Base-Stacking and Base-Pairing Contributions to Helix Stability: Thermodynamics of Double-Helix Formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp†

Matthew Petersheim‡ and Douglas H. Turner*

ABSTRACT: The thermodynamics of double-helix formation in 1 M NaCl have been measured spectrophotometrically for CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. The results indicate additional double-helical stability is conferred by the terminal unpaired bases. The 3′ A stabilizes the double helix more than the 5′ A or the 3′ U. The increased stability is due to a more favorable enthalpy change for double-helix formation. Comparison of the thermodynamics for CCGG, ACCGGp, CCGGUp, and ACCGGUp indicates stacking interactions are somewhat more important than pairing interactions in determining the stability of the terminal AU base pairs in ACCGGUp.

Due to the development of rapid sequencing techniques, there has been an explosion in our knowledge of nucleic acid sequences (Maxam & Gilbert, 1977; Sanger & Coulson, 1975; Sanger et al., 1977; Wu, 1978; Orcutt et al., 1982). This knowledge provides a foundation for understanding the functions and mechanisms of these macromolecules. However, such understanding requires information about the structures of the molecules, which is still difficult to acquire. In principle, it is possible to predict structure from primary sequence. Progress has been made in predicting both secondary structure (Borer et al., 1974; Tinoco et al., 1973) and local conformation (Levitt, 1978; Hogan et al., 1978; Kollman et al., 1981). However, there are large gaps in our knowledge of the interactions that are important to consider in making these predictions.

Forces thought to be important for nucleic acid structure include electrostatics, solvent, base stacking, and base pairing (Cantor & Schimmel, 1980; Bloomfield et al., 1974). Electrostatic contributions are reasonably well understood, both theoretically and experimentally (Manning, 1978; Record et al., 1978). Recent work from our laboratory suggests bulk solvent contributions are relatively unimportant (Turner et al., 1981; Freier et al., 1981; Albergo & Turner, 1981; Dewey & Turner, 1980). The contributions of base stacking and pairing are uncertain. In this paper, we report the thermodynamics of double-helix formation for CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp, where p denotes a 3′-terminal phosphate. The results provide insight into the relative importance of base stacking and pairing in determining the stability of double helices.

The molecules studied all contain the double-helical core CCGG. This sequence was chosen because previous NMR[1] studies indicate all four base pairs melt cooperatively (Arter et al., 1974). Our NMR studies indicate similar behavior for the other oligomers (Petersheim & Turner, 1983). Thus a two-state model can be used to derive the thermodynamics. Three of the oligonucleotides contain terminal unpaired bases ("dangling ends"). Such bases are known to stabilize double helices, presumably by stacking (Martin et al., 1971; Romaniuk et al., 1978; England & Neilson, 1976; Neilson et al., 1980; Alkema et al., 1981a,b). However, the only previous thermodynamic comparison of a particular sequence with and without dangling ends indicated the greater stability was due to an entropy term (Martin et al., 1971). Stacking is commonly thought to be associated with a favorable enthalpy term, although the magnitude is controversial (Cantor & Schimmel, 1980; Freier et al., 1981; Leng & Felsenfeld, 1966; Stannard & Felsenfeld, 1975; Applequist & Damle, 1966; Neumann & Ackermann, 1969; Breslauer & Sturtevant, 1977; Filimonov & Privalov, 1978; Rawitscher et al., 1963; Suurkuusk et al., 1977; Brahms et al., 1967; Pörschke, 1976). The previous thermodynamic measurements on ribooligonucleotide dangling ends involved molecules of the type A_nU_m (Martin et al., 1971). When $n \neq m$, there may be more than one way to form the maximum number of base pairs in these duplexes. This could give rise to an unusual entropy effect. Moreover, it is unlikely that all the base pairs in A_nU_m melt cooperatively. The CCGG series does not have these ambiguities, and our results indicate terminal base stacking is driven by a favorable enthalpy.

Many biological processes require base pairing of short RNA sequences. For example, the structures of tRNA, mRNA, and rRNA, the associations of mRNA with tRNA and rRNA, and possibly mRNA splicing depend on formation of double helices (Cantor & Schimmel, 1980; Steitz & Jakes, 1975; Oxender et al., 1979; Crawford & Stauffer, 1980; Lerner et al., 1980). These double helices usually terminate with unpaired bases. For the tRNA–mRNA association, it has been suggested that these unpaired bases are crucial to the stability of the complex (Grosjean et al., 1976; Yoon et al., 1976). Presumably this is also true for many of the other associations. The thermodynamics reported here give further insight into these stabilization effects.

Experimental Procedures

Oligonucleotide Synthesis. CCGGp, CCGG, and CCGGAp were originally synthesized enzymatically from CC (Sigma) by methods described previously (Thach, 1966;

† From the Department of Chemistry, University of Rochester, Rochester, New York 14627. *Received June 24, 1982.* This work was supported by National Institutes of Health Grant GM22939. D.H.T. is an Alfred P. Sloan Fellow. M.P. was a National Institutes of Health Predoctoral Trainee (Grant 5T32 GM 07230).

‡ Present address: Department of Chemistry, Yale University, New Haven, CT 06520.

[1] Abbreviations: GDP, guanosine 5′-diphosphate; PNPase P, primer-dependent polynucleotide phosphorylase; RNase T1, ribonuclease T1; Tris, tris(hydroxymethyl)aminomethane; BAPase, bacterial alkaline phosphatase; pNp, nucleoside 3′,5′-bisphosphate; HPLC, high-performance liquid chromatography; DEAE, diethylaminoethyl; TEAB, triethylammonium bicarbonate; EDTA, ethylenediaminetetraacetate; BSA, bovine serum albumin; CD, circular dichroism; NMR, nuclear magnetic resonance.

0006-2960/83/0422-0256$01.50/0

Martin, 1969; Borer, 1972; Borer et al., 1973; Uhlenbeck & Cameron, 1977; England & Uhlenbeck, 1978). The trimer CCGp was obtained by incubating 4 mM CC and 40 mM GDP with 8 units/mL PNPase P (P-L Biochemicals) and 300 units/mL RNase T1 at 37 °C for 1–5 days in 0.5 M NaCl, 5–10 mM $MgCl_2$, and 50 mM Tris-HCl, pH 8.2. CCG was obtained by incubating ~1 mM CCGp with 50 units/mL BAPase (Millipore, BAPC) for 1–2 h. Subsequent additions were made by using T4 RNA ligase (P-L Biochemicals) to add a nucleoside 3′,5′-bisphosphate (pNp) to the free 3′-hydroxyl of an acceptor oligomer. Terminal phosphates were removed from intermediate oligomers with BAPase. Oligomers were then purified prior to further addition. Typical T4 RNA ligase reaction mixtures contained ≤1 mM acceptor (e.g., CCG), 4–8 mM pNp, 8–16 mM ATP, and 20–200 units/mL T4 RNA ligase in 20 mM $MgCl_2$, 10 mM dithiothreitol, 20 μg/mL BSA, and 200 mM Tris-HCl, pH 8.5. Reactions were incubated at 37 °C for 1–48 h. The reactions and product purity were monitored by HPLC on a reverse-phase column (Altex Ultrasphere-Octyl, 5 μm, 0.46 × 14 cm with Bownlee RP-8 MPLC guard column) with a methanol gradient. The gradient was generated from 0.1 M NaH_2PO_4 in water, pH 4, and a 1:1 (v/v) mixture of 0.1 M NaH_2PO_4 in water with absolute methanol (Burdick & Jackson, distilled in glass). Apparent impurities were less than 1%.

CCGG was also obtained from Collaborative Research and had properties identical with that synthesized from CC. The commercial tetramer was used to make CCGGAp and CCGGUp by the T4 RNA ligase reaction. ACCGGp and ACCGGUp were prepared from ACC (Sigma) by successive application of the T4 RNA ligase and BAPase reactions.

Anion-exchange chromatography with DEAE-Sephadex (A-25, 120 mesh, 40–120-μm particles) was used to purify the starting materials, CC, ACC, and CCGG, the intermediate product oligomers after reaction with BAPase, and the final products. Gradients contained either NaCl/7 M urea or TEAB. When NaCl was used, the oligomer was desalted on a DEAE-Sephadex column with TEAB. The TEAB and water were removed under vacuum. Addition of methanol after most or all of the water was evaporated accelerated the decomposition of TEAB. The above procedures have been described in greater detail by Petersheim (1982).

Extinction Coefficients. Extinction coefficients were calculated with the nearest-neighbor approximation (Cantor & Tinoco, 1965; Fasman, 1975). It was also assumed that $\epsilon_{280}/\epsilon_{260}$ for GpA is the same as that for ApG and that for GGC and GGU, $\epsilon_{280}/\epsilon_{260}$ is an average of the ratios for the component dimers. The latter was used to estimate ϵ_{280} for GG. For CCGG, the extinctions at 260 nm measured at 70 °C (Arter et al., 1974) and calculated are 34×10^3 and 34.9×10^3 cm^{-1} M^{-1} (strand), respectively. The calculated extinction coefficients ($\times 10^{-3}$ cm^{-1} M^{-1}) at 280 nm for the oligomers are CCGG (24.4), ACCGGp (26.3), CCGGAp (26.7), CCGGUp (27.4), and ACCGGUp (29.3).

Melting Curves. Absorbance vs. temperature curves were measured at 280 nm on a Gilford 250 spectrometer and PDP 11/34 computer, as previously described (Albergo et al., 1981; Petersheim, 1982). The buffer was 1.0 M NaCl, 1 mM EDTA, and 10 mM sodium cacodylate, pH 7.0. Absorbances were not corrected for thermal expansion since the correction was linear and 3% or less from 0 to 80 °C.

Curve Fitting. All the oligomers are self-complementary, and the cooperative part of the melting curves reflects the single-strand to double-helix equilibrium:

$$2A \rightleftarrows A_2 \qquad (1)$$

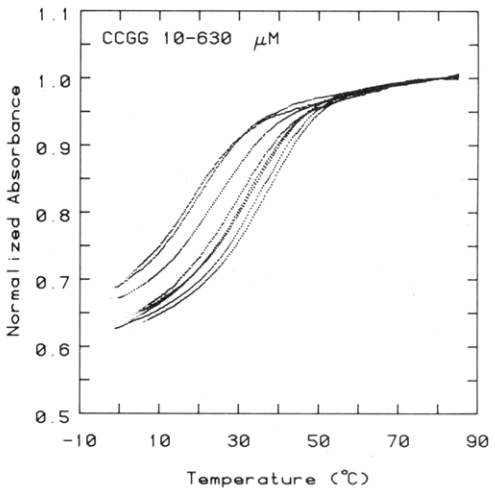

FIGURE 1: Normalized absorbance at 280 nm vs. temperature for CCGG. Absorbances were normalized by dividing by the absorbance at 80 °C. Oligomer strand concentrations range from 10 to 630 μM. All solutions contained 1.0 M NaCl, 1 mM EDTA, and 10 mM cacodylate, pH 7.0.

The Marquardt nonlinear least-squares method (Bevington, 1969) was used to fit the melting curves with the equation

$$\epsilon(T) = A(T)/(lc_T) = \alpha\epsilon_{ds} + (1 - \alpha)\epsilon_{ss} \qquad (2)$$

where $\epsilon(T)$ and $A(T)$ are the extinction coefficient and absorbance of the solution, respectively, at temperature T, l is the path length, c_T is the total strand concentration, and ϵ_{ds} and ϵ_{ss} are extinction coefficients of the double- and single-strand species, respectively, and are assumed to be linear functions of temperature:

$$\epsilon_{ds} = m_{ds}T + b_{ds} \qquad (3)$$

$$\epsilon_{ss} = m_{ss}T + b_{ss} \qquad (4)$$

α is the fraction of strands in the double-strand state and is related to the changes in enthalpy, $\Delta H°$, and entropy, $\Delta S°$, for the reaction by

$$K = \frac{\alpha}{2(1 - \alpha)^2 c_T} = \exp\left(-\frac{\Delta H°}{RT} + \frac{\Delta S°}{R}\right) \qquad (5)$$

The total strand concentration, c_T, is determined from the absorbance at 70 °C (80 °C for ACCGGUp). The program fits the experimental melting curve treating $\Delta H°$, $\Delta S°$, m_{ds}, b_{ds}, m_{ss}, and b_{ss} as variable parameters. In general, the root-mean-square difference between the data and calculated curve is less than 0.5%, the approximate error in the absorbance readings.

The above treatment assumes that a linear approximation is adequate for the temperature dependences of the single- and double-strand extinctions. The assumption for single strands is consistent with absorption vs. temperature experiments on 7×10^{-7} M CCGGp in a 10-cm cell and with previous studies of single-stranded oligomer melting (Martin, 1969; Nelson, 1981; Adler et al., 1967; Brahms et al., 1967). The assumption for double strands is supported by the low-temperature shapes of melting curves for oligomers with high melting temperatures [see Figure 2 and Breslauer et al. (1975) and Albergo et al. (1981)].

Results

Figures 1 and 2 show melting curves for CCGG and ACCGGUp, the least and most stable oligomers, respectively. Several methods have been used to derive thermodynamics from the melting curves. All are based on a two-state model with linearly sloping base lines (Breslauer et al., 1975; Albergo et al., 1981).

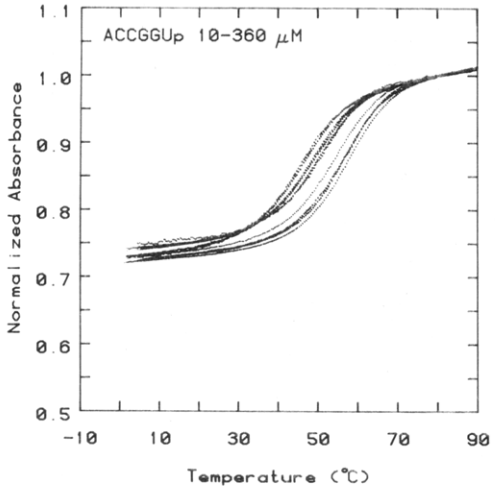

FIGURE 2: Normalized absorbance at 280 nm vs. temperature for ACCGGUp. Absorbances were normalized by dividing by the absorbance at 80 °C. Oligomer strand concentrations range from 10 to 360 μM. All solutions contained 1.0 M NaCl, 1 mM EDTA, and 10 mM cacodylate, pH 7.0.

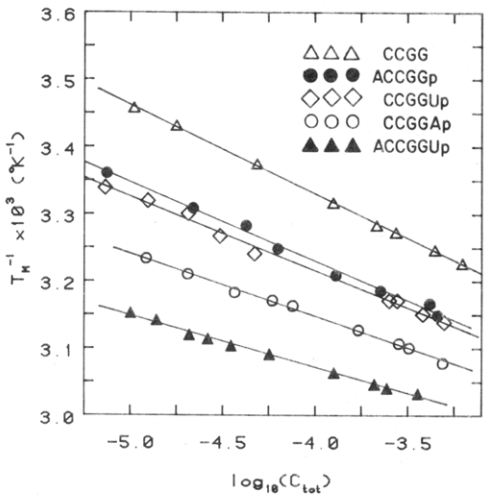

FIGURE 3: Plots of T_m^{-1} vs. log c_T for CCGG (△), ACCGGp (●), CCGGUp (◇), CCGGAp (○), and ACCGGUp (▲). The lines are least-squares fits of the data. All solutions contained 1.0 M NaCl, 1 mM EDTA, and 10 mM cacodylate, pH 7.0.

One way to obtain thermodynamics from melting curves is from the concentration dependence of the melting temperature, T_m (Borer et al., 1974):

$$1/T_m = (2.3R/\Delta H°) \log c_T + \Delta S°/\Delta H° \quad (6)$$

T_m is defined as the temperature where the fraction of strands in double helix is 0.5. For each concentration, T_m is determined from the parameters derived from fitting the melting curve (see Experimental Procedures). Figure 3 contains plots of T_m^{-1} vs. log c_T. The plots are linear as expected for two-state behavior. The derived thermodynamics are listed in Table I. Alternatively, log c_T could be plotted vs. T_m^{-1}. The latter treatment is appropriate if errors in log c_T are larger than errors in T_m^{-1}. Plots of log c_T vs. T_m^{-1} yield thermodynamic parameters within 1% of those from T_m^{-1} vs. log c_T, suggesting the errors in log c_T and T_m^{-1} are comparable.

A second method for deriving the thermodynamics is to average the $\Delta H°$ and $\Delta S°$ parameters obtained from fitting the individual melting curves. These values are also listed in Table I.

The above treatments assume the thermodynamics of helix formation are temperature independent. However, the fitted

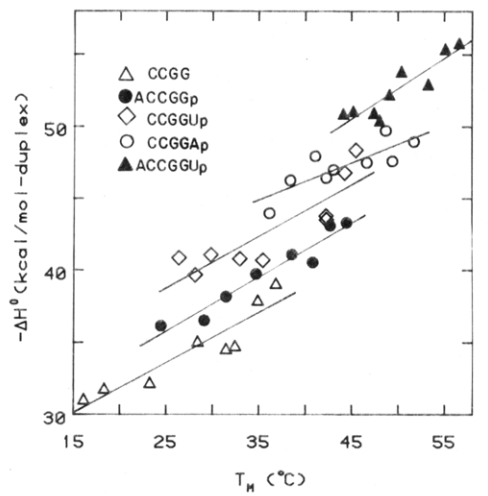

FIGURE 4: Plots of $\Delta H°$ vs. T_m for CCGG (△), ACCGGp (●), CCGGUp (◇), CCGGAp (○), and ACCGGUp (▲). The solid lines are least-squares fits of the data for a given oligomer.

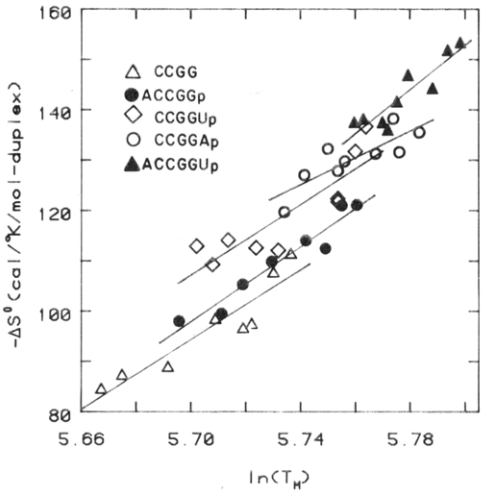

FIGURE 5: Plots of $\Delta S°$ vs. ln T_m for CCGG (△), ACCGGp (●), CCGGUp (◇), CCGGAp (○), and ACCGGUp (▲). The solid lines are least-squares fits of the data for a given oligomer.

thermodynamics are not random with respect to T_m. It would not be surprising if $\Delta H°$ and $\Delta S°$ are temperature dependent, since the sloping base lines probably reflect temperature-dependent double- and single-stranded states. For example, the degree of stacking in single strands is temperature dependent (Pörschke et al., 1973; Appleby & Kallenbach, 1973). In general, the temperature dependence of ΔH and ΔS is given by (Lewis & Randall, 1961)

$$d\Delta H/dT = \Delta C_p \quad (7)$$

and

$$d\Delta S/d \ln T = \Delta C_p \quad (8)$$

Here ΔC_p is the difference in heat capacity between single and double strands. If one assumes the $\Delta H°$ and $\Delta S°$ from a fit are approximately the values for the transition at the T_m (Petersheim, 1982), plots suggested by eq 7 and 8 are shown in Figures 4 and 5. Within the scatter of the data, the plots are linear, suggesting ΔC_p is approximately constant over the range of T_m's for a given oligomer. Thus the data have been fit to the following equations:

$$\Delta H°(T_m) = \Delta H°(T_0) + \Delta C_p°(T_m - T_0) \quad (9)$$

$$\Delta S°(T_m) = \Delta S°(T_0) + \Delta C_p° \ln (T_m/T_0) \quad (10)$$

Here T_0 is an arbitrary reference temperature. The slopes of

Table I: Thermodynamics for Single-Strand to Double-Helix Transitions

method	eq	parameter	units	CCGG	ACCGGp	CCGGUp	ACCGGUp	CCGGAp	CCGGp
T_m^{-1} vs. log c_T	6	$\Delta H°$	kcal/mol of duplex	−34.2 ± 0.2	−38.6 ± 0.4	−41.4 ± 0.2	−59.5 ± 0.6	−48.9 ± 0.4	−31.0
	6	$\Delta S°$	cal/(K mol of duplex)	−95.6 ± 0.8	−106.4 ± 1	−114.9 ± 0.5	−164.5 ± 2	−135.6 ± 1	−86.5
		$\Delta G°$(39 °C)	kcal/mol of duplex	−4.36 ± 0.2	−5.39 ± 0.08	−5.56 ± 0.1	−8.17 ± 0.6	−6.58 ± 0.2	−4.02
		T_m(10^{-4} M strands)	°C	27.1	36.4	37.8	52.4	44.6	22.8
average of fitted parameters	5	$\Delta H°$	kcal/mol of duplex	−34.7 ± 3	−39.7 ± 3	−42.8 ± 3	−53.2 ± 3	−47.3 ± 2	−33.0
	5	$\Delta S°$	cal/(K mol of duplex)	−97.1 ± 10	−109.9 ± 9	−119.3 ± 10	−144.9 ± 9	−130.4 ± 5	−93.1
		$\Delta G°$(39 °C)	kcal/mol of duplex	−4.41 ± 0.1	−5.44 ± 0.09	−5.60 ± 0.08	−7.98 ± 0.08	−6.57 ± 0.06	−3.92
$\Delta H°$ vs. T_m	7	$\Delta H°$(39 °C)	kcal/mol of duplex	−39.0 ± 1	−41.0 ± 0.7	−43.8 ± 2	−48.0 ± 0.9	−45.9 ± 1	
$\Delta S°$ vs. ln T_m	8	$\Delta S°$(39 °C)	cal/(K mol of duplex)	−111.2 ± 4	−114.1 ± 2	−122.4 ± 5	−127.9 ± 3	−126.1 ± 3	
		$\Delta G°$(39 °C)	kcal/mol of duplex	−4.32 ± 0.04	−5.41 ± 0.09	−5.59 ± 0.08	−8.09 ± 0.05	−6.58 ± 0.04	

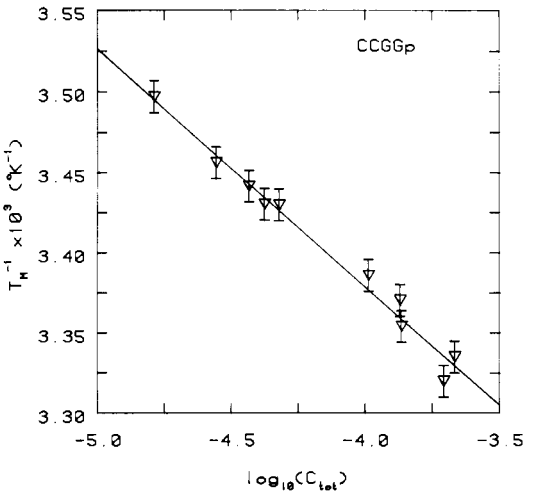

FIGURE 6: Plot of T_m^{-1} vs. log c_T for CCGGp. In the determination of T_m's, m_{ds} in eq 3 was fixed at 20 (cm K mol of strand)$^{-1}$. All solutions contained 1.0 M NaCl, 1 mM EDTA, and 10 mM cacodylate, pH 7.0.

the fitted lines give $\Delta C_p°$. For a given oligomer, the $\Delta C_p°$'s derived from $\Delta H°$ and $\Delta S°$ data are identical within 5%. The average derived $\Delta C_p°$'s for CCGG, ACCGGp, CCGGUp, ACCGGUp, and CCGGAp are respectively −382, −393, −355, −434, and −263 cal K^{-1} (mol of duplex)$^{-1}$. These magnitudes are consistent with heat capacities expected from the temperature dependence of single-strand stacking (Freier et al., 1981) and with the heat capacity measured for poly(A)–poly(U) duplex formation (Rawitscher et al., 1963; Suurkuusk et al., 1977).

The apparent temperature dependence of the fitted thermodynamics suggests the single- to double-strand transition is not pure two state. However, two-state fits to a given melting curve are excellent, and T_m^{-1} vs. log c_T plots are linear. Presumably, this is due to the narrow temperature ranges for each curve and for the log c_T plots. The assumption of a temperature-independent, nonzero $\Delta C_p°$ in eq 9 and 10 is a first-order correction for the two-state model that incorporates the observed temperature dependence. In principle, more complex treatments with additional adjustable parameters are possible (Appleby & Kallenbach, 1973). However, the modified two-state model appears adequate for the data presented here.

In the work described above, we have determined the thermodynamics of double-helix formation in 1 M NaCl for CCGG, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. The tetramer, pentamers, and hexamer have three, five, and six phosphates, respectively. In comparing the thermodynamics of the different oligomers, it would be useful to know if an additional phosphate on CCGG alters the thermodynamics significantly. In earlier work, we measured the thermodynamics for CCGGp. The errors in these measurements were somewhat larger than the errors in the experiments reported above. Moreover, there was less lower base line available than for any of the other oligomers. As a result of insufficient data at lower temperatures, the standard fitting program often gave negative slopes for the lower base line. Therefore, we fixed the lower base-line slope at 20 (cm K M strand)$^{-1}$, which is approximately the average slope for the other oligomers with terminal phosphates. The average $\Delta H°$ and $\Delta S°$ from these fits are −33.0 kcal/mol of duplex and −93.1 cal/(K mol of duplex), respectively (see Figure 6 and Table I). When the lower base line is a variable parameter, the average $\Delta H°$ and $\Delta S°$ from the fits are −30.8 kcal/mol

Table II: Differences in Thermodynamics between Various Oligomers and CCGG per Mole of *Strand*

method	parameter	units	ACCGGp	CCGGUp	ACCGGUp	CCGGAp
T_m^{-1} vs. log c_T	$\Delta\Delta H^\circ$	kcal/mol of strand	−2.2	−3.6	−12.7	−7.4
	$\Delta\Delta S^\circ$	cal/(K mol of strand)	−5.4	−9.7	−34.5	−20.0
	$\Delta\Delta G^\circ$(39 °C)	kcal/mol of strand	−0.52	−0.60	−1.91	−1.11
average of fitted parameters	$\Delta\Delta H^\circ$	kcal/mol of strand	−2.5	−4.1	−9.3	−6.3
	$\Delta\Delta S^\circ$	cal/(K mol of strand)	−6.4	−11.1	−23.9	−16.7
	$\Delta\Delta G^\circ$(39 °C)	kcal/mol of strand	−0.52	−0.60	−1.79	−1.08
ΔH° vs. T_m	$\Delta\Delta H^\circ$(39 °C)	kcal/mol of strand	−1.0	−2.4	−4.5	−3.5
ΔS° vs. ln T_m	$\Delta\Delta S^\circ$(39 °C)	cal/(K mol of strand)	−1.5	−5.6	−8.4	−7.5
	$\Delta\Delta G^\circ$(39 °C)	kcal/mol of strand	−0.55	−0.64	−1.89	−1.13

of duplex and −86.2 cal/(K mol of duplex), respectively. Moreover, with other oligomers, changing the base-line slope by ±20 (cm K M strand)$^{-1}$ changes the thermodynamics by 10% or less. Therefore, although a detailed comparison with the other oligomers is not possible, a qualitative comparison is justified. The results in Table I indicate that in 1 M NaCl, the additional phosphate group induces no large thermodynamic effect.

Discussion

The goals of this work are (1) to determine the effect of dangling ends and terminal base pairs on helix stability, (2) to resolve the thermodynamics of these effects, and (3) to estimate the relative contributions of base stacking and base pairing to helix stability. One concern in deriving thermodynamics from melting curves is that aggregation may affect the results (Nelson et al., 1981). NMR measurements indicate aggregation is not important for CCGGAp, ACCGGp, CCGGUp, and ACCGGUp since proton chemical shifts and line widths show no inordinate changes at low temperatures (Petersheim & Turner, 1983). The transition temperatures of CCGG and CCGGp are too low to use NMR to assay aggregation. A second assay for aggregation is circular dichroism. For example, the temperature dependences of the CD spectra of both GGCC and (dG-dC)$_3$ are anomalous under conditions where aggregation is present (Freier et al., 1983; S. M. Freier, unpublished results). No anomalous temperature dependence of CD spectra is observed with CCGG and CCGGp at concentrations of 4×10^{-4} and 2×10^{-4} M, respectively (M. Petersheim, unpublished results). Moreover, if aggregation were important, it should be inhibited by the 3′ terminal phosphate in CCGGp. The thermodynamics for CCGG and CCGGp reported in Table I are very similar. Thus there is no indication that aggregation affects the thermodynamic results reported here.

The melting curves have been analyzed by using a two-state model with linearly sloping base lines. There are four reasons for considering this an adequate model. First, analysis of optical melting curves for (dG-dC)$_3$ using this model gave enthalpies in good agreement with those obtained from microcalorimetry (Albergo et al., 1981; Freier et al., 1983). Second, individual melting curves are fit very well by the model. Third, the thermodynamic parameters are relatively insensitive to how the two-state approximation is applied. That is, T_m^{-1} vs. log c_T gives the same result as the average of the curve fitting parameters within 11% at worst and better than 5% in general (see Table I). Fourth, NMR melting curves indicate that all the base pairs in a given oligomer break at essentially the same temperature (Arter et al., 1974; Petersheim & Turner, 1983) and that there are only minor deviations from two-state behavior. Thus the two-state analysis appears reasonable.

One observation that is not consistent with the pure two-state model is the apparent temperature dependence of ΔH° and ΔS° shown in Figures 4 and 5. If this is real, then it is probably best to compare the thermodynamics of the oligomers at a single temperature. This will reduce differences due to the varying melting temperatures of the oligomers. The plots in Figures 4 and 5 have been used to estimate the thermodynamics for all the oligomers at 39 °C. This is the average T_m for all the data and is also midway between the T_m's for the tetramer and hexamer. Thus, extrapolation is minimized. The thermodynamics are listed in Table I.

The error limits in Table I for ΔH° and ΔS° derived from T_m^{-1} vs. log c_T data reflect errors in the slopes and intercepts (Meyer, 1975). The errors for the other ΔH°'s and ΔS°'s are standard deviations. Because the correlation factor between ΔH° and ΔS° is greater than 0.99, ΔG° is a more accurate parameter than either ΔH° or ΔS° individually. The error estimates for ΔG° include this correlation (Snedecor & Cochran, 1971).

The temperature-independent thermodynamics for CCGG and ACCGGUp in Table I can be compared with the literature. For CCGG, our ΔH° of −34.0 kcal/mol of duplex is in excellent agreement with the −33 kcal/mol of duplex measured by Arter et al. (1974). However, it is 16% lower than the ΔH° of −40.4 kcal/mol of duplex predicted from the thermodynamics of other oligomers containing GC base pairs (Borer et al., 1974). For ACCGGUp, the measured and predicted ΔH° values of, respectively, −56.1 and −52.2 kcal/mol of duplex and ΔS° values of −153.8 and −135.3 cal/(K mol of duplex) are in good agreement.

The difference of the thermodynamics for ACCGGUp and CCGG provides thermodynamics for the addition of an AU base pair to a CG base pair. These are listed in Table II. The ΔH° and ΔS° are very dependent on the method used to derive the thermodynamics. None of the methods are in agreement with the predicted values of −5.9 kcal/mol (AU base pair stacked on GC base pair) and −12.7 cal/K mol (AU base pair stacked on GC base pair) for ΔH° and ΔS°, respectively. Nevertheless, the free energy change at 39 °C, ΔG°(39 °C), is −1.9 kcal/mol of duplex for temperature-independent, temperature-dependent, and predicted thermodynamics. ΔG° is the parameter used for predicting stability of base-paired regions at a given temperature (Tinoco et al., 1973). The results suggest ΔG° will be reasonably accurate when calculated near the melting temperatures of the model oligonucleotides. However, long extrapolations to other temperatures will be more difficult.

Table II also contains the thermodynamics associated with addition of single, unpaired A and U residues to the ends of the CCGG duplex. The qualitative trends in the table are not dependent on the method used to derive the thermodynamics. These trends lead to several conclusions that are discussed in detail below.

Dangling Bases Stabilize the Double Helix. The free energy changes for helix formation with the pentamers are all more favorable than those with CCGG. This is consistent with the

literature (Martin et al., 1971; Alkema et al., 1981a,b; Neilson et al., 1980; Romaniuk et al., 1978). Neilson's group, in particular, has done extensive work on the effects of dangling ends. In all cases but two, a dangling base stabilizes the helix. One exception is AAGCU vs. AGCU (Alkema et al., 1981a,b). In this study, concentrations corresponding to the T_m's are not reported, and aggregation may be important. The second exception is GAGC/AGCUC vs. GAGC/GCUC (Neilson et al., 1980). In this system, AGCUC can form a self-complementary duplex thus complicating the T_m measurement. Thus, these two cases may not be exceptions.

3′ Adenosine Stabilizes More Than 3′ Uridine. At 10^{-4} M, the calculated T_m's for CCGG, CCGGUp, and CCGGAp are 27, 38, and 45 °C, respectively. A similar effect has been observed by Neilson et al. (1980) with NMR. For CAUG, CAUGU, and CAUGA, they measure T_m's of 25, 30, and 36 °C, respectively. This effect may be due to the larger size of adenine, which provides greater potential for cross-strand stacking. Alternatively, the electronic structure of uracil may make stacking less favorable. Single-stranded poly(U) has a random-coil conformation above 10 °C, indicating weak stacking interactions (Young & Kallenbach, 1978; Lipsett, 1960; Richards et al., 1963; Gukovsky et al., 1981). However, the stacking properties of uracil in single strands of heterogeneous sequence are not as well understood (Lee & Tinoco, 1980; Stone & Borer, 1981a,b; Stone et al., 1981).

3′ Adenosine Stabilizes More Than 5′ Adenosine. At 10^{-4} M, the calculated T_m of CCGGAp is 8 °C higher than that of ACCGGp. The difference in stabilization could be due to the positioning of the dangling end, to nearest-neighbor sequence effects, or to the terminal phosphates. However, Neilson's group has observed the following ordering in T_m's measured by NMR: GCA ≫ AGC, AGCUA > AAGCU, CAUGA > ACAUG (Neilson et al., 1980; Alkema et al., 1981a,b). The combination of results suggests the effect is likely due to the 3′ vs. 5′ positioning of the adenosine.

The primary difference between the 3′ adenosine and the 5′ adenosine is the connection of the ribose group to the helix. The 3′ adenosine is anchored to the sugar–phosphate chain by its 5′ carbon; the 5′ adenosine by its 3′ carbon. Greater stabilization could occur if the 3′ adenine is more constrained to a favorable position by the connection of its ribose than is the 5′ adenine, or the 5′ adenine is restricted to less favorable interactions by its ribose connection. NMR experiments indicate that as temperature is increased, the 3′ ribose is the first to undergo a structural change (Petersheim & Turner, 1983). This implies a less restricted ribose. This suggests the 3′ connection of the 5′ adenosine in some way restricts the cross-strand stacking. A separate possibility is that 5′ and 3′ dangling ends affect the core helix differently. NMR results indicate some alteration in the conformational flexibility of the core ribose groups of ACCGGp (Petersheim & Turner, 1983).

Stabilization of CCGG by Dangling Ends Is Enthalpic. The additional favorable free energy of duplex formation for the pentamers is associated with an additional favorable enthalpy change. The entropy change is more unfavorable for the pentamers than for CCGG. The only directly comparable previous study of the thermodynamics of dangling ends involved the sequences A_nU_m (Martin et al., 1971). In that case, the enhanced stabilization due to dangling bases was associated with a more favorable $\Delta S°$ of helix formation. It may be that dangling ends interact differently with GC and AU base pairs. It is more likely, however, that the A_nU_m results reflect a deviation from two-state behavior. Multiple double-helical states could involve bulging out or intercalation of interior bases or fraying of terminal base pairs. The results suggest A_nU_m molecules may have interesting conformational diversity.

The enthalpic stabilization of the CCGG series requires some increased interstrand interaction due to the dangling base. NMR experiments suggest little, if any, change in the core helical structure (Petersheim & Turner, 1983). The stabilization must then be due to either cross-strand stacking of the dangling base or a favorable change induced by the dangling base in the enthalpy of solvent interactions with the helix. These experiments cannot distinguish these two possibilities. However, solvent-perturbation studies on other nucleic acids indicate bulk solvent has little effect on stability (Dewey & Turner, 1980; Freier et al., 1981; Albergo & Turner, 1981). Thus it seems reasonable to assume the extra stability is due to interstrand stacking effects.

The results in Table II can be used to quantitatively estimate the thermodynamics associated with dangling-end stacking on CCGG. For example, the enthalpy change, ΔH_S, associated with stacking a single 3′ adenosine on a CCGG core is given by $\Delta H_S = 0.5[\Delta H(\text{CCGGA}) - \Delta H(\text{CCGG})]$. The value obtained for ΔH_S depends on which thermodynamics are used. These thermodynamics are for a dangling base going from its conformation on the end of a CCGG helix to its conformation in the separated single strands. Presumably, the dangling bases are more "stacked" in the double helix than in the single strand, thus giving rise to the enthalpy change. The enthalpy changes associated with the transition from completely unstacked to completely stacked conformations in single-stranded poly(C) and poly(A) have previously been estimated as −9 and −7 kcal/mol, respectively (Freier et al., 1981). If the single strands of the CCGG oligomers are roughly half-stacked at 39 °C, then the thermodynamics for 3′ dangling ends in Table II are consistent with these signs and magnitudes (Turner et al., 1981). Stacking on the 3′ side of the double helix seems to be phenomenologically similar to intrastrand stacking. The $\Delta\Delta H$ and $\Delta\Delta S$ associated with stacking of the 5′ A in ACCGGp have smaller magnitudes than might be expected. As discussed above, this may reflect constraints on the 5′ A or effects of the 5′ A on the CCGG core.

Stacking Contributes Roughly Three-Fifths of the Favorable Free Energy of Helix Formation Associated with Terminal AU Base Pairs in ACCGGUp. The thermodynamics of stacking derived above can be used to estimate the relative contributions of base stacking and base pairing to the stability of the terminal AU base pairs in the ACCGGUp duplex. In doing this, we use an empirical definition for base pairing. The pairing contribution is defined as the thermodynamic effect associated with having two complementary bases opposite each other, minus the sum of the effects each base has separately. For example, the free energy change associated with pairing, ΔG_p, is given by

$$\Delta G_p = 0.5[\Delta G°(\text{ACCGGUp}) - \Delta G°(\text{ACCGGp}) - \Delta G°(\text{CCGGUp}) + \Delta G°(\text{CCGG})]$$

The stacking contribution is just the sum of the 5′-adenosine and 3′-uridine dangling ends as calculated above. These definitions are empirical since it is not known if the stacking of a terminal base is the same in a base pair as it is when not base paired. The derived thermodynamics associated with "stacking" and "pairing" in the terminal AU base pairs of ACCGGUp at 39 °C are listed in Table III. Evidently, stacking contributes roughly three-fifths of the favorable free energy associated with a terminal AU base pair at 39 °C.

The data for CCGGAp suggest a similar result is expected for the terminal AU base pairs of UCCGGAp. If the terminal

Table III: Thermodynamic Changes Associated with Stacking and Pairing of a Terminal AU Base Pair in ACCGGUp

	ΔH (kcal/mol of base pair)		ΔS [cal/(K mol of base pair)]		ΔG(39 °C) (kcal/mol of base pair)	
analysis method	pairing	stacking	pairing	stacking	pairing	stacking
T_m^{-1} vs. log c_T	−6.9	−5.8	−19.4	−15.1	−0.79	−1.12
fitted parameters	−2.7	−6.6	−6.4	−17.5	−0.68	−1.11
temperature dependent	−1.1	−3.4	−1.3	−7.1	−0.71	−1.18

AU base pairs in UCCGGAp and ACCGGUp are assumed to provide the same free energy increment (Borer et al., 1974), then the 3′-A stacking would account for three-fifths of the $\Delta\Delta G$. Any contribution from the 5′ U would increase the stacking contribution.

These results are consistent with the view that hydrogen bonding is not the major source of helix stabilization since hydrogen bonds formed in the helix only replace hydrogen bonds to water (Crothers & Zimm, 1964). Theoretical calculations of helix stability under vacuum also indicate that hydrogen bonding and base stacking make comparable contributions to double-helix stability (Pullman & Pullman, 1969; Kollman et al., 1981). However, the experiments cannot be directly compared with calculations since the experiments measure the difference in stability between double and single strands in water. The calculations do not consider either the water or the stability of single strands.

Presumably, the first base pair formed in the nucleation of a double helix is stabilized only by pairing interactions. Thus the enthalpy of pairing derived above provides an estimate of the enthalpy change associated with this first base pair. The latter is generally assumed to be 0 kcal/mol without experimental justification (Borer et al., 1974). Our pairing ΔH ranges from −1.1 to −6.9 kcal/mol of base pair. Thus there may be a nonzero enthalpy associated with the formation of the initial base pair.

One purpose of this work is to gain insight into the forces stabilizing double helices. The above results indicate that stabilization due to stacking of dangling ends is associated with a favorable enthalpy term. This is consistent with the stabilization being due to electronic interactions such as dipole–dipole and London-dispersion forces. These stacking effects seem to be more important than pairing in determining the stability of a terminal AU base pair. However, both terms make significant contributions. The results also support the suggestion that stacking of a base on the 5′ side of a helix differs from stacking on the 3′ side (Alkema et al., 1981a,b).

The results also provide thermodynamics that will improve predictions for the stabilities of RNA duplexes with dangling ends. Such duplexes are important for many biological processes (Grosjean et al., 1976; Yoon et al., 1976; Steitz & Jakes, 1975). Inclusion of dangling-end effects is essential. For example, addition of two 3′ dangling adenosines to CCGG makes the $\Delta H°$ of helix association more favorable by 20–40% and increases the melting temperature at 10^{-4} M by 17 °C.

The codon–anticodon association is one case where such an effect could be particularly important. In the crystal structure of phenylalanine tRNA (Kim et al., 1974; Robertus et al., 1974; Sigler, 1975; Rich & RajBhandary, 1976), the anticodon is positioned so that the Y base would be a 3′ dangling end for the codon–anticodon helix. The Y base consists of three aromatic rings and is expected to stack strongly. However, there is evidence from relaxation kinetics (Urbanke & Maass, 1978; Labuda & Pörschke, 1980, 1982) that suggests the anticodon loop can also exist in a conformation in which the Y base is not stacked on the anticodon (Fuller & Hodgson, 1967). In this alternate conformation, a 5′ dangling U from the anticodon loop would stack on the codon–anticodon helix. The results presented in this paper indicate that if such a conformational change occurred during translocation, it would greatly reduce the stability of the codon–anticodon association.

Acknowledgments

We thank Dr. S. M. Freier for stimulating discussions and B. J. Burger and D. Mellberg for assistance with the analysis and figures.

Registry No. CCGG, 55048-62-9; CCGGp, 83831-15-6; CCGGAp, 83831-16-7; ACCGGp, 83831-17-8; CCGGUp, 83831-18-9; ACCGGUp, 83831-19-0.

References

Adler, A., Grossman, L., & Fasman, G. D. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 423–430.

Albergo, D. D., & Turner, D. H. (1981) *Biochemistry 20*, 1413–1418.

Albergo, D. D., Marky, L. A., Breslauer, K. J., & Turner, D. H. (1981) *Biochemistry 20*, 1409–1413.

Alkema, D., Bell, R. A., Hader, P. A., & Neilson, T. (1981a) *J. Am. Chem. Soc. 103*, 2866–2868.

Alkema, D., Bell, R. A., Hader, P. A., & Neilson, T. (1981b) *Biomolecular Stereodynamics, Proceedings of a Symposium* (Sarma, R. H., Ed.) Vol. I, pp 417–428, Adenine Press, Guilderland, NY.

Appleby, D. W., & Kallenbach, N. R. (1973) *Biopolymers 12*, 2093–2120.

Applequist, J., & Damle, V. (1966) *J. Am. Chem. Soc. 88*, 3895–3900.

Arter, D. B., Walker, G. C., Uhlenbeck, O. C., & Schmidt, P. G. (1974) *Biochem. Biophys. Res. Commun. 61*, 1089–1094.

Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, Chapter 11, McGraw-Hill, New York.

Bloomfield, V. A., Crothers, D. M., & Tinoco, I., Jr. (1974) *Physical Chemistry of Nucleic Acids*, Harper and Row, New York.

Borer, P. N. (1972) Ph.D. Thesis, University of California, Berkeley, CA.

Borer, P. N., Uhlenbeck, O. C., Dengler, B., & Tinoco, I., Jr. (1973) *J. Mol. Biol. 80*, 759–771.

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

Brahms, J., Maurizot, J. C., & Michelson, A. M. (1967) *J. Mol. Biol. 25*, 465–480.

Breslauer, K. J., & Sturtevant, J. M. (1977) *Biophys. Chem. 7*, 205–209.

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

Cantor, C. R., & Tinoco, I., Jr. (1965) *J. Mol. Biol. 13*, 65–77.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules*, Chapter 6, W. H. Freeman, San Francisco.

Crawford, I. P., & Stauffer, G. V. (1980) *Annu. Rev. Biochem. 49*, 163–195.

Crothers, D. M., & Zimm, B. H. (1964) *J. Mol. Biol. 9*, 1–9.

Dewey, T. G., & Turner, D. H. (1980) *Biochemistry 19*, 1681–1685.

England, T. E., & Neilson, T. (1976) *Can. J. Chem. 54*, 1714–1721.
England, T. E., & Uhlenbeck, O. C. (1978) *Biochemistry 17*, 2069–2076.
Fasman, G. D., Ed. (1975) *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, Vol. I, 3rd ed., CRC Press, Cleveland, OH.
Filimonov, V. V., & Privalov, P. L. (1978) *J. Mol. Biol. 122*, 465–470.
Freier, S. M., Hill, K. O., Dewey, T. G., Marky, L. A., Breslauer, K. J., & Turner, D. H. (1981) *Biochemistry 20*, 1419–1426.
Freier, S. M., Albergo, D. D., & Turner, D. H. (1983) *Biopolymers* (in press).
Fuller, W., & Hodgson, A. (1967) *Nature (London) 215*, 817–821.
Grosjean, H., Soll, D. G., & Crothers, D. M. (1976) *J. Mol. Biol. 103*, 499–519.
Gukovsky, I. Y., Gukovskaya, A. S., Sukhomudrenko, A. G., & Sukhorukov, B. I. (1981) *Nucleic Acids Res. 9*, 4061–4079.
Hogan, M., Dattagupta, N., & Crothers, D. M. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 195–199.
Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H. J., Seeman, N. C., & Rich, A. (1974) *Science (Washington, D.C.) 185*, 435–440.
Kollman, P. A., Weiner, P. K., & Dearing, A. 1981) *Biopolymers 20*, 2583–2621.
Labuda, D., & Pörschke, D. (1980) *Biochemistry 19*, 3799–3805.
Labuda, D., & Pörschke, D. (1982) *Biochemistry 21*, 49–53.
Lee, C.-H., & Tinoco, I., Jr. (1980) *Biophys. Chem. 11*, 283–294.
Leng, M., & Felsenfeld, G. (1966) *J. Mol. Biol. 15*, 455–466.
Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L., & Steitz, J. A. (1980) *Nature (London) 283*, 220–224.
Levitt, M. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 640–644.
Lewis, G. N., & Randall, M. (1961) *Thermodynamics*, McGraw-Hill, New York.
Lipsett, M. N. (1960) *Proc. Natl. Acad. Sci. U.S.A. 46*, 445–446.
Manning, G. S. (1978) *Q. Rev. Biophys. 11*, 179–246.
Martin, F. H. (1969) Ph.D. Thesis, Harvard University.
Martin, F., Uhlenbeck, O. C., & Doty, P. (1971) *J. Mol. Biol. 57*, 201–215.
Maxam, A., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.
Meyer, S. L. (1975) *Data Analysis for Scientists and Engineers*, Chapter 14, Wiley, New York.
Neilson, T., Romaniuk, P. J., Alkema, D., Everett, J. R., Hughes, D. W., & Bell, R. A. (1980) *Nucleic Acids Res., Spec. Publ. No. 7*, 293–311.
Nelson, J. W., Martin, F. H., & Tinoco, I., Jr. (1981) *Biopolymers 20*, 2509–2531.
Neumann, E., & Ackermann, T. (1969) *J. Phys. Chem. 73*, 2170–2178.
Orcutt, B. C., George, D. G., Fredrickson, J. A., & Dayhoff, M. O. (1982) *Nucleic Acids Res. 10*, 157–174.
Oxender, D. L., Zurawski, G., & Yanofsky, C. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5524–5528.
Petersheim, M. (1982) Ph.D. Thesis, University of Rochester.
Petersheim, M., & Turner, D. H. (1983) *Biochemistry* (third paper of three in this issue).
Pörschke, D. (1976) *Biochemistry 15*, 1495–1499.
Pörschke, D., Uhlenbeck, O. C., & Martin, F. H. (1973) *Biopolymers 12*, 1313–1335.
Pullman, B., & Pullman, A. (1969) *Prog. Nucleic Acid Res. Mol. Biol. 9*, 327–402.
Rawitscher, M. A., Ross, P. D., & Sturtevant, J. M. (1963) *J. Am. Chem. Soc. 85*, 1915–1918.
Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103–178.
Rich, A., & RajBhandary, U. L. (1976) *Annu. Rev. Biochem. 45*, 805–860.
Richards, E. G., Flessel, C. P., & Fresco, J. R. (1963) *Biopolymers 1*, 431–446.
Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C., & Klug, A. (1974) *Nature (London) 250*, 546–551.
Romaniuk, P. J., Hughes, D. W., Gregoire, R. J., Neilson, T., & Bell, R. A. (1978) *J. Am. Chem. Soc. 100*, 3971–3972.
Sanger, F., & Coulson, A. R. (1975) *J. Mol. Biol. 94*, 441–448.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Sigler, P. B. (1975) *Annu. Rev. Biophys. Bioeng. 4*, 477–527.
Snedecor, G. W., & Cochran, W. G. (1971) in *Statistical Methods*, p 190, The Iowa State University Press, Ames, IA.
Stannard, B. S., & Felsenfeld, G. (1975) *Biopolymers 14*, 299–307.
Steitz, J. A., & Jakes, K. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4734–4738.
Stone, M. P., & Borer, P. N. (1981a) *Biophys. Chem. 14*, 363–368.
Stone, M. P., & Borer, P. N. (1981b) *Biophys. Chem. 14*, 369–374.
Stone, M. P., Johnson, D. L., & Borer, P. N. (1981) *Biochemistry 20*, 3604–3610.
Suurkuusk, J., Alvarez, J., Freire, E., & Biltonen, R. (1977) *Biopolymers 16*, 2641–2652.
Thach, R. E. (1966) in *Procedures in Nucleic Acid Research* (Cantoni, G. L., & Davies, D. R., Eds.) pp 520–524, Harper and Row, New York.
Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature (London), New Biol. 246*, 40–41.
Turner, D. H., Petersheim, M., Albergo, D. D., Dewey, T. G., & Freier, S. M. (1981) *Biomolecular Stereodynamics, Proceedings of a Symposium* (Sarma, R. H., Ed.) Vol. I, pp 429–438, Adenine Press, Guilderland, New York.
Uhlenbeck, O. C., & Cameron, V. (1977) *Nucleic Acids Res. 4*, 85–98.
Urbanke, C., & Maass, G. (1978) *Nucleic Acids Res. 5*, 1551–1560.
Wu, R. (1978) *Annu. Rev. Biochem. 47*, 607–634.
Yoon, K., Turner, D. H., Tinoco, I., Jr., von der Haar, F., & Cramer, F. (1976) *Nucleic Acids Res. 3*, 2233–2241.
Young, P. R., & Kallenbach, N. R. (1978) *J. Mol. Biol. 126*, 467–479.