, with oligomer concentrations at 10 μM.

Absorbance-Temperature Profiles. For each oligomer at least four A_{260}-T profiles were collected over a twenty-fold concentration range of about 4 to 80 μM. The normalized melting curves of X-9, Y-9, X-11, and Y-11 are shown in Figure 1; each curve was obtained from the A_{260}-T profile of a 10 μM oligonucleotide sample by dividing the absorbance at every temperature by the value at 85 °C. Melting curves derived from other concentrations and normalized in this way were indistinguishable from the corresponding 10 μM plots within the limits of experimental error. The absorptivity-temperature curves for X-9 and X-11 shown in Figure 2 were also derived from the 10 μM plots.

Composite Absorptivity-Temperature Curves. The calculated ε_{260}-temperature plots for X-9 and X-11 shown in Figure 2 were obtained from the corresponding curves for their respective subfragments, X-9a, X-9b and X-11a, X-11b with

$$\varepsilon(T) = \varepsilon_a(T) + \varepsilon_b(T) + I(T) \qquad (1)$$

where $\varepsilon(T)$ is the value of the calculated molar absorptivity at T °C; $\varepsilon_a(T)$ and $\varepsilon_b(T)$ are the corresponding values obtained from the subfragments and the term I(T) accounts for the disruption of interactions (associated with the internal T-T junction) that would result from converting X-9 or X-11 into the respective subfragments. Approximate values for I(T) were estimated from absorptivity-temperature measurements on oligomer T-T_{10}-T (T-12) and thymidine 3'-phosphate (T-1):

$$I(T) = [\varepsilon_{(T-12)}(T) - 12 \cdot \varepsilon_{(T-1)}(T)]/11 \qquad (2)$$

I(T) was almost linear with values of -800 and -500 M^{-1} cm^{-1} at 20° and 80°C, respectively. Equations 1 and 2 follow from a semi-empirical approach first used by Cantor and Tinoco (13) to calculate the optical properties of

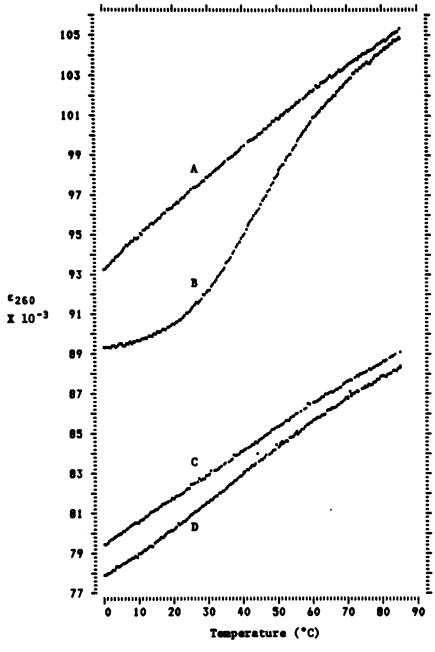

FIGURE 2: Absorptivity-temperature profiles obtained from oligomers dissolved in 0.1 M NaCl-10 mM sodium phosphate-0.1 mM EDTA, pH 7.0 A, Composite curve from X-11a + X-11b + internal T-T stack; B, X-11; C, Composite curve from X-9a + X-9b + internal T-T stack; D, X-9.

trinucleotides from those of dinucleotides.

Effect of Ionic Strength on Melting Curves. The effect of ionic strength on the absorbance-temperature profiles of X-11 and Y-11 is shown in Figure 3. At the lower ionic strength the X-11 curve shows a substantial shift to lower temperatures while the same change in sodium ion concentration has little effect on the Y-11 profile. The curves for Y-9 and X-9 (not shown) are simil-

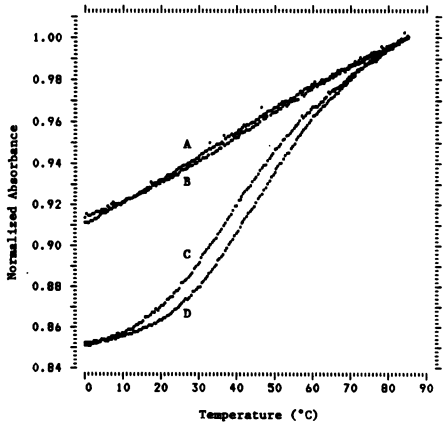

FIGURE 3: Effect of ionic strength on normalized 260 nm absorbance-temperature plots for Y-11 and X-11. Plots A (Y-11) and C (X-11) were derived from samples dissolved in 10 mM sodium phosphate-0.1 mM EDTA, pH 7.0. Plots B (Y-11) and D (X-11) were derived from samples dissolved in 0.1 M NaCl-10 mM sodium phosphate-0.1 mM EDTA, pH 7.0.

ar to the latter in exhibiting insensitivity to changes in ionic strength.

Analysis of X-11's Structural Transition. In order to estimate the standard enthalpy and entropy changes associated with the cooperative transition in X-11, the ε_{260}-T profiles were analyzed in terms of a two-state (hairpin-coil) equilibrium. The fraction (α) of molecules in the hairpin state was obtained from:

$$\alpha = [\varepsilon_c - \varepsilon(T)]/[\varepsilon_c - \varepsilon_h]$$

where $\varepsilon(T)$ is the measured absorptivity at temperature T; ε_c and ε_h are the corresponding values for coil and hairpin, respectively. Estimates for α were used to derive thermodynamic parameters by means of the standard equilibrium expression for a two-state transition:

$$K_h = \frac{\alpha}{(1-\alpha)} = \exp\left[\frac{-\Delta H^\circ}{RT} + \frac{\Delta S^\circ}{R}\right] \quad (3)$$

In general, ε_c and possibly ε_h are expected to vary with temperature in response to non-cooperative structural changes in the coil and hairpin states, respectively. The availability of measurements carried out on subfragments X-11a and X-11b (whose local conformations and hence optical properties should be similar to those of the X-11 coil) permitted the temperature dependence of ε_c to be estimated:

$$\varepsilon_c(T) = \varepsilon_a(T) + \varepsilon_b(T) + m_c T + n_c \quad (4)$$

where $\varepsilon_a(T)$ and $\varepsilon_b(T)$ are the molar absorptivities of X-11a and X-11b, respectively, at T °C. The linear term $m_c T + n_c$ was included to approximate interactions present in the X-11 coil that are not represented in the subfragments. The absorptivity of the hairpin was assumed to be a linear function of temperature: $\varepsilon_h(T) = m_h T + n_h$. The non-linear least squares algorithm of Marquart (14) was used to fit eq 3 to the measured profiles by treating ΔH°, ΔS°, m_c, m_h, n_c, and n_h as adjustable parameters.

The results of this analysis carried out on melting curves of X-11 samples in solutions at various ionic strengths are shown in Table II. For the two-state equilibrium, Tm is defined as the temperature at which $\alpha = 0.5$ and is given by $Tm = \Delta H^\circ/\Delta S^\circ$. The values for samples in 0.115 M Na^+ are averages obtained from five separate melting curves representing a 20-fold range of oligonucleotide concentration; there was no apparent systematic variation in the values of these parameters with concentration. Each of the values given for the other ionic strengths is an average derived from curves of three samples measured at oligonucleotide concentrations of about 10 μM.

Excellent agreement was obtained between experimental curves and the two-state fits to the data when ε_c and ε_h were allowed to change with temperature

TABLE II: Thermodynamic Parameters for the X-11 Hairpin

[Na^+]	$\Delta H°$ (kcal/mol)[a]	$\Delta S°$ (e.u.)	Tm (°C)
0.015 M	-18.6 ± 1	-60.6 ± 3	34.2 ± 0.9
0.115 M	-19.2 ± 1	-61.3 ± 4	40.3 ± 0.6
1.015 M	-17.5 ± 1	-55.4 ± 2	43.5 ± 0.7

[a]The uncertainties in the values for $\Delta H°$ and $\Delta S°$ are correlated because of the theoretical relationship between the two parameters, and this is responsible for the relatively high precision of Tm. The free energy change should be comparatively free of random errors for the same reason and significant figures in excess of those warranted by the standard deviations are given for $\Delta H°$ and $\Delta S°$ to improve the accuracy of $\Delta G°$ calculations.

as described above. A typical melting profile (after baseline subtraction) and its theoretical fit are shown in Figure 4. The contributions of the linear term in the expression for ε_c (eq 4) were found to be quite small and roughly equivalent to the values for I(T) calculated from T-T_{10}-T and thymidine 3'-phosphate (eq 2). The average slope of the lower baselines was 33 M^{-1} cm^{-1} $°K^{-1}$. The fits were significantly less satisfactory when the baselines were not allowed to vary with temperature. Although use of a temperature-dependent upper baseline is justified by the behavior of subfragments X-11a and X-11b (Curve A, Figure 2), the linear approximation for the lower baseline is less easily rationalized, despite the fact that it does yield slightly better fits. It is therefore important to note that the magnitudes of $\Delta H°$ and $\Delta S°$ decreased by 9 and 8 %, respectively, when ε_c was allowed to vary with temperature (in the above manner) but ε_h was not.

In view of earlier studies on oligonucleotide duplex melting (15-17), it is likely that changes in ε_c and ε_h with temperature reflect the deviation of the actual hairpin-coil transition from two-state behavior. In general, the two-state model will underestimate the total enthalpy change of a multi-state transition. However, a comparison of the enthalpy changes determined from spectroscopic and calorimetric data for the duplex-coil transition of the

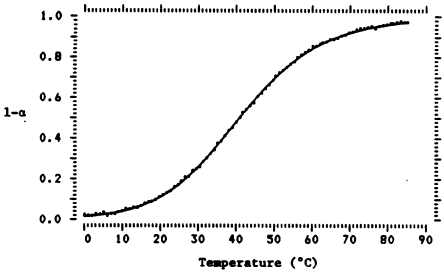

FIGURE 4: Plot of the fraction of X-11 molecules in the coil state (1-α) v. temperature. The dotted curve is the experimental melting profile after adjustment for the temperature-dependent baselines. The solid curve represents a two-state fit to the data. The oligomer concentration was 10 μM with 0.1 M NaCl-10 mM sodium phosphate-0.1 mM EDTA, pH 7.0, as the buffer.

deoxyribo-oligomer GCGCGC shows that the two-state approximation gives accurate results if the temperature dependencies of both baselines are taken into account (18). In addition, significant deviation from two-state behavior should yield systematic variation between the experimental melting curves and two-state fits to the data. Since no such variation was observed for the X-11 transition (Figure 4), it is unlikely that the thermodynamic values (Table II) contain substantial inaccuracies deriving from the two-state approximation.

Duplex-Coil Transitions. The melting curves of duplexes X-9:Y-9 and X-11:Y-11 were analyzed with a modified expression for the standard two-state equilibrium constant (K_d) corresponding to the formation of a duplex from two non-identical complementary strands both in the coil state:

$$K_d = \frac{2\alpha_d(1+K_h)}{(1-\alpha_d)^2 C_t} = \exp\left[\frac{-\Delta H^\circ}{RT} + \frac{\Delta S^\circ}{R}\right] \quad (5)$$

where C_t is the total oligonucleotide strand concentration. α_d, the fraction of strands in the duplex state, was obtained from:

$$\alpha_d = [\varepsilon_s - \varepsilon(T)]/[\varepsilon_s - \varepsilon_d]$$

where $\varepsilon(T)$ is the measured total strand absorptivity for the oligonucleotide mixture and ε_s and ε_d are the average absorptivities of the two strands in the coil and duplex states, respectively. The value of ε_s was calculated from the melting profiles of the single strands (Figure 1), and ε_d was treated as a linear function of temperature: $\varepsilon_d(T) = m_d T + n_d$. The term $1+K_h$ in eq 5 was included to allow for the concomitant participation by one of the strands in a coil-hairpin equilibrium. Since hairpins were not detected for the components of the duplex X-9:Y-9, $K_h = 0$ for this analysis. However, in the case of the helix-coil transition of the duplex X-11:Y-11, where the X-11 strand can form a hairpin, the value for K_h (the hairpin-coil equilibrium constant) was calculated from eq 3 with the enthalpy and entropy values for X-11 given in Table II. The Marquart algorithm (14) was used to fit experimental duplex melting curves with equation 5; ΔH°, ΔS°, m_d, and n_d were the adjustable parameters. For each sample, the oligonucleotide concentration (C_t) was determined from the temperature-dependent absorptivities of the component single strands and from multiple absorbance readings on the melting curves above the points where the duplex-to-coil transitions were complete. For each duplex, the melting curves of 3 samples measured in the standard Tm buffer (0.115 M Na^+) were analyzed to give the average values (± standard deviation):

X-9:Y-9 ΔH° = -65 ± 2 kcal/mol ΔS° = -189 ± 7 e.u.

X-11:Y-11 ΔH° = -86 ± 1 kcal/mol ΔS° = -242 ± 4 e.u.

With the assumption that the nucleation free energy values are the same for

the formation of these two duplexes, the difference in their stabilities must arise from the presence, in the larger duplex, of the two base-paired 5'-CG-3'/5'-CG-3' stacking interactions. Thus a comparison of the parameters of these two duplexes yields the values for this type of stacking interaction:

$$\Delta H^\circ(CG) = -10.5 \text{ kcal/mol} \qquad \Delta S^\circ(CG) = -26.5 \text{ e.u.}$$

Since the possible occurrence of differential end effects in the two duplexes was not taken into account in this calculation, the values may contain systematic errors. However, based on the reported stacking enthalpy for the corresponding interaction in RNA (19,20) and the previous observation that enthalpies of analogous nearest neighbor combinations in RNA and DNA are roughly the same (21), values for the CG stack estimated above are probably accurate to within about 10%.

DISCUSSION

The experimental basis of this work consisted of the accurate measurement and comparison of molar absorptivities and absorptivity-temperature profiles belonging to a set of related oligodeoxyribonucleotides. The absorptivities were determined by a method that exploits the exonucleolytic activity of spleen phosphodiesterase to measure oligonucleotide hyperchromicities. The enzyme chosen for this purpose has two special advantages. It possesses high activity in a pH range that is at least 2 pH units away from any of the pKa values associated with the nucleotide bases. In this range mononucleotide absorptivities are insensitive to moderate changes in pH, and determinations of oligonucleotide absorptivities are therefore unaffected by small experimental inaccuracies in setting the pH. A second advantage lies in the types of products formed by the enzyme. These are readily separated and quantified by a simple ion-exchange chromatographic system, and this permits the subsequent use of the solution resulting from the hyperchromicity measurement for confirmation of the oligomer's nucleotide content and purity.

The absorptivity of each oligomer examined in this study increases with temperature (Figures 1,2), indicating the occurrence of thermally induced conformational changes. These changes must be intramolecular because, in each case, the ε_{260}-T curve is independent of the oligomer's concentration over at least a twenty-fold range. Also, since none of the oligonucleotides listed in Table I contain sequences that would allow extensive intramolecular W-C base pairing, the predominant conformational feature of the molecules should be the stacking of neighboring bases. In general, the formation of this type of interaction is an essentially non-cooperative process (22), and, since the

enthalpy change for each independent structural transition is relatively small, the fraction of bases in the stacked state changes in a nearly linear fashion over a broad range of temperature. The virtually linear ε_{260}-T profiles of oligonucleotides Y-9, X-9, and Y-11 (Figure 1) are therefore consistent with the expectation that such interactions constitute a major portion of the conformational structure in these molecules. In contrast, the profile of X-11 has a sigmoidal shape, implying that this oligomer undergoes a more cooperative transition.

The sequence of the undecamer X-11 differs from that of X-9 simply by the presence of an extra C and G at its terminals (Table I); the sequence of Y-11 (the W-C complement of X-11) differs from that of Y-9 (the W-C complement of X-9) in exactly the same way. In view of these relationships and the fact that Y-11 and Y-9 have virtually identical normalized melting profiles (Figure 1), X-11's profile is unusual in differing so markedly from that of X-9.

The source of the apparent cooperativity in X-11 was determined by examining the subfragments of X-9 and X-11. These subfragments are pairs of shorter oligomers whose sequences (Table I), when appropriately combined, constitute the sequences of the nonamer (X-9a + X-9b) and the undecamer (X-11a + X-11b); they were readily available from the corresponding half-length synthetic blocks used in the solution-phase construction of the longer species. Since the UV hypochromicities of oligonucleotides are sensitive primarily to interactions between base chromophores that are in close proximity (13), the observed value of this optical property for an oligomer should be similar to the combined values of the same property of its subfragments, unless local conformations in the corresponding sequences differ significantly.

The composite melting curve D shown in Figure 2 is derived from measurements on the subfragments X-9a and X-9b along with a term that approximates the effects of local interactions, associated primarily with the central T-T stack, that exist in the nonamer X-9 but not in its two subfragments. The curve corresponds well with the measured profile for X-9, and this similarity, together with the linear shapes, supports the assignment of short-range base stacking as the major structural feature of X-9 and its subfragments. The small discrepencies displayed in the magnitudes of the absorptivities probably result from inadequacy in the approximation used for the central T-T stack, from end effects, and perhaps from errors associated with the values calculated for the absorptivities of the oligomers. Because of the similarity between the sequences of X-9 and X-11, these small quantitative differences should also appear in the comparison of the X-11 curve with that derived from the

corresponding subfragments. However, there is a marked difference in shape between the profile measured from X-11 and the one calculated from X-11a, X-11b and the term for the T-T stack (Figure 2). At high temperatures, the absorptivities of the two curves are nearly identical, but, while the absorptivity of X-11 drops sigmoidally with temperature, the composite curve remains essentially linear. The distinctive shape of the X-11 profile and the apparent cooperativity of the transition must therefore arise from interactions between the segments of the molecule represented by the two subfragments. Moreover, it is apparent that at least one of the two terminal nucleotides of X-11 must play a critical role in the behavior of this molecule, because X-9 (whose sequence differs from that of the undecamer only in lacking these two residues) does not adopt a similar conformation. The CD spectra (data not shown) and proton NMR spectra (23) of the various species also provide evidence that X-11 assumes a low-temperature structure quite unlike its subfragments and X-9.

The most reasonable explanation for these results is that, at low temperature, X-11 folds into a hairpin conformation that involves base pairing between nucleotides at opposite ends of the sequence. The sigmoidal shape of the melting curve would then reflect the cooperative transition between this structure and the (partially stacked) random coil form prevalent at high temperatures. This conclusion is supported by the influence of ionic strength on X-11's melting profile. An eight-fold reduction in sodium ion concentration causes a 6° shift to lower temperatures with essentially no change in the shape of the curve (Figure 3 and Table II), and this implies that the low-temperature form of X-11 has a greater density of phosphate charges than the high-temperature form. In comparison, the melting curves of molecules Y-9, X-9, and Y-11 are relatively insensitive to similar changes in ionic strength; this is consistent with the apparent absence of long-range structure in these molecules.

The Structure and Stability of the X-11 Hairpin. There are several structures that can be considered for X-11, but most of these would require substantial contributions from relatively low stability arrangements such as G:A and G:T base pairs, together with loop sizes of 3 or less. The most attractive structure is a hairpin that contains the contiguous base pairs C^1:G^{11} and G^2:C^{10}, with a loop of 7 nucleotides. This arrangement is consistent with the observed differences in the optical properties of X-9 and X-11 that implicate the terminal nucleotides of the undecamer as being important contributors to the hairpin's stability. Moreover, spectra of this

oligomer obtained in a preliminary ^{1}H-NMR study (23) display an imino proton resonance with a chemical shift (12.6 ppm relative to TSP) characteristic of G:C base pairing. In addition, resonances from non-exchangeable base protons of both of the C residues undergo large chemical shift changes with temperature (23), an observation that cannot be accounted for by any of the other possible structures.

However, thermodynamic analysis of X-11's melting curves in terms of a two-state equilibrium (Table II) indicates that the oligomer's structure is much more stable than predicted from earlier studies on oligonucleotide hairpins. This is evident from a comparison of X-11 with several deoxyribo-oligomers with the sequence ATCCTAT$_n$TAGGAT, each of which has been shown to exist as a hairpin consisting of a double helical stem of six base pairs formed by the two complementary ends together with a loop of four (n = 4) to seven (n = 7) contiguous T residues (5,24). The reported Tm values (at 0.015 M Na^+) range from 44° for the hexadecamer (n = 4) to 32° for the nonadecamer (n = 7), with the stability of these structures decreasing monotonically as loop size is increased (5). The undecamer X-11, in the form of a two G:C base-pair stem and seven-membered loop, would be expected to display a substantially lower stability than the nonadecamer (n = 7) with its six base-pair stem (2G:C,4A:T) and seven-membered loop. In fact, the undecamer has a Tm (Table II) that is 2-3° higher than the longer molecule. However, the parameter Tm, by itself, is a good quantitative measure of stability only for temperatures near Tm. The magnitude of the transition enthalpy for X-11 (19 kcal/mol) is less than half that reported for the nonadecamer ATCCTAT$_7$TAGGAT (42 kcal/mol), and it is the comparatively favorable distribution of free energy between $\Delta H°$ and $\Delta S°$ that endows X-11's hairpin with the higher Tm.

One factor that might partly account for this comparative stability is the existence of electrostatic end effects. The relative influence of such effects has been shown to decrease with increasing duplex size, and they manifest themselves as a lower dTm/dlog[Na^+] for shorter helices (25,26). At higher salt concentrations (ca. 1 M Na^+), the influence of these effects should be small. Unfortunately, the Tm of the nonadecamer in high salt has not been reported; however, by comparing the Tm of X-11 (44 °C) in 1 M Na^+ with that of the hexadecamer (58 °C) under the same conditions, it is possible to estimate that X-11's Tm would still be roughly equal to that of the nonadecamer in the higher salt environment. Therefore, factors other than differences in electrostatic effects within the stems of hairpins must be responsible for the comparative stability of X-11.

The contribution of the GC pairs of the stem to the stability of the X-11 hairpin can be assigned a value corresponding to a single 5'-CG-3'/5'-CG-3' base-paired stacking interaction. This value, obtained from measurements on two duplexes as described in Results, permits the derivation of the parameters for the formation of a loop consisting of the remaining 7 nucleotides:

$\Delta H°(loop) = -8.7$ kcal/mol and $\Delta S°(loop) = -35$ e.u. (0.115 M Na^+).

The striking feature of this result is the apparent large negative enthalpy of loop closure; previously reported loop enthalpies for oligonucleotide hairpins range from 0 to large positive values (2,3,24,27,28). This effect must arise from interactions within the loop that are energetically more favorable than those experienced by the same nucleotides in the coil state, and while it is not possible to derive detailed structural information from absorption measurements, it seems likely that base pairing between the two A residues and/or between G^4 and G^8, together with associated stacking interactions, is responsible for the observed negative enthalpy. Although base pairing of this kind, when located internally within double helical regions, tends to destabilize oligomer duplexes (29), it is possible that the unusual nucleotide conformations required for the formation of purine-purine base pairs would be readily accommodated, and perhaps even favored, at the stem terminus adjacent to the loop of a hairpin. In the case of X-11, the formation of these non W-C pairs, besides adding to the strength of the stem, would increase hairpin stability by reducing the effective loop size. The existence of such special interactions would account for the observed difference in behavior between X-11 and its W-C complement Y-11. The failure of the latter molecule to display cooperative structure, even though it is theoretically capable of forming a two GC base pair stem identical with that of X-11, must be due to its internal sequence, which contains pyrimidines at the positions flanking the C-G terminals. However, there may be other factors contributing to the lack of cooperative structure in Y-11; one of these could be an unfavorable free energy change associated with the destacking of the three central A residues that might be a consequence of hairpin formation. Some of these uncertainties may be resolved by a more extensive study of the contribution of loop sequence and size effects to hairpin stability. Preliminary results from a larger series of oligomers indicate that stabilization can also occur with certain other loop purine arrangements located next to hairpin stems (M. P. Koleck, unpublished).

While the assignment of a detailed structure for X-11 must await the completion of a proton and phosphorus NMR study, it is clear from the results

presented here that non W-C interactions between loop nucleotides can have a marked effect on hairpin stability, and consequently these interactions will have to be taken into account if reliable models of secondary structure for single-stranded nucleic acids are to be derived from thermodynamic data. The magnitude of the effect in X-11 can be evaluated by comparing its hairpin with the analogous structure, mentioned above, that can be assigned to Y-11, namely a stem composed of C^1:G^{11} and G^2:C^{10} pairs with a loop of 7 nucleotides. If only W-C pairing is considered, these hairpins have identical stems and loop sizes and differ only in loop sequences. Differences in the normalized absorbance-temperature profiles of X-9 and X-11 (Figure 1) can be easily detected when α, the fraction of X-11 molecules in the hairpin state, is greater than 0.1. From a similar comparison of the absorbance-temperature profiles for Y-9 and Y-11 (Figure 1), it is clear that the latter oligomer shows no tendency to form a hairpin above 0 °C, a conclusion that is also supported by a preliminary NMR study of the molecule (J. G. Nadeau and E. L. Ulrich, unpublished observations). For the X-11 hairpin at 25 °C and in the presence of 0.115 M Na^+, ΔG° = -0.9 kcal/mol (Table II). With the conservative assumption that $\alpha < 0.1$ for Y-11 at 25 °C, the ΔG° for this hypothetical hairpin must be greater than +1.3 kcal/mol. Thus, in the particular case of X-11 and Y-11, the existence of non W-C pairing and/or other loop sequence effects can produce a difference in stability of at least 2.2 kcal/mol, an amount roughly equal to that contributed to a DNA duplex region by a CG/CG base-paired stacking interaction.

ACKNOWLEDGEMENTS

We thank Dr. G. R. Gough and M. J. Brunden for advice and assistance with syntheses of oligonucleotides. The work was supported by the National Institutes of Health (GM 11518, GM 19395, and Training Grant GM 07211).

REFERENCES

1. Scheffler, I. E., Elson, E. L., and Baldwin, R. L. (1968) J. Mol. Biol. 36, 291-304.
2. Gralla, J. and Crothers, D. M. (1973) J. Mol. Biol. 73, 497-511.
3. Uhlenbeck, O. C., Borer, P. N., Dengler, B., and Tinoco, I., Jr. (1973) J. Mol. Biol. 73, 483-496.
4. Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1973) Nature (London), New Biol. 246, 40-41.
5. 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., and van Boom, J. H. (1983) J. Biomol. Str. and Dyn. 1, 115-129.
6. Wickstrom, E. and Tinoco, I., Jr. (1974) Biopolymers 13, 2367-2383.
7. van Knippenberg, P. H. and Heus, H. A. (1983) J. Biomol. Str. and Dyn. 1, 371-381.

8. Gough, G. R., Brunden, M. J., Nadeau, J. G., and Gilham, P. T. (1982) Tetrahedron Letters 23, 3439-3442.
9. Nadeau, J. G., Singleton, C. K., Kelly, G. B., Weith, H. L., and Gough, G. R. (1984) Biochemistry 23, 6153-6159.
10. Lawson, T. G., Regnier, F. E., and Weith, H. L. (1983) Anal. Biochem. 133, 85-93.
11. Taras, M. J. (1958) in Colorimetric Determination of Non-Metals (Boltz, D. F., Ed.) pp 75-122, Interscience New York.
12. Black, D. M. and Gilham, P. T. (1985) Nucleic Acids Res. 13, 2433-2442.
13. Cantor, C. R. and Tinoco, I. Jr. (1965) J. Mol. Biol. 13, 65-77.
14. Bevington, P. R. (1969) Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, New York.
15. Martin, F. H., Uhlenbeck, O. C., and Doty, P. (1971) J. Mol. Biol. 57, 201-215.
16. Appleby, D. W. and Kallenbach, N. R. (1973) Biopolymers 12, 2093-2120.
17. Petersheim, M. and Turner, D. H. (1983) Biochemistry 22, 256-263.
18. Albergo, D. D., Marky, L. A., Breslauer, K. J., and Turner, D. H. (1981) Biochemistry 20, 1409-1412.
19. Borer, P. N., Dengler, B., Tinoco, I. Jr., and Uhlenbeck O. C. (1974) J. Mol. Biol. 86, 843-853.
20. Freier, S. M., Petersheim, M., Hickey, D. R., and Turner, D. H. (1984) J. Biomol. Str. and Dyn. 1, 1229-1242.
21. Marky, L. A. and Breslauer, K. J. (1982) Biopolymers 21, 2185-2194.
22. Bloomfield, V. A., Crothers, D. M., and Tinoco, I. Jr. (1974) Physical Chemistry of Nucleic Acids, Harper and Row, New York.
23. Ulrich, E. L., Nadeau, J. G., and Gilham, P. T. (1985) Fed. Proc. Fed. Am. Soc. Exp. Biol. 44, 1616.
24. van Boom, J. H., van der Marel, G. A., Westerink, H., van Boeckel, C. A. A., Mellema, J. R., Altona, C., Hilbers, C. W., Haasnoot, C. A. G., de Bruin, S. H., and Berendsen, R. G. (1983) Cold Spring Harbor Symp. Quant. Biol. 47, 403-409.
25. Elson, E. L., Scheffler, I. E., and Baldwin, R. L. (1970) J. Mol. Biol. 54, 401-415.
26. Record, M. T. Jr. and Lohman, T. M. (1978) Biopolymers 17, 159-166.
27. Porschke, D. (1974) Biophys. Chem. 1, 381-386.
28. Marky, L. A., Blumenfeld, K. S., Kozlowski, S., and Breslauer, K. J. (1983) Biopolymers 22, 1247-1257.
29. Aboul-ela, F., Koh, D., Tinoco, I. Jr., and Martin, F. H. (1985) Nucleic Acids Res. 13, 4811-4824.