Guo, Q., Lu, M., Sharestanifar, M., Sheardy, R. D., & Kallenbach, N. R. (1991) *Biochemistry* (in press).
Hardwick, J. M., Von Sprecken, R. S., Yielding, K. L., & Yielding, L. W. (1984) *J. Biol. Chem. 259*, 11090–11097.
Hattman, S., Gribbin, C., & Hutchison, C. A., III (1979) *J. Virol. 32*, 845–851.
Hosur, R. V., Chary, K. V. R., Kumar, A., & Govil, G. (1985a) *J. Magn. Reson. 62*, 123–127.
Hosur, R. V., Ravikumar, M., & Sheth, A. (1985b) *J. Magn. Reson. 65*, 375–381.
Kilpatrick, M. W., Wei, C.-F., Gray, H. B., Jr., & Wells, R. D. (1983) *Nucleic Acids Res. 11*, 3811–3822.
Kruger, D. H., Bareak, G. J., Reuter, M., & Smith, H. O. (1988) *Nucleic Acids Res. 16*, 3997–4008.
Krugh, T. R., Hook, J. W., III, Balakrishnan, M. S., & Chen, F.-M. (1980) in *Nucleic Acid Structure and Dynamics* (Sarma, R. H., Ed.) Pergamon Press, New York.
Lavery, R. (1988) in *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds.) Springer-Verlag, New York.
Lown, J. W., Morgan, A. R., Yen, S.-F., Wang, Y.-H., & Wilson, W. D. (1985) *Biochemistry 24*, 4028–4035.
Lu, M., Guo, Q., Pasternak, R. F., Wink, D. J., Seeman, N. C., & Kallenbach, N. R. (1990) *Biochemistry 29*, 1614–1624.
Mallamaci, M. A., Bascoy, M., Brown, J., Combates, N. P. & Winkle, S. A. (1991) *J. Biomol. Struct. Dyn.* (submitted for publication).
Manning, G. S. (1988) *Biopolymers 27*, 1529–1545.
Oller, A. R., Vanden Brock, W., Conrad, M. & Topal, M. D. (1991) *Biochemistry 30*, 2543–2549.
Pein, C.-D., Reuter, M., Cech, D., & Kruger, D. (1989) *FEBS Lett. 245*, 141–144.
Ravikumar, M., Hosur, R. V., Roy, K. B., Miles, H. T., & Govil, G. (1985) *Biochemistry 24*, 7703–7711.
Royer-Pokora, B., Gordon, L. K., & Haseltine, W. A. (1981) *Nucleic Acids Res. 9*, 4595–4609.
Shafer, R. H., Brown, S. C., Delbane, A., & Wack, D. (1984) *Nucleic Acids Res. 12*, 4679–4690.
Sheardy, R. D. (1988) *Nucleic Acids Res. 16*, 1153–1167.
Sheardy, R. D., & Winkle, S. A. (1989) *Biochemistry 28*, 720–725.
Suh, D., Sheardy, R. D., & Chaires, J. B. (1991) *Biochemistry 30*, 8722–8726.
Ushay, H. M., Tullius, T. D., & Lippard, S. J. (1981) *Biochemistry 20*, 3744–3748.
Vardimon, L., & Rich, A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3268–3272.
Walker, G. T., Stone, M. P., & Krugh, T. R. (1985a) *Biochemistry 24*, 7462–7471.
Walker, G. T., Stone, M. P., & Krugh, T. R. (1985b) *Biochemistry 24*, 7471–7479.
Winkle, S. A., & Sheardy, R. D. (1990) *Biochemistry 29*, 6514–6521.
Winkle, S. A., Alayo, M., Lee-Chee, T., Graves, D. E., & Sheardy, R. D. (1991) *Biophys. J. 59*, 572a.

Articles

Thermodynamic and Structural Properties of Pentamer DNA·DNA, RNA·RNA, and DNA·RNA Duplexes of Identical Sequence†

Kathleen B. Hall*
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Box 8231, 660 South Euclid Avenue, St. Louis, Missouri 63110
Larry W. McLaughlin*
Department of Chemistry, Boston College, 140 Commonwealth Avenue, Chestnut Hill, Massachusetts 02167
Received January 24, 1991; Revised Manuscript Received August 21, 1991

ABSTRACT: Four pentamers with the general sequence $^{5'}$CU(T)GU(T)G/$^{5'}$CACAG have been prepared by chemical synthesis in order to generate duplex structures with common sequences. The four duplexes studied include the DNA·DNA duplex ($^{5'}$dCACAG/$^{5'}$dCTGTG) and the RNA·RNA duplex ($^{5'}$rCUGUG/$^{5'}$rCACAG) as well as the two corresponding DNA·RNA heteroduplexes ($^{5'}$rCUGUG/$^{5'}$dCACAG and $^{5'}$CACAG/$^{5'}$dCTGTG). The measured entropy, enthalpy, and free energy changes upon melting are reported for each pentamer and compared to the predicted values where possible. Results show that the two DNA·RNA heteroduplexes are destabilized ($\Delta G°_{25} = -4.2 \pm 0.4$ kcal/mol) relative to either the DNA·DNA duplex ($\Delta G°_{25} = -4.8 \pm 0.5$ kcal/mol) or the RNA·RNA duplex ($\Delta G°_{25} = -5.8 \pm 0.6$ kcal/mol). Circular dichroism spectra indicate that the RNA and the two heteroduplexes adopt an A-form conformation, while the DNA conformation is B-form. Imino proton NMR spectra also show that the heteroduplex structures resemble the RNA·RNA duplex.

Double-stranded RNA sequences adopt an A-form conformation, while double-stranded DNA is usually B-form. DNA can exhibit B to A transitions either through sequence effects (Leslie et al., 1980) or as a result of the solvent conditions (Drew et al., 1980), while the corresponding A to B transition for RNA has not been observed. Significant changes in the salt conditions for solutions of either DNA or RNA can additionally effect a B(A) to Z transition (Pohl & Jovin, 1972;

† Supported by the Lucille P. Markey Charitable Trust No. 88-30 (K.B.H.) and by a grant to L.W.M. from the National Science Foundation (DMB-8904306).
* Correspondence can be directed to either author.

0006-2960/91/0430-10606$02.50/0

Hall et al., 1984). Although one DNA and one RNA strand are present in the DNA·RNA heteroduplex, the limited studies reported to date indicate that A-form structures are generally found (Milman et al., 1967; Mellema et al., 1983; Reid et al., 1983; Sanger, 1984; Chou et al., 1989), the sole exception being poly(rA)·poly(dT) which under some solution conditions can adopt a B-form conformation (Zimmerman & Pheiffer, 1981). Several studies compare the thermodynamic stabilities (Chamberlin & Berg, 1964; Martin & Tinoco, 1980; Nelson & Tinoco, 1982) or structural parameters (Pardi et al., 1981; Reid et al., 1983; Chou et al., 1989) of DNA, RNA, and hybrid duplexes with common sequences.

The physical chemical questions of the relative stabilities of RNA·DNA heteroduplexes, and the sequence dependence of duplex formation are important to understand the priming step in DNA replication (Kornberg, 1980), transcription termination processes (Farnham & Platt, 1980; Martin & Tinoco, 1980; Yager & von Hippel, 1987), and the use of heteroduplex molecules for antisense regulation of gene expression (Cohen, 1989). Previous studies with synthetic homopolymers have shown that there is a sequence dependence to the relative stabilities of DNA·DNA, RNA·RNA, and RNA·DNA duplexes (Riley et al., 1966; Chamberlin & Patterson, 1965; Martin & Tinoco, 1980). Other work with naturally occurring sequences (Chamberlin & Berg, 1964; Birnsteil et al., 1972) showed that hybrid duplexes were less stable and that the GC content influenced the magnitude of the relative stabilities.

For the experiments described here, we have used pentamer sequences to form DNA·DNA, RNA·RNA, and RNA·DNA hybrid molecules. The specific RNA·RNA sequence comprises the T-stem of yeast $tRNA^{Phe}$, and its structure based on NMR experiments was described previously (Clore et al., 1985). This sequence is useful as a model system for different duplex forms since it contains pyrimidine/purine steps, purine/pyrimidine steps, and purine/purine (pyrimidine/pyrimidine) steps and thus avoids the potential problems of homopolymer tracts.

Materials and Methods

RNA Synthesis. The RNA pentamers were synthesized in solution using hydroxybenzotriazole-activated phosphotriesters by techniques described previously (van der Marel et al., 1982; Wreesman et al., 1983). The exocyclic amino groups of guanosine and adenosine were protected as benzoyl amides, and cytidine was protected as the *p*-anisoyl amide. The 2′-hydroxyl groups were protected as the 2-tetrahydropyranyl ether derivatives except for the 3′-terminal residues which were protected as the 2′,3′-diacetyl esters. The 5′-terminal residue was protected as the 4,4′-dimethoxytrityl derivative. As an example of the coupling procedure, the final reaction between the protected derivatives of CpUpG and UpG yielding the fully protected CpUpGpUpG pentamer is described below.

To 613 mg (0.31 mmol) of the protected trimer (CpUpG) containing a free 3′-terminal hydroxyl, coevaporated from pyridine (3×), was added 1.85 mL of a 0.2 M solution of *o*-chlorophenyl bis(1-benzotriazolyl) phosphate (0.37 mmol) (van der Marel et a., 1982) in anhydrous dioxane, and the reaction was stirred under an anhydrous atmosphere. After 30 min, this solution was transferred to a dry flask containing 317 mg (0.30 mmol) of the protected dimer (UpG) with a free 5′-hydroxyl which had been dried by coevaporation from pyridine (3×). To this solution was added 0.12 mL of dry *N*-methylimidazole, and the mixture was stirred for 2 h at ambient temperature. The reaction was stopped by addition of 2 mL of water, worked up in the normal fashion, and purified by column chromatography on silica gel. Fractions containing pure fully protected pentamer were pooled and evaporated to dryness. Yield: 592 mg (0.18 mmol), 61%.

Deprotection of the RNA pentamers occurred in three steps. As an example, 100 mg (31 μmol) of the fully protected CpUpGpUpG pentamer was dissolved in 2 mL of dioxane/acetonitrile (1/1). To 183 mg (1.3 mmol) of pyridinaldoxime in 4 mL of dioxane/acetonitrile (1/1) was added 150 mg (1.3 mmol) of tetramethylguanidine, and 3.5 mL of this mixture (0.325 M) was added to the pentamer solution. After incubation at ambient temperature overnight, 25 g of Dowex 50W-X8 (NH_4^+ form) was added, and the suspension was filtered 5 min later. The filtrate was evaporated to dryness, dissolved in 40 mL of concentrated ammonia, and incubated at 50 °C overnight followed by 24 h at ambient temperature. This solution was evaporated to dryness, the residue was dissolved in 60 mL of 0.01 M HCl, and the pH was adjusted to 2.0. After 48 h at ambient temperature, this solution was added directly to a 1.5 × 15 cm column of Sephadex A-25 and eluted with a 2-L linear gradient of 0.05–0.7 M triethylammonium bicarbonate, pH 7.5. The product was the major peak which eluted between fractions 86 and 106 (about 15-mL fractions). These fractions were pooled, evaporated to dryness, and desalted on a Sephadex G-10 column. The appropriate fractions from the G-10 column were collected and lyophilized to dryness. Yield: 985 A_{260} units (21 μmol), 68%.

DNA Synthesis. The DNA pentamers were synthesized on CPG supports using phosphite triester chemistry as described previously (Beaucage & Caruthers, 1981). Each pentamer synthesis employed 10 μmol of CPG-bound 2′-deoxyguanosine in a glass column 13 mm in diameter by 15 mm long, containing plastic Leurlock fittings with Teflon filters. The pressure on the DNA synthesizer was raised to 4.2 bar which produced a flow of 3.0 mL/min of acetonitrile to the column. For each coupling, 100 μmol of the appropriate phosphoramidite dissolved in 1.0 mL of dry acetonitrile was delivered to the column with tetrazole in three aliquots. After addition of each aliquot, the coupling was allowed to proceed for 90 s. Each cycle also contained the normal capping, oxidation, and acid treatment with appropriate column washes. The support was mixed at various times by allowing dry argon to bubble through the CPG support. Coupling yields exceeded 98% at each step as determined by collection of the acid wash and measurement of the amount of DMT cation present by absorption at 498 nm.

After completion of the synthesis, the CPG support was treated with concentrated ammonia and the product isolated by HPLC as described elsewhere (McLaughlin & Piel, 1984). Yields of purified pentamer from a single 10-μmol synthesis: d(CpTpGpTpG), 2.76 μmol; d(CpApCpApG), 3.98 μmol.

Oligonucleotide Solutions. Absorption in the thermal denaturation experiments was measured at 280 nm because absorbance at 260 nm was too great to allow accurate measurement at the concentrations required. Since extinction coefficients for DNA are not available at 280 nm, the values at 260 nm for the RNA and DNA were compared to determine the difference. At 260 nm, the extinction coefficients of the RNA and DNA single strands were calculated from published dimer and monomer coefficients (Richards, 1975; T'so, 1974), using the nearest-neighbor approximation (Borer et al., 1975). Extinction coefficients determined: rCUGUG, ϵ_{260} = 45 600 M^{-1} cm^{-1}; dCTGTG, ϵ_{260} = 44 300 M^{-1} cm^{-1}; dCACAG, ϵ_{260} = 50 400 M^{-1} cm^{-1}; rCACAG, ϵ_{260} = 50 000 M^{-1} cm^{-1}. Values for the analogous DNA and RNA sequences are within 1–3%, so the available values for the RNA at 280

Table I: Enthalpy and Entropy of Transition and the Free Energy of Duplex Formation

(a) Enthalpy and Entropy of Transition[a]

pentamer duplex	log C_T parameters $-\Delta H°$ (kcal/mol)	log C_T parameters $-\Delta S°$ (cal/K·mol)	temperature-independent parameters $-\Delta H°$ (kcal/mol)	temperature-independent parameters $-\Delta S°$ (cal/K·mol)	T_M (1×10^{-4} M)[b]
rCACAG·rCUGUG	40.2	115.4	41.1	118.2	22
dCACAG·dCTGTG	34.0	98.1	33.4	96.0	12
rCACAG·dCTGTG	31.3	90.6	34.5	101.2	9
dCACAG·rCUGUG	31.0	90.6	35.2	104.7	6

(b) Free Energy of Duplex Formation[d]

pentamer duplex	log C_T parameters[c] $-\Delta G°_{37}$	log C_T parameters[c] $-\Delta G°_{25}$	log C_T parameters[c] $-\Delta G°_{13}$	temperature-independent parameters[c] $-\Delta G°_{37}$	temperature-independent parameters[c] $-\Delta G°_{25}$	temperature-independent parameters[c] $-\Delta G°_{13}$
rCACAG·rCUGUG	4.7	5.8	7.2	4.4	5.9	7.3
dCACAG·dCTGTG	3.6	4.8	5.9	3.6	4.8	5.9
rCACAG·dCTGTG	3.2	4.3	5.4	3.1	4.3	5.5
dCACAG·rCUGUG	2.9	4.0	5.1	2.7	3.9	5.2

[a] Error is estimated at ±10%. The extra decimal places are reported for subsequent calculation of the free energy. [b] Temperature is in degrees Celsius. [c] Error is estimated at ±10%. [d] Values are given in kilocalories per mole.

nm and 90 °C were used for DNA as well. Those values are $\epsilon_{280} = 24\,340\ M^{-1}\ cm^{-1}$ for rCUGUG and $\epsilon_{280} = 24\,190\ M^{-1}\ cm^{-1}$ for rCACAG, with an average for the duplex mixture of $24\,260\ M^{-1}\ cm^{-1}$. Solutions for melting curves contained 1 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7. Solutions for the NMR experiments contained 0.5 M NaCl and 10 mM sodium cacodylate, pH 7.

Thermodynamic Analysis. Melting curves were measured on a Gilford 250 spectrophotometer, equipped with a 2527 thermoprogrammer in the laboratory of Prof. D. Turner at the University of Rochester. Concentrations of the pentamer duplexes varied from 6×10^{-5} M to 3×10^{-3} M for the RNA and from 2×10^{-4} to 4×10^{-3} for the DNA and the heteroduplexes. The low T_M values observed for the pentamers dictated the available concentration ranges. The heating rate was 1 deg/min, and each sample was heated to 90 °C and slowly cooled before recording the melting curve.

Thermodynamic parameters were derived from absorbance vs temperature curves as described previously (Petersheim & Turner, 1983; Freier et al., 1983). The data were analyzed using programs (Freier et al., 1983) modified by A. Peritz. Two methods were used for calculating thermodynamic parameters: $\Delta H°$ and $\Delta S°$ can be obtained from plots of log (C_T) vs the reciprocal of the melting temperature (T_M^{-1}) using the following relationship (Borer et al., 1974):

$$T_M^{-1} = \frac{2.33}{\Delta H°} \log (C_T) + \frac{\Delta S°}{\Delta H°}$$

where C_T is the total strand concentration. Each melting curve was also fit to a two-state model to obtain $\Delta H°$ and $\Delta S°$, which were then averaged over the number of curves (see supplementary material). The predicted thermodynamic parameters for duplex formation for the RNA and DNA pentamers were calculated from the respective published dimer values of Freier et al. (1986) and Breslauer et al. (1986).

Circular Dichroism Spectra. Spectra were recorded on a Jobin/Yvon Mark V autodichrograph spectropolarimeter in the laboratory of Prof. G. Fasman at Brandeis University, using the samples from the thermal denaturation experiments which had duplex concentrations from 0.5 to 1.0 mM. Spectra were recorded from 230 to 330 nm at 6.5 °C, using a water-jacketed circular cuvette with a path length of 0.5 mm. Two spectra were averaged for each pentamer duplex to give the recorded spectrum. Ellipticity was reported as $A_L - A_R$ rather than as molar ellipticity due to the uncertainty in the pentamer concentrations.

NMR Experiments. The pentamer single strands were dissolved in 0.5 mL of sterilized water, and the concentration was determined by the absorbance at 260 nm of a series of dilutions, using the calculated extinction coefficients. Strands were then mixed in equimolar proportions to form the duplexes. To each solution, 5 M NaCl and 1 M sodium cacodylate, pH 7.0, were added to give a final concentration of 0.5 M NaCl and 10 mM sodium cacodylate. Imino proton data were obtained using either a 2–1–4 pulse or a 1–1 pulse to suppress the water resonance. For the 2–1–4 pulse, the carrier was centered in the imino proton region, with a spectral width of 8 kHz and a recovery time of 500 ms. NOE difference spectra were recorded in the interleaved mode with the off-resonance spectrum. To observe nonexchangeable protons, samples were lyophilized three times from D_2O. Spectra were recorded with a spectral width of 3–6 kHz, with a recovery time of 2–4 s. NOESY, TOCSY, and DQF-COSY spectra were recorded at various temperatures for all samples. Spectra were recorded on the 500-MHz spectrometer in the laboratory of Prof. A. Redfield at Brandeis University and on the 600-MHz spectrometer at Washington University.

Results

The RNA pentamer sequence chosen for this study corresponds to the sequence of the TΨC stem in yeast $tRNA^{Phe}$. It thus reflects a naturally occurring pentameric helix within the structure of the larger tRNA molecule. Although the sequence is relatively short, previous NMR studies (Clore et al., 1985) have indicated that it exists as a duplex structure under a variety of conditions. Previous reports (Petersheim & Turner, 1983) have also indicated that short RNA sequences of this type are amenable to thermodynamic analyses.

Temperature-Independent Thermodynamic Parameters. Plots of T_M^{-1} vs log C_T are shown in Figure 1. These data were used to calculate the thermodynamic parameters given in Table I. The error in the data is estimated to be about ±10% because the low melting temperature of these pentamers decreased the accuracy of the measurements. The error in the data from the heteroduplexes is greatest, due to the lower stability of these helices. Table I gives the values obtained both from averaging the fits to a two-state model and from the T_M^{-1} vs log C_T plots, as described (Freier et al., 1983; Petersheim & Turner, 1983). The values calculated from the two methods agree within 5%, indicating that the assumption of a two-state transition is valid for these molecules (Santa-Lucia et al., 1991).

There is a clear trend in relative stabilities of these pentamer duplexes. The RNA duplex is the most stable, with a calcu-

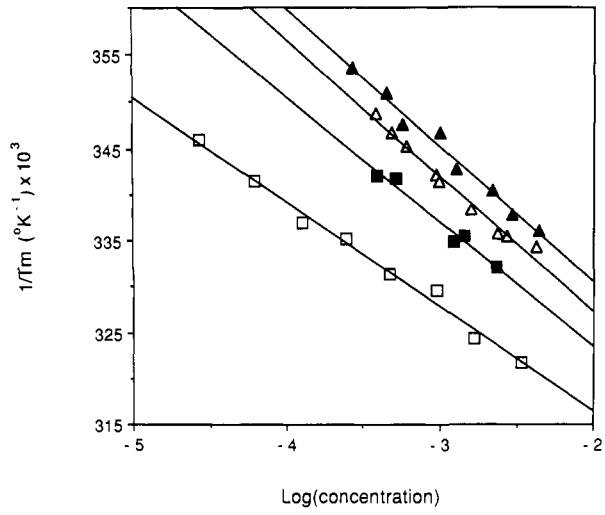

FIGURE 1: Plot of $1/T_M$ vs log C_T for the four pentamer duplexes in 1 M NaCl/10 mM sodium cacodylate/0.5 mM EDTA, pH 7: □, RNA duplex; ■, DNA duplex; △, rCACAG·dCTGTG; ▲, rCUGUG·dCACAG.

lated T_M (at 1×10^{-4} M) of 22 °C. The T_M of the analogous DNA helix is lower by 10 °C. The two heteroduplexes are the least stable (9 °C and 6 °C). The low melting temperature complicates the T_M measurements which increases the error introduced into the $\Delta H°$ and $\Delta S°$ determination for the hybrid pentamers. The similarity of the T_M values observed for the two heteroduplexes suggest that a particular ribo or deoxy strand is not a major determinant influencing the stabilization of the duplex.

CD Spectra. The circular dichroism spectra of the four pentamer duplexes were recorded at 6.5 °C and 1 M NaCl. The concentrations of the strands were approximately 1 mM for the DNA and heteroduplexes and about 0.5 mM for the RNA. At this temperature and these concentrations, all the molecules should be present as duplex structures.

The CD spectra of the four duplexes are shown in Figure 2. The spectrum of the DNA duplex is not the simple conservative spectrum normally observed for DNA (Tunis-Schneider & Maestre, 1970), for it contains a negative band at 295 nm. Such deviation from classical behavior could be due to the nature of a short duplex, which may not be completely stacked, or could arise from the individual transitions of the bases which are averaged in longer molecules. However, the DNA CD spectrum does show the characteristic B-form cross-over near 260 nm with a large negative ellipticity at 250 nm. The spectra of the other three duplexes show the characteristic nonconservative A-form CD spectrum, with a large positive ellipticity near 270 nm and a small negative band near 290 nm. The nearly identical CD spectra of the RNA and the two DNA·RNA duplexes suggest that these three duplexes all adopt an A-form conformation under these experimental conditions.

Imino Proton NMR. The imino proton region of the NMR spectrum of the four duplexes can be compared in Figure 3. The concentration of the strands varies from 3 to 5 mM; at the lower temperatures shown the strands should be paired.

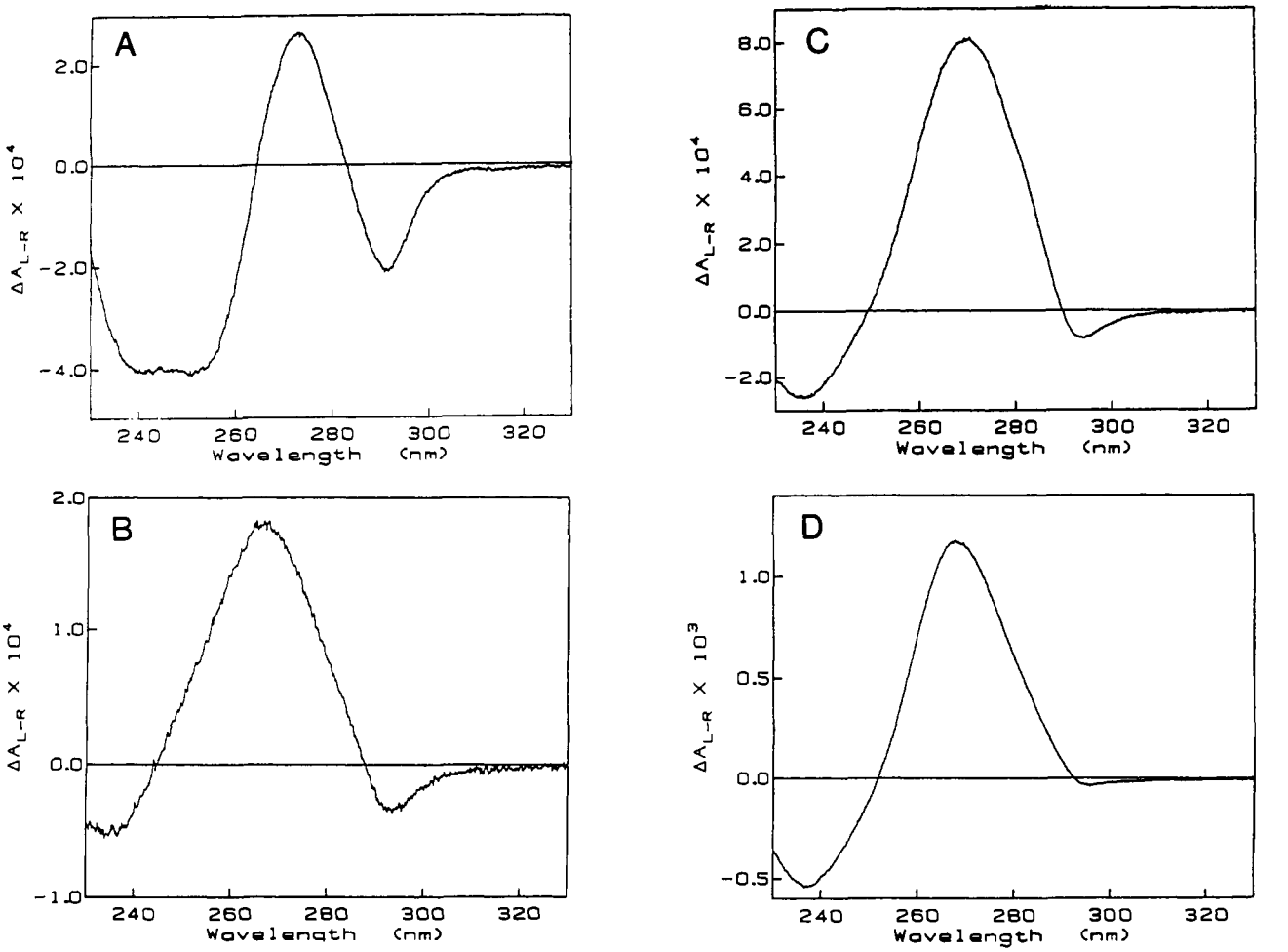

FIGURE 2: Circular dichroism spectra of the four pentamer duplexes at 6.5 °C in 1 M NaCl/10 mM sodium cacodylate/0.5 mM EDTA, pH 7: (A) DNA; (B) RNA; (C) rCUGUG·dCACAG; (D) rCACAG·dCTGTG.

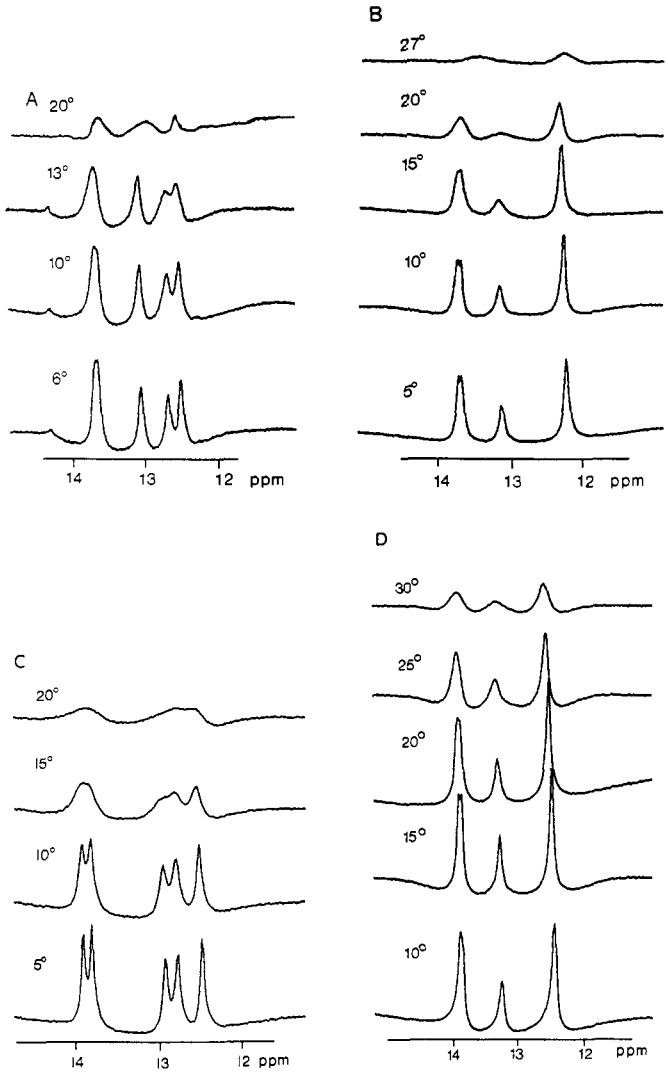

FIGURE 3: Imino proton NMR spectra of the pentamer duplexes as a function of temperature: (A) rCUGUG·dCACAG; (B) rCACAG·dCTGTG; (C) DNA; (D) RNA. Spectra were measured at 500 MHz using a 2-1-4 pulse for water suppression. Carrier was centered in the middle of the imino proton region (11 ppm) with a spectral width of 8000 kHz.

There are two patterns of spectra observed. The DNA duplex exhibits one pattern, in which the resonances of all five imino protons are clearly separated. The imino proton spectra of rCUGUG/dCACAG looks similar to this. The second pattern includes the RNA and the rCACAG/dCTGTG duplex, in which the dispersion is far less.

Assignment of the imino proton resonances used imino–imino NOEs to define the order of protons in the sequence. The downfield resonances, near 13.8 ppm, arise from the uridine (thymidine) imino proton hydrogen-bonded in each AU(T) base-pair. These assignments are based upon the strong NOE observed between these resonances and their respective AH2 resonances in the aromatic region. The other three resonances are assigned to imino protons from the three GC base-pairs. The overlap of these resonances in the spectrum of the RNA and that of the rCACAG/dCTGTG heteroduplex complicated assignment of specific resonances. Assignment of the imino proton resonances to a specific base-pair in the helix of an A-form structure uses the distinctive AH2 to H1′ NOEs described previously (Wuthrich, 1986). In canonical A-form helices (Arnott, 1970), the nonexchangeable AH2 protons, located in the minor groove, are within 4.1 Å of the (deoxy)ribose H1′ protons of two nucleotides: the adjacent intrastrand nucleotide on the 3′ side of A; and the adjacent interstrand nucleotide on the 3′ side of the U(T) base. Two cross-peaks, corresponding to the frequencies of these two H1′ protons, are thus observed in NOESY experiments at the frequency of each AH2 proton (if the helix is regular and the structure is stable). Each AH2 proton is additionally correlated with a specific AU(T) imino proton on the basis of the reciprocal strong NOE observed between them, since the imino proton in an AU base-pair is 2.8 Å from the AH2 proton. After assignment of all the sugar H1′ resonances together with their attached base (data not shown), the imino protons were assigned to the base-pairs in the sequence.

Table II: Assignments of Imino Proton Resonances[a]

base-pair	DNA[b]	RNA	rCACAG·dCTGTG	dCACAG·rCUGUG
C_1G_{10}	12.82	12.42	12.28	12.52
$T(U)_2A_9$	13.78	13.87	13.73	13.68
A_9H2	7.72	6.97	7.14	7.14
G_3C_8	12.54	12.42	12.28	13.06
$T(U)_4A_7$	13.89	13.87	13.70	13.68
A_7H2	7.53	7.35	7.25	7.56
G_5C_6	12.96	13.12	13.12	12.68

[a] Resonances are given at 10 °C in parts per million relative to H_2O at 4.8 ppm. Buffer was 0.5 M NaCl/10 mM sodium cacodylate, pH 7.0. Error is ±0.03 ppm. Imino protons were ordered in the helix sequence using imino–imino NOEs and were ordered 5′-3′ for RNA using the nonexchangeable proton assignments (data not shown). [b] The given DNA order is arbitrary, for the 5′-3′ order cannot be determined from these experiments.

The assignments for the pentamers are given in Table II. The DNA imino protons in the B-form helix cannot be assigned directly to specific base-pairs since there is no spectroscopic correlation between imino protons and other resonances. However, it is possible to identify those imino protons arising from the terminal base-pairs on the basis of the melting behavior and the imino–imino proton NOE pattern.

Imino proton spectra of the pentamers as a function of temperature are shown in Figure 3. It is evident that the DNA duplex and the heteroduplexes exhibit more fraying at the ends prior to thermally induced cooperative denaturation than does the RNA. Both terminal base-pairs of the DNA were relatively unstable, as measured by the temperature dependence of the imino proton resonances. In contrast, only one imino proton resonance of the heteroduplexes melts first (at 13.12 ppm in the spectrum of rCACAG/dCTGTG and at 12.68 ppm in that of dCACAG/rCUGUG): this imino proton resonance has been assigned to the G_5C_6 base-pair, on the basis of the AH2 to H1′ data (data not shown).

Discussion

Stoichiometry of Strands. These pentamer sequences are not self-complementary, and so they might be expected to form double-stranded complexes only with the complementary strand. However, it is also possible that strand self-association could occur; this interaction would be observed by NMR measurements of exchangeable and nonexchangeable protons as well as suggested by analysis of the thermal melting data.

All single strands were examined by NMR experiments in which the chemical shifts of the nonexchangeable protons were observed as a function of temperature. The chemical shifts of the aromatic protons (data not shown) were observed to vary smoothly with temperature from 65 to 10 °C at concentrations of 3–5 mM. At lower temperatures, there was evidence of aggregation for all single strands, in particular the RNA strands, as shown by the increased line width. Such aggregation was apparent in the spectra of the duplexes as well at

Table III: Comparing Predicted and Measured Thermodynamic Parameters

	predicted[a]					measured[b]	
duplex	$\Delta H°$ (kcal/mol)	$\Delta S°$ [c] (eu)	$\Delta G°_{37}$ [d] (kcal/mol)	T_M(10^{-4} M) (°C)	T_M(10^{-3} M) (°C)	$\Delta G°_{37}$ (kcal/mol)	T_M (°C)
RNA	−38.8	−111.8	−4.0	18.9	29.3	−4.7	30
DNA	−25.9	−80.7	−0.8	−18.5	−6.6	−3.6	25
dCACAG/rCUGUG						−2.9	15
rCACAG/dCTGTG						−3.2	19

[a] Predicted values were calculated from the nearest-neighbor approximation (Borer et al., 1974). Values for dimer pairs come from Freier et al. (1986) (RNA) and from Breslauer et al. (1986) (DNA). [b] Values were measured in these experiments. T_M values were measured at the following concentrations: RNA, 9.46×10^{-4} M; DNA, 1.23×10^{-3} M; dCACAG/rCUGUG, 1.02×10^{-3} M; rCACAG/dCTGTG, 9.87×10^{-4} M. [c] Values include −10.8 eu for RNA initiation and −16.8 eu for DNA initiation. [d] Values include +3.4 kcal/mol for RNA initiation and +5.2 kcal/mol for DNA initiation.

lower temperatures and might thus account for the large hyperchromicity observed for the RNA at a concentration of 3 mM.

There was no evidence of additional new resonances appearing at any temperature—such new resonances would be characteristic of alternative structures (with different chemical shifts and magnetic environments). Also, if the strands were not present in equimolar concentrations, then the nonexchangeable proton spectra would contain a disproportionate number of resonances. The spectra of single strands and of duplexes (not shown) were apparently free of any extraneous resonances.

The melting curves were accurately fit using a two-state model and exhibited no evidence of a biphasic transition even at the highest concentrations (3 mM). The agreement between the two methods of calculating thermodynamic parameters (from fits of the melting data and from T_M^{-1} vs log C_T plots) also shows that the two-state algorithm is appropriate. In cases where the association is not two-state, the results from the two methods will differ significantly [see SantaLucia et al. (1991) for an example].

From these data, we conclude that no self-association is occurring for any of the strands, nor is there any evidence to suggest that the interaction is anything but a double-stranded duplex.

Predicted Thermodynamic Parameters. Values for enthalpy ($\Delta H°$), entropy ($\Delta S°$), free energy ($\Delta G°$), and the melting temperatures (T_M) were calculated for two of the pentamer duplexes using the nearest-neighbor approximation (Borer et al., 1974) with values for RNA dimers from Freier et al. (1986) and for DNA dimers from Breslauer et al. (1986). These predicted values (Table III) are compared with the $\Delta G°_{37}$ calculated from the experiments and the measured T_M at a concentration near that of the calculated example.

The predicted and experimental thermodynamic values for the RNA duplex are in moderate agreement. The free energy values agree within 15%. The source of the discrepancy in the two values is not obvious but could result from the inherent inaccuracy of the measurements using molecules with low T_M values. However, the data for the DNA duplexes are quite different, in that there is little correlation between the experimental and predicted values for $\Delta G°$ or for the melting temperature. The source of this discrepancy is unclear. It is noteworthy that the experimental data base of thermodynamic values includes a far greater number of RNA sequences than DNA sequences. The DNA pentamer sequence may include nearest-neighbor sequences which are not adequately represented in the available data base.

No attempt was made to predict thermodynamic values for the hybrid duplexes, since there is no data base available for these structures. The lower melting temperatures of these hybrid duplexes indicate that neither the RNA nor DNA values would be appropriate for prediction.

Duplex Properties. Several aspects of the thermodynamic data presented here should be considered. First is the observation that both heteroduplexes are less stable than the DNA·DNA or RNA·RNA duplexes. Second is the apparent sequence dependence of local structural defects in the hybrid duplexes.

Relative Stabilities of Duplexes. The data reported for this general pentamer sequence indicate that these heteroduplexes are less stable than either the corresponding RNA·RNA or DNA·DNA duplex. The heteroduplexes are destabilized by 1.5–2.0 kcal/mol ($\Delta\Delta G°$) relative to the RNA·RNA duplex or by approximately 0.3 kcal/mol per base-pair. This destabilization is similar for either combination of DNA and RNA strands, suggesting that the instability for this sequence is not associated exclusively with one strand.

Earlier experiments with naturally occurring sequences showed the same relative stabilities reported here: DNA·RNA < DNA·DNA < RNA·RNA. Those experiments used ϕX DNA as a model system and showed that the RNA·DNA form had a lower T_M and broader melting transition than the corresponding DNA·DNA form (Chamberlin & Berg, 1964). Melting temperatures of the two ϕX forms were also measured as a function of the sodium concentration, and results showed that the relative stabilities were unchanged over a 100× range of sodium ion concentration. The ϕX results suggest that the stabilities reported for the pentamers may not depend on the length of the helix or on the solution conditions used. Other experiments with natural sequences (Casey & Davidson, 1977) showed that addition of high concentrations of formamide did stabilize RNA·DNA hybrids relative to the DNA·DNA duplexes, however.

Previous experiments with homopolymers have shown that some heteroduplexes are also less stable than the analogous RNA or DNA duplexes. Examples of such observations include the following relative stabilities: poly(dI)·poly(rC) < poly(dI)·poly(dC) < poly(rI)·poly(dC) < poly(rI)·poly(rC) (Chamberlin & Patterson, 1965); poly(rA)·poly(rU) < poly(rA)·poly(dT) < poly(dA)·poly(dT) (Riley et al., 1966); and poly(rAdU) ≪ poly(dAT) < poly(rAU) (Chamberlin, 1969). Additionally, results reported for heptameric sequences indicate relative stabilities of $rC(A)_5G·rC(U)_5G$ < $rC(A)_5G·dC(T)_5G$ < $dC(A)_5G·dC(T)_5G$ (Martin & Tinoco, 1980). Notably, poly(dA)·poly(rU) and $dC(A)_5G·rC(U)_5G$ preferentially exist as the triple helix [Riley et al. (1966) and Martin and Tinoco (1980), respectively]. Many of these homopolymer duplexes are known to adopt peculiar structures, however, so that the relative stabilities of mixed duplexes are unpredictable. Together, all the data on homopolymers, alternating polymers, and mixed sequences indicate that additional information relating sequence of hybrid duplexes to thermodynamic parameters is needed.

An examination of the thermodynamic parameters derived from these pentamer experiments indicates that, relative to the RNA duplex, the heteroduplex formation has a more favorable entropy term and a less favorable enthalpy term (part a of Table I). The heteroduplexes resemble DNA in both the relative and absolute values of their enthalpy and entropy terms (part a of Table I) although the $\Delta G°$ values of the heteroduplexes are consistently a little lower than those of DNA ($\Delta\Delta G° = 0.5–1$ kcal/mol). Previous experiments with the heptamers $dC(A)_5G·dC(T)_5G$, $rC(A)_5G·dC(T)_5G$, and $rC(A)_5G·rC(U)_5G$ gave different relative thermodynamic parameters (Martin & Tinoco, 1980); in 0.2 M NaCl, the $\Delta H°$ values for DNA and RNA duplex formation were similar, while the $\Delta H°$ value for heteroduplex formation was slightly more favorable. Variations in sequence and solvent conditions could account for these differences in relative stabilities. The heptamer duplexes contain a central $A_5·T(U)_5$ tract which (in DNA) is known to induce modulation of the duplex structure by cooperative deflection (curvature) of the helix axis (Ulanovsky et al., 1986; Zahn & Blattner, 1987).

Local Structural Variation. Since base stacking contributes to a more favorable $\Delta H°$ value, conformations in which base–base interactions are optimized can be expected to exhibit more favorable enthalpies for helix formation than conformations which have less base overlap. There is more extensive base–base overlap in A-form conformations (Sanger, 1984), such that the $\Delta H°$ term for an A-form helix might be expected to be greater than that of a B-form helix, in the absence of other mitigating variables. Our data for the DNA and RNA pentamers support this. However, the heteroduplexes are also A-form helices, yet their enthalpies are unfavorable relative to the RNA (or even to the B-form DNA). Since base stacking should enhance stability of the heteroduplexes in the A-form, other compensating interactions must be present. One possible explanation is that destabilization may result from the presence of the 2′-OH in only one of the duplex strands. In A-form RNA, hydrogen bonds involving the 2′-OH have been suggested to participate in structure stabilization (Kim & Jhon, 1979).

Local structural variation in the DNA·RNA duplexes could also account for their less favorable enthalpy of duplex formation. Differences in the measured hyperchromicities of the strands offer some support for this hypothesis: the RNA duplex (3.3 mM) exhibits a 58% hyperchromic effect, while the DNA duplex (3.3 mM) shows only a 21% change. The heteroduplexes are intermediate: rCACAG·dCTGTG (2.7 mM) exhibits a 34% hyperchromicity, while rCUGUG·dCACAG (2.9 mM) shows only a 27% change. These hyperchromic data suggest that the hybrid structures do not adopt a perfect A-form structure with the corresponding base–base overlap. Other NMR data (imino proton data and NOESY experiments of the nonexchangeable protons) indicate that one terminal base-pair is apparently fraying, which results in inefficient or unobservable NOEs (data not shown). The instability of this terminal base-pair could also account for the lower observed hyperchromicity, as well as contribute to the unfavorable enthalpy of helix formation.

This terminal base-pair of the RNA duplex and the heteroduplexes, G_5C_6, is less stable than other base-pairs, as suggested by the loss of intensity of that imino proton resonance as a function of temperature, as well as by the NOE patterns. The G_5C_6 imino proton of the heteroduplexes is the first to exhibit temperature-dependent broadening. This base-pair is apparently flexible or transiently stacked over the adjacent $T(U)_4A_7$ base-pair, since the characteristic $A_7H2–G_5H1'$ NOE in the heteroduplexes was not observed in NOESY experiments, even with very long mixing times at lower temperatures (data not shown). In contrast, the other terminal base-pair C_1G_{10}, is stacked over the adjacent $(U)T_2A_9$, based upon both the characteristics of the imino proton resonance and the efficiency of the $A_9H2–G_{10}H1'$ NOE (although NMR experiments using the nonexchangeable protons indicate that the structures of the two heteroduplexes are not identical in this region; data not shown). The RNA·RNA A-form duplex does not suffer from any such apparent instability: each AH2 resonance shows NOEs to two H1′ resonances in NOESY spectra (data not shown). Thermodynamic parameters (Freier et al., 1986) for this pyrimidine/purine dimer, $^5U_4G_5/^3A_7C_6$, are not unusual. Measured in 1 M NaCl, the $\Delta H°$, $\Delta S°$, and $\Delta G°$ of propagation at a $^5UG/^3AC$ dimer are −10.5 kcal/mol, −27.8 eu, and −1.8 kcal/mol, respectively (Freier et al., 1986). These values can be compared to those of the other terminal dimer, $^5C_1U_2/^3AG$, which has measured values of $\Delta H° = -7.6$, $\Delta S° = -19.2$, and $\Delta G° = -1.7$. The internal U_2G_3/A_9C_8 dimer of the heteroduplexes is stably stacked as determined by the imino proton data and by NMR analysis of the nonexchangeable protons (data not shown), as predicted from the thermodynamic data. However, Calladine's rules (Calladine, 1982; Dickerson, 1983) predict a large base plane roll angle and twist angle for A-form helices at a pyrimidine/purine step to relieve the purine/purine minor groove clash. Helix distortion at this step with accompanying destabilization would be more severe at the end of a helix than in the middle where flanking base-pairs would anchor the structure. Why this effect is more severe in the hybrids than in the DNA·DNA or RNA·RNA duplexes is not understood [but see Calladine and Drew (1984) for a discussion of pyrimidine/purine steps and their relation to the A to B transition].

Hybrid Destabilization. For these hybrid pentamers, there is a net ΔG change of 1.5–2 kcal/mol, or 0.3–0.4 kcal/mol per base-pair compared to the RNA duplex. For other systems, many factors will affect the value of the relative $\Delta\Delta G$. The actual value of the destabilization per base-pair is expected to be sequence dependent, as Birnstiel et al. (1972) showed that with increased GC content the difference in the T_M of related RNA and hybrid duplexes decreased. Solution conditions can also affect the relative stabilities of these duplexes. While counterions may not alter the relative stabilities (Chamberlin & Berg, 1964), formamide was shown to stabilize RNA·DNA hybrid duplexes preferentially over the RNA·RNA or DNA·DNA duplexes (Birnstiel et al., 1972; Casey & Davidson, 1977). One explanation for this result is that DNA can more easily adopt an A-form conformation in the presence of denaturing solvents that lower water activity. This explanation is attractive because DNA is known to adopt an A-form conformation at low relative humidity (Tunis-Schneider & Maestre, 1970; Arnott & Selsing, 1974; Leslie et al., 1980). Lowering the water activity could approximate these conditions and facilitate the formation of the A-form hybrid in a sequence-dependent manner (Calladine & Drew, 1984).

Summary. The pentamer duplexes used for these experiments share common sequences. The relative stabilities RNA·RNA > DNA·DNA > RNA·DNA are identical to that shown previously for analogous duplexes of ϕX sequences but differ from many of the results obtained for homopolymer duplexes. These results may be relevant for certain biological systems as well as the process of antisense regulation. These results also demonstrate that DNA·RNA structures are

qualitatively and quantitatively different from DNA and RNA structures and warrant additional detailed study.

Acknowledgments

We thank Doug Turner, University of Rochester, for his hospitality and the generous use of his melting apparatus. The help of John SantaLucia and Adam Peritz was crucial for these measurements. We also thank Prof. Al Redfield, Brandeis University, for his support (National Institutes of Health GM20168) and interest, and for the use of his NMR spectrometer.

Supplementary Material Available

A plot of the absorbance vs temperature and of the calculated fit to the two-state model at a given concentration for each duplex (3 pages). Ordering information is given on any current masthead page.

References

Arnott, S. (1970) *Prog. Biophys. Mol. Biol. 21*, 265–319.
Arnott, S., & Selsing, E. (1974) *J. Mol. Biol. 88*, 551–552.
Beaucage, S. L., & Caruthers, M. H. (1981) *Tetrahedron Lett. 22*, 1859–1862.
Birnstiel, M. L., Sells, B. H., & Purdom, I. F. (1972) *J. Mol. Biol. 63*, 21–39.
Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843–853.
Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746–3750.
Calladine, C. R. (1982) *J. Mol. Biol. 161*, 343–352.
Calladine, C. R., & Drew, H. R. (1984) *J. Mol. Biol. 178*, 773–782.
Casey, J., & Davidson, N. (1977) *Nucleic Acids Res. 4*, 1539–1552.
Chamberlin, M. J. (1969) *Proced. Nucleic Acid Res.* 513–519.
Chamberlin, M. J., & Berg, P. (1964) *J. Mol. Biol. 8*, 297–313.
Chamberlin, M. J., Baldwin, R. L., & Berg, P. (1963) *J. Mol. Biol. 7*, 334–349.
Chou, S.-H., Flynn, P., & Reid, B. (1989) *Biochemistry 28*, 2435–2443.
Clore, G. M., Gronenborn, A., & McLaughlin, L. M. (1985) *Eur. J. Biochem. 151*, 153–165.
Cohen, J. S. (1989) *Oligonucleotides, Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL.
Dickerson, R. E. (1984) *J. Mol. Biol. 166*, 419–441.
Drew, H., Takano, T., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature 286*, 567–573.
Farnham, P. J., & Platt, T. (1980) *Cell 20*, 739–748.
Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., & Turner, D. H. (1983a) *Biochemistry 22*, 1698–6206.
Freier, S. M., Petersheim, M., Hickey, D. R., & Turner, D. H. (1983b) *Biopolymers 22*, 1107–1131.
Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373–9377.
Hall, K. B., Cruz, P. C., van de Sande, J., Jovin, T. M., & Tinoco, I., Jr. (1984) *Nature 311*, 584–586.
Kim, K., & Jhon, M. S. (1979) *Biochim. Biophys. Acta 565*, 131–147.
Kornberg, A. (1980) *DNA Replication*, W. H. Freeman.
Leslie, A. G. W., Arnott, S., Chandrasekaran, R., & Ratliff, R. L. (1980) *J. Mol. Biol. 143*, 49–72.
Martin, F. H., & Tinoco, I., Jr. (1980) *Nucleic Acids Res. 8*, 2295–2299.
McLaughlin, L. W., & Piel, N. (1984) in *Oligodeoxynucleotide Synthesis: A Practical Approach* (Gait, M. J., Ed.) pp 199–218, IRL Press, Oxford.
Mellema, J.-R., Haasnoot, C. A. G., van der Marel, G. A., Wille, G., van Boeckel, C. A. A., van Boom, J. H., & Altona, C. (1983) *Nucleic Acids Res. 11*, 5717–5738.
Milman, G., Langridge, R., & Chamberlin, M. J. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 1804–1810.
Nelson, J. N., & Tinoco, I., Jr. (1982) *Biochemistry 21*, 5289–5295.
Pardi, A., Martin, F. H., & Tinoco, I. Jr. (1981) *Biochemistry 20*, 3986–3996.
Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 256–263.
Pohl, F., & Jovin, T. M. (1972) *J. Mol. Biol. 67*, 375–396.
Reid, D. G., Salisbury, S. A., Brown, T., Williams, D., Vasseur, J.-J., Rayner, R. B., & Imbach, J.-L. (1983) *Eur. J. Biochem. 135*, 307–314.
Richards, E. G. (1975) *Handb. Biochem. Mol. Biol., 3rd Ed. I.*
Riley, M., Maling, B., & Chamberlin, M. J. (1966) *J. Mol. Biol. 20*, 359–389.
SantaLucia, J., Kierzek, R., & Turner, D. H. (1991) *J. Am. Chem. Soc. 113*, 4313–4322.
Ts'o, P. O. P., Ed. (1974) *Basic Principles in Nucleic Acid Chemistry*, Academic Press, New York.
Tunis-Schneider, M. J., & Maestre, M. F. (1970) *J. Mol. Biol. 52*, 521–541.
Ulanovsky, L., Bodner, M., Trifonov, E. N., & Chodev, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 862–866.
van der Marel, G. A., van Boeckel, C. A. A., Wille, B., & van Boom, J. H. (1982) *Nucleic Acids Res. 10*, 2337–2351.
Wreesman, C. T. J., Fidder, A., van der Marel, G. A., & van Boom, J. H. (1983) *Nucleic Acids Res. 11*, 8389–8405.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Yager, T. D., & von Hippel, P. H. (1987) in *Escherichia coli and Salmonella typhimurium*, Vol. 2, pp 1241–1275, American Society for Microbiology, Washington, D.C.
Zahn, K., & Blattner, F. R. (1987) *Science 236*, 416–422.
Zimmerman, S. B., & Pheiffer, B. H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 78–82.