Effects of Flanking G·C Base Pairs on Internal Watson–Crick, G·U, and Nonbonded Base Pairs within a Short Ribonucleic Acid Duplex†

D. Alkema, P. A. Hader, R. A. Bell, and T. Neilson*

ABSTRACT: A series of pentaribonucleotides, ApGpXpGpU (where X ≡ A, G, C, or U), was synthesized to investigate the effects of flanking G·C pairs on internal Watson–Crick, G·U, and nonbonded base pairs. Sequences ApGpApCpU (T_m = 26 °C) and ApGpCpCpU (T_m = 25 °C) were each found to form a *duplex* with non-base-paired internal residues that stacked with the rest of the sequence but were *not* looped out. ApGpGpCpU also forms a duplex (T_m = 30 °C) but with dangling *terminal* nonbonded adenosines rather than *internal* nonbonded guanosines. ApGpUpCpU prefers a stacked *single*-strand conformation. In addition, contribution to duplex stability from an internal A·U or G·C base pair is enhanced by 6 °C when flanked by G·C base pairs as compared to A·U base pairs. G·C base pairs flanking an internal G·U base pair were found to be more tolerant to the altered conformation of a G·U pair and result in an increase to stability comparable with that found for an internal A·U base pair.

Major advances in sequencing have led to the elucidation of a large number of native RNA primary sequences; however, a thorough knowledge of secondary and tertiary RNA structure, essential to a complete undertanding of RNA biological function, is still lacking. RNA secondary structure is known to play an important role in the initiation of translation (Fiers et al., 1975, 1976), posttranscriptional processing (Abelson, 1979), RNA–protein interactions (Krol et al., 1978), and RNA–RNA interactions (Sitz et al., 1978). So that secondary and tertiary RNA structures may be resolved, factors affecting conformation and conformational stability, such as base pairing and base stacking, must be evaluated.

Circular dichroism studies have provided much data on the thermodynamic properties of both single strands and duplex molecules of RNA (Lomant & Fresco, 1975; Gray et al., 1980; Uhlenbeck et al., 1973). Temperature-jump studies (Grosjean et al., 1978, 1976), as well as base-pairing rules derived from model studies (Tinoco et al., 1971, 1973; Studnicka et al., 1978) and empirical energy calculations (Giessner-Prettre & Pullman, 1976), have also contributed to the understanding of RNA base stacking.

Recently, ^{1}H NMR[1] spectroscopy has allowed the observation of changes in local microenvironments of RNA molecules by following chemical shift changes of specific proton resonances as a function of temperature (Borer et al., 1975; Romaniuk et al., 1979), concentration (Krugh et al., 1975, 1976), and magnesium or sodium ion concentration (Johnston & Redfield, 1978) or through drug binding and intercalation studies (Early et al., 1978; Lee & Tinoco, 1980). This technique is nondestructive, permitting many different experiments to be carried out on a single sample. NMR has also found wide application in model duplex systems of both deoxyribonucleotide (Cross & Crothers, 1971; Crothers et al., 1973; Patel, 1976, 1977, 1979; Early et al., 1978) and ribonucleotide sequences (Arther et al., 1974; Borer et al., 1975; Kan et al., 1975; Hughes et al., 1978; Romaniuk et al., 1978a,b; Alkema et al., 1981a,b).

The existence of G·U base pairs, first suggested by Crick to explain the degeneracy of the genetic code (Crick, 1966), has become an accepted feature of RNA secondary structure. This base pair has been found within the double-helical regions of a number of native RNAs (Fiers et al., 1975, 1976; Gross et al., 1978; Rich, 1977; Hurd & Reid, 1979; Johnston & Redfield, 1981). CD studies on poly[d(GT)] (Lezuis & Domin, 1973) and poly[r(GU)] (Gray et al., 1972; Gray & Ratliff, 1977) suggested formation of an intermolecular structure but provided insufficient proof to conclude duplex formation.

It was suggested that G·U and G·T base pairs could occur in helical structures containing predominately normal Watson–Crick base pairs (Lomant & Fresco, 1975). This hypothesis was supported by X-ray studies of yeast $tRNA^{Phe}$, which confirmed the existence of the G4-U69 base pair, in the wobble conformation proposed by Crick (Quigley et al., 1975; Ladner et al., 1975). However, it was the discovery of hydrogen-bonded imino-ring proton resonances in the ^{1}H NMR spectra of tRNA that led to unequivocable proof for G·U base pair formation (Johnston & Redfield, 1978, 1981; Robillard et al., 1976). NMR and ethidium bromide binding studies of poly[d(GT)] (Early et al., 1978) found that the polymer formed a stable double helix where G·T forms a wobble base pair: observation of N–H resonances indicated —NH---O= hydrogen bond formation, expected for a wobble base pair, rather than –NH---N– or –OH---N– hydrogen bonds. Further NMR and nuclear Overhauser effect studies firmly established the existence of G·U base pairs in the double-helical regions of tRNA (Reid et al., 1979; Hurd & Reid, 1979; Johnston & Redfield, 1978, 1981).

Although the existence of G·U base pairs is established, the relative stability of a G·U base pair within a Watson–Crick double helix is still unclear. Uhlenbeck and co-workers found that terminal G·U base pairs were as stable as terminal A·U base pairs (Uhlenbeck et al., 1970) but that an internal G·U base pair did not increase duplex stability as much as an

† From the Departments of Biochemistry and Chemistry, McMaster University, Hamilton, Ontario L8N 3Z5, Canada. *Received October 2, 1981.*

[1] Abbreviations: AGCU, tetranucleotide triphosphate ApGpCpU. In addition to the abbreviations recommended by the IUPAC–IUB Commission (1970), the following are also used: trac, triphenylmethoxyacetyl; t, tetrahydropyranyl; pO^-, 3′-*O*-(2,2,2-trichloroethyl) phosphate; p between two characters, 3′,5′-(2,2,2-trichloroethyl) phosphotriester; MST, mesitylenesulfonyl 1,2,4-triazolide; DSS, 4,4-dimethyl-4-silapentane-1-sulfonate; NMR, nuclear magnetic resonance; CD, circular dichroism; T_m, melting temperature (degrees Celsius). All oligoribonucleotides are written in the 5′ → 3′ sequence, and the bases are numbered from the 5′ end. For double helices, the base pairs are numbered from left to right:

0006-2960/82/0421-2109$01.25/0 © 1982 American Chemical Society

internal A·U base pair (Uhlenbeck et al., 1971). This discrepancy was not explained. A review by Clarke (1977) suggests that G·U base pairs could be accommodated internally by an RNA double helix without causing gross helical distortion. However, it may result in a center of potential weakness since it will require certain accommodations of the helix backbone. An NMR study confirming the formation of an internal G·U base pair in an otherwise normal short RNA duplex, CAUUG:CAGUG, indicated that such a G·U base pair did not stabilize the duplex and was in fact a distinct center of instability in the helix (Romaniuk et al., 1979). Since hydrogen bonding in a wobble G·U base pair requires a shift in the glycosyl torsion angles from those usually associated with Watson–Crick base pairing, slight changes in the stacking interactions could lead to an overall decrease in duplex stability. Also, an internal G·U base pair may be unable to adopt its preferred conformation and thus would create a region of local instability affecting the nucleation and closing of the entire duplex (Romaniuk et al., 1979).

Duplex CAGUG:CAUUG contains a G·U base pair flanked by standard A·U base pairs. This communication reports an extension of this earlier work by studying the duplex AGGCU:AGUCU containing a G·U base pair flanked by Watson–Crick G·C base pairs. The G·C pairs are considerably stronger than A·U base pairs and are expected to allow less conformational perturbation of the helix backbone and may significantly affect the stability of the G·U base pair. Four pentaribonucleotides, AGACU, AGUCU, AGGCU, and AGCCU, were used to study the effect of the single insertion of an additional A·U (AGACU:AGUCU), G·C (AGGCU:AGCCU), or G·U (AGGCU:AGUCU) base pair on the stability of the duplex formed relative to reference duplex AGCU:AGCU (Neilson et al., 1980).

Each of the four pentaribonucleotides was also studied individually by ^{1}H NMR to observe the effects of nonbonded base opposition on duplex stability. Nonbonded residues are also a common feature in RNA duplex regions (Lomant & Fresco, 1975; Clarke, 1977). Lomant & Fresco (1975) found that mixtures of homopolymers and copolymers adopt extrahelical conformations with noncomplementary base pair opposition. Similarily, studies with longer sequences, of the type $A_nC_mU_n$, showed bulges with loops of two or three bases (m = 2 or 3) being more stable than loops of one base (m = 1) (Uhlenbeck et al., 1973). Recently, dodecamer DNA duplexes containing a single nonbonded T·T opposition have been found to adopt bulge-loop conformation with little loss of stability (Haasnoot et al., 1979, 1980). Romaniuk (1979) found no evidence for looping out of bases, even when U·U opposition occurred, and suggested that the pentaribonucleotides used prefer a single-stranded stacked conformation.

Materials and Methods

Materials. Nucleosides were purchased from Terochem Laboratories, Ltd., Edmonton, Alberta, Canada; mesitylenesulfonyl chloride, 1,2,4-triazole, and 2,2,2-trichloroethanol were from Aldrich, Milwaukee, WI; methylene chloride and silica gel (40–140 mesh) were from J. T. Baker, Philipsburg, NJ; Silica gel G (250 μm) thin-layer chromatography plates were from Analtech, Inc., Newark, DE.

Chemical Synthesis of Oligoribonucleotides. The sequences used in this study were synthesized by the phosphotriester method developed in this laboratory and fully described in the literature (Neilson & Werstiuk, 1974; England & Neilson, 1976; Werstiuk & Neilson, 1976). The fully deblocked sequences were characterized by ^{1}H NMR. The base ratios of the oligomers were confirmed from the distinctive aromatic proton resonances of the spectra.

NMR Methodology. The NMR spectra were obtained with Bruker WH-90, WM-250, and WH-400 spectrometers operating in the Fourier transform mode and equipped with quadrature detection. Probe temperatures were controlled to within ±1 °C by Bruker variable-temperature units and calibrated by thermocouple measurements. *tert*-Butyl alcohol-*d* was used as an internal reference (1.231 ppm), and the resonance positions reported are in parts per million (ppm) relative to the standard DSS (sodium 4,4-dimethyl-4-silapentane-1-sulfonate). The field frequency lock was provided by the deuterium signal of D_2O. The spectra were acquired in 200–300 scans. The samples were lyophilized twice from D_2O and then dissolved in 100% D_2O (Aldrich), which contained 0.01 M sodium phosphate buffer (pD 7.0) and 1.0 M sodium chloride. The sample concentrations were as follows: AGACU, 6.34 mM; AGUCU, 6.13 mM; AGGCU, 6.38 mM; AGCCU, 6.53 mM. Total strand concentration for each of the duplex samples was 6.25 mM.

^{1}H NMR analysis of short oligoribonucleotides is facilitated by the characteristic signals displayed by each of the four regular nucleosides. An incremental procedure developed by Borer et al. (1975) and refined by Everett et al. (1980) is used. Since the chemical shifts of protons on each nucleoside residue are affected by the neighboring residues, it is advantageous to assemble a desired sequence in a stepwise fashion, studying each intermediate to aid in the assignment of the final sequence. The stepwise phosphotriester method employed readily provides the final sequence and the required intermediates. Use of self-complementary sequences, which halves the number of signals, further simplifies this procedure.

Results

Synthesis and Characterization of Oligoribonucleotides. To study the effect one additional internal base pair has on the stability of the self-complementary duplex AGCU (Neilson et al., 1980), we synthesized a series of non-self-complementary pentaribonucleotides, AGXCU (where X is A, G, C, or U). The three complementary duplexes containing the extra A·U, G·C, or G·U pairs were prepared by appropriate mixing experiments. An outline of the synthetic scheme is presented in Figure 1. The preparative data for the protected oligoribonucleotides are contained in Table I. Full deblocking of the sequences was carried out by using a three-step procedure (England & Neilson, 1976). Deblocked oligoribonucleotides were isolated with paper chromatography (Table II). Sequence integrity was checked by ^{1}H NMR analysis of the products after each synthetic step. The results of these analyses are presented in Table III.

Chemical Shift Assignments of the Proton Resonances. The low-field nonexchangeable proton chemical shift assignments for the four pentamers were determined by two methods: by comparison with the NMR data on AGCU (Neilson et al., 1980) and by the procedure of incremental analysis (Borer et al., 1975; Everett et al., 1980). The 70 °C spectrum for AGACU has two additional singlets at 8.281 and 8.134 ppm and a doublet at 5.998 ppm corresponding to the A(3)H-8, A(3)H-2, and A(3)H-1′ proton resonances, respectively. The upfield shifts exhibited by the proton resonances of the neighboring residues, when compared to those of AGCU, were explained by the strong ring-current shielding effect of adenosine (Giessner-Prettre et al., 1976). The two GH-8 resonances in AGGCU were difficult to distinguish since they occurred at 7.928 and 7.914 ppm. The GH-8 resonance in AGCU occurs at 7.628 ppm (Neilson et al., 1980); however, the G(2)H-8 of AGGCU was assigned to the higher field

Table I

reactants						products		
	quantity			quantity			quantity	
compd[a]	mg	mmol	compd[b]	mg	mmol	compd[c]	mg	% yield
A	2000	2.650	G	1500	3.180	AG	2200	59
AG	500	0.350	A	191	0.420	AGA	370	51
AGA	250	0.121	C	43	0.169	AGAC	208	64
AGAC	160	0.056	U	22	0.067	AGACU	113	59
AG	500	0.350	C	181	0.420	AGC	470	65
AGC	350	0.170	C	109	0.250	AGCC	292	64
AGCC	220	0.082	U	33	0.098	AGCCU	160	61
AG	500	0.350	U	138	0.420	AGU	340	50
AGU	250	0.130	C	69	0.160	AGUC	213	65
AGUC	213	0.083	U	33	0.100	AGUCU	127	50
AG	700	0.490	G	278	0.590	AGG	480	47
AGG	200	0.072	C	62	0.144	AGGC	150	58
AGGC	110	0.041	U	16	0.049	AGGCU	61	46

[a] Column 1 contains the 5′-trityloxyacetyl reactants and A is trac-bzAtOH. [b] Column 4 contains the incoming nucleosides, all of which are protected at the 2′-OH group with a tetrahydropyranyl residue and all of which except U are protected at their NH_2 groups with a benzoyl residue. [c] Column 7 contains the trityloxyacetyl products and, for example, AG stands for trac-bzAt-*p*-bzGtOH. Two equivalents of pyridinium 2,2,2-trichloroethyl monophosphate activated by 4 equiv of mesitylenesulfonyl 1,2,4-triazolide (MST) in anhydrous pyridine is used in each phosphorylation step. The coupling step to the incoming nucleoside derivative is driven by 1.2 equiv of MST.

Table II: Chromatographic Data and Deprotection Yields for the AGXCU Series

compd[a]	R_f[b]	yield (%)[c]
ApG	0.40	60
ApGpA	0.29	73
ApGpU	0.34	85
ApGpC	0.36	60
ApGpG	0.32	26
ApGpApC	0.19	79
ApGpUpC	0.20	49
ApGpCpC	0.20	62
ApGpGpC	0.21	77
ApGpApCpU	0.15	62
ApGpUpCpU	0.16	81
ApGpCpCpU	0.17	77
ApGpGpCpU	0.17	32

[a] Refers to the free, deblocked oligoribonucleotide. [b] Chromatography system: 1.0 M ammonium acetate–ethanol (50:50 v/v) on Whatman No. 40 paper. [c] Calculated from UV spectrophotometric data assuming a 90% hypochromicity factor.

Table III: Chemical Shifts[a] of the Oligoribonucleotides in D_2O[b] at 70 °C

proton	AGCU	AGACU	AGGCU	AGUCU	AGCCU
A(1)H-8	8.248	8.196	8.210	8.237	8.245
A(1)H-2	8.176	8.121	8.164	8.185	8.170
G(2)H-8	7.929	7.903	7.914	7.953	7.926
A(3)H-8		8.281			
A(3)H-2		8.134			
G(3)H-8			7.928		
U(3)H-6				7.752	
C(3)H-6					7.713
C(4)H-6	7.751	7.711	7.752	7.800	7.809
U(5)H-6	7.768	7.757	7.773	7.794	7.796
A(1)H-1′	5.978	5.931	5.942	5.973	5.977
G(2)H-1′	5.791	5.731	5.780	5.780	5.780
A(3)H-1′		5.998			
G(3)H-1′			5.839		
U(3)H-1′				5.878	
C(3)H-1′					5.844
C(4)H-1′	5.903	5.821	5.910	5.924	5.892
U(5)H-1′	5.892	5.874	5.892	5.927	5.912
U(3)H-5				5.779	
C(3)H-5					5.799
C(4)H-5	5.850	5.827	5.853	6.014	5.979
U(5)H-5	5.821	5.801	5.823	5.855	5.843

[a] Chemical shifts are in parts per million (ppm) relative to DSS, with *tert*-butyl alcohol-*d* as an internal reference, and are accurate to ±0.005 ppm. [b] pD 7.0.

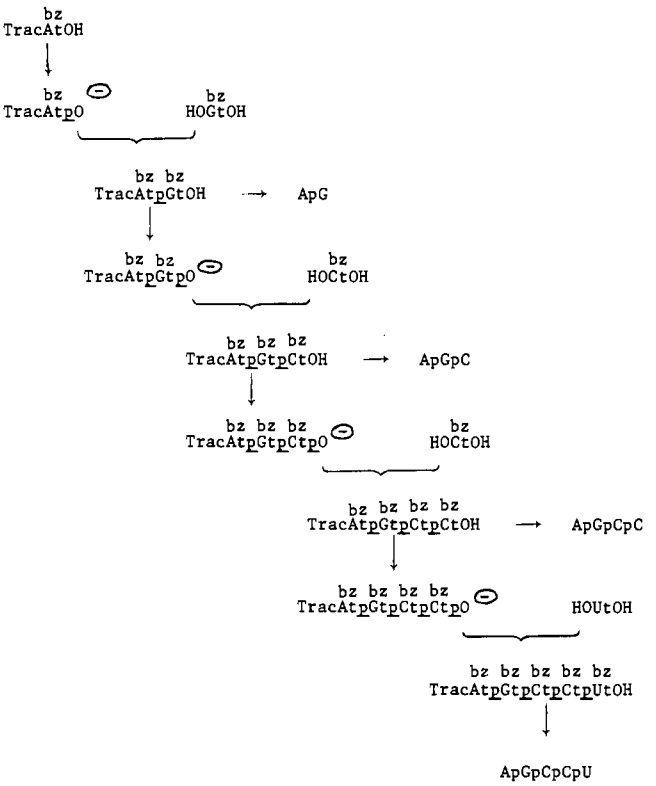


FIGURE 1: Scheme for the chemical synthesis of ApGpCpCpU. Abbreviations and details of synthesis are described in the text.

resonance (7.914 ppm) because the G(3) residue would result in some shielding of its neighbors (Bell et al., 1981). The G(2)H-1′ and G(3)H-1′ signals were more readily distinguishable at 5.780 and 5.839 ppm, respectively. In the spectrum of AGUCU the additional resonances were readily assigned to the U(3) protons at 7.752 (H-6), 5.878 (H-1′), and 5.779 ppm (H-5). The C(3)H-6, H-5, and H-1′ protons in AGCCU resonated at 7.713, 5.799, and 5.844 ppm while the C(4)H-6, H-5, and H-1′ protons occurred at 7.809, 5.979, and 5.892 ppm. Simple comparison of this sequence with AGCU was insufficient, so the method of incremental analysis, developed by Borer et al. (1975), was used to confirm the assignments (Table IV).

Effect of a Noncomplementary Base Opposition on the Stability of a Short Duplex. Theoretically, each AGXCU

Table IV: Incremental Analysis of AGCCU in D_2O at 70 °C[a]

proton	AG	AGC	AGCC	AGCCU
A(1)H-8	8.238	8.238	8.247	8.245
A(1)H-2	8.186	8.170	8.168	8.170
G(2)H-8	7.942	7.934	7.927	7.926
C(3)H-6		7.748	7.722	7.713
C(4)H-6			7.791	7.809
U(5)H-6				7.796
A(1)H-1′	5.967	5.969	5.979	5.977
G(2)H-1′	5.842	5.809	5.785	5.780
C(3)H-1′		5.889	5.873	5.844
C(4)H-1′			5.886	5.892
U(5)H-1′				5.912
C(3)H-5		5.872	5.823	5.799
C(4)H-5			5.978	5.979
U(5)H-5				5.843

[a] See the footnotes found in Table III.

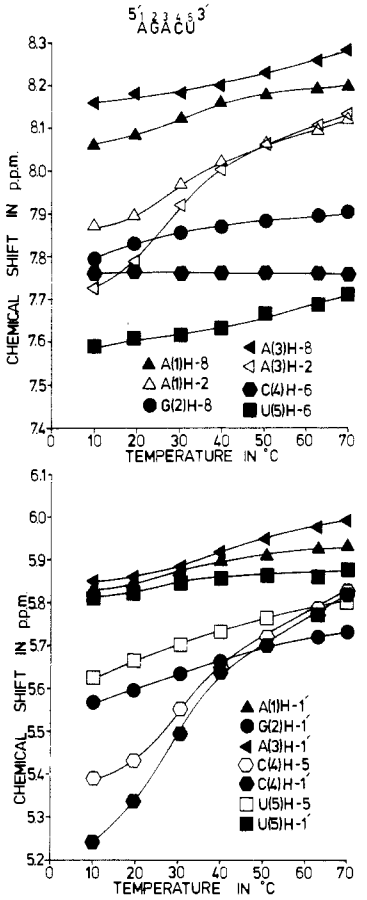
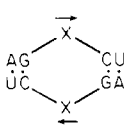


FIGURE 2: Chemical shifts vs. temperature plots for AGACU.

pentaribonucleotide can form a duplex containing a looped-out region:



X = A, U, C, or G

A duplex of this nature would provide valuable information on the effects that single-base bulge loops have on the formation of adjacent base pairs. Figures 2–5 provide the results of the variable-temperature experiments on each pentaribonucleotide.

Unlike the CAXUG sequences, which displayed only minor upfield shifts with temperature (Romaniuk et al., 1979), we find that the AGXCU series of compounds displays significant upfield shifts. Some proton curves display sigmoidal behavior, which could indicate base pairing.

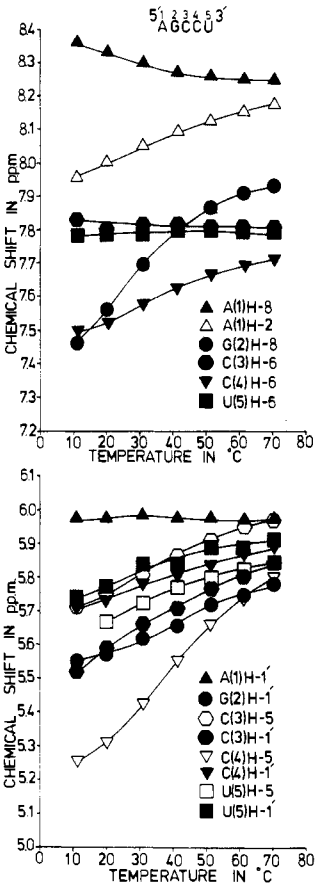


FIGURE 3: Chemical shift vs. temperature plots for AGCCU.

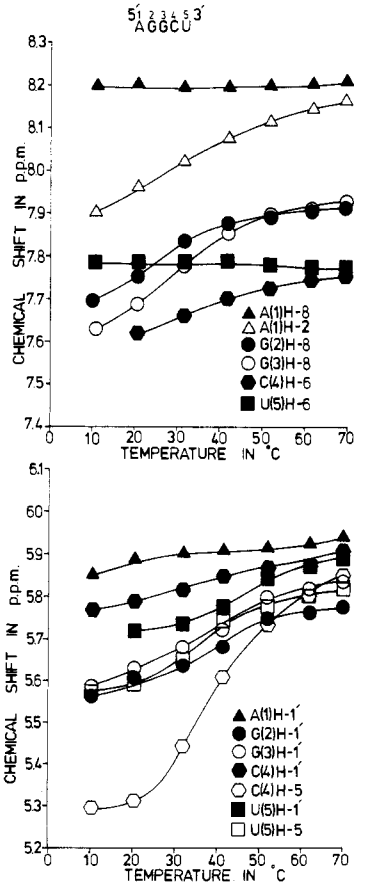


FIGURE 4: Chemical shift vs. temperature plots for AGGCU.

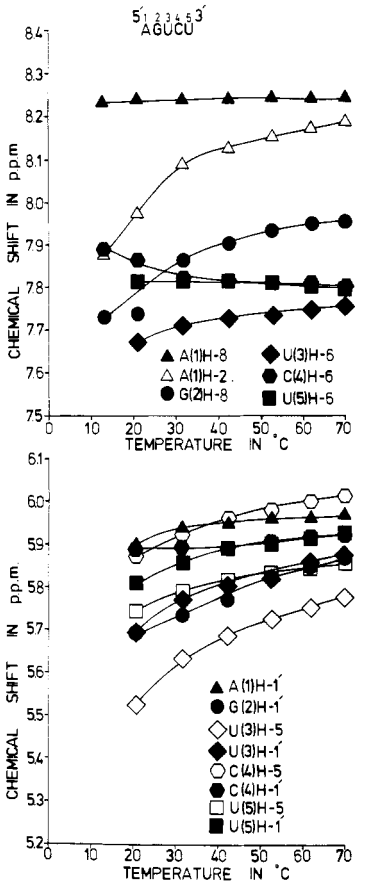


FIGURE 5: Chemical shift vs. temperature plots for AGUCU.

Studies on Duplex Formation in Complementary Sets. Before observing the formation and stability of an internal G·U base pair, we studied the effects of other internal Watson–Crick base pairs on the T_m of AGCU. The duplex formed by AGACU:AGUCU illustrated the effects of an additional internal A·U base pair. Duplex AGACU:AGUCU has a T_m of 45 °C (averaged from the 15 sigmoidal curves) in Figure 6. This represents a difference of 11 °C higher than that obtained for AGCU (T_m = 34 °C).

Similarly, the effect of an additional internal G·C base pair on duplex stability was determined by using the complementary set AGGCU:AGCCU. Figure 7 shows the results of this experiment. An average of 18 curves displaying sigmoidal character resulted in a T_m of 54 °C, a full 20 °C higher than the parent tetramer duplex AGCU.

Formation of a G·U Wobble Base Pair within a Double Helix. By mixing the single-strand pentamers AGGCU and AGUCU to make the resultant duplex, we were able to observe the formation of an internal G·U base pair. The sigmoidal behavior of the chemical shift vs. temperature plots, given in Figure 8, clearly indicates that duplex AGGCU:AGUCU forms, and averaging the 10 curves gave a T_m of 44 °C. Comparison with duplex AGACU:AGUCU (T_m = 45 °C) demonstrates that the internal G·U base pair is virtually equivalent to the internal A·U base pair in its ability to stabilize duplex formation. This result contrasts with our earlier studies with the series CAXUG where a G·U base pair neither stabilized nor destabilized the duplex formed by CAGUG:CAUUG, which had a T_m of 24 °C, similar to that of the parent tetramer CAUG (T_m = 24 °C).

Discussion

This work presents an extension of a study that established the existence of an internal G·U base pair within a normal RNA duplex (Romaniuk et al., 1979), although the G·U base pair appeared to result in a distinct center of instability. In Romaniuk's study, the internal G·U base pair in the duplex CAUUG:CAGUG was flanked by two Watson–Crick A·U base pairs. In this study, we are interested in determining the effect of G·C, A·U, and G·U base pairs when flanked by two Watson–Crick G·C base pairs and comparing these results with those from the earlier study where the internal base pair was flanked by two weaker A·U base pairs. Since the T_m for reference duplex AGCU is higher than that for duplex CAUG by ~10 °C, the possibility of duplex formation by the single strands AGXCU, where X = A, G, C, or U, is conceivable.

Four pentaribonucleotide sequences, AGACU, AGUCU, AGGCU, and AGCCU, were prepared. Each of these sequences has the potential to form a duplex of the type

with internal noncomplementary base opposition and can be compared to the reference duplex AGCU:AGCU (Neilson et al., 1980). Subsequent mixing of two of these pentamers provided a duplex with an additional internal A·U, G·C, or G·U base pair. By studying these different duplex systems over the temperature range 10–70 °C, it was possible to determine the effects of (1) noncomplementary base oppositions or (2) an additional base pair on duplex stability when flanked by G·C base pairs.

In our study of the AGXCU series, we obtained some unexpected results. In Figure 2, a significant number of the chemical shift vs. temperature plots for AGACU resulted in curves that showed large *upfield* shifts with strong sigmoidal characteristics, which are normally associated with duplex formation (Gralla & Crothers, 1973) or strong cooperative stacking interactions (Stannard & Fesenfeld, 1975). This suggested that either AGACU was forming a highly stacked single-stranded structure, such as the rodlike structure seen with poly(rA) at very low temperatures (Stannard & Fesenfeld, 1975), or duplex formation had occurred with a non-base-paired A·A opposition in the center of the duplex. However, the *upfield* shifts of the chemical shifts seem to indicate that looping out did not occur. Perhaps duplexing of the sequence with the adenosine residues stacking cooperatively into the sequence is being observed. The A(1)H-2, A(1)H-1', C(4)H-5, and C(4)H-1' protons all show sigmoidal behavior, and these bases could base pair. The A(3)H-2 also displays sigmoidal behavior but cannot form a base pair. Adenine, however, stacks better than other bases and is perhaps drawn cooperatively into a stacked conformation, allowing duplex formation to occur with the rest of the sequence. In related studies (Alkema et al., 1981a), we have shown that terminal nonpaired adenosines can, in fact, determine stacking direction and duplex stability. In duplex AGACU:AGACU, the central adenosine, instead of looping out of the sequence and thus disrupting the stacking and base pairing of its neighbors, stacks with the other bases and, while still reducing the overall duplex stability, does not totally disrupt duplex formation.

In their review, Lomant & Fresco (1975) gave a rough estimate for calculating the duplex instability caused by a noncomplementary base opposition. In a duplex containing 50% A·U and 50% G·U base pairs, one noncomplementary base opposition for every 100 normal base pairs reduced the T_m by ~1 °C. Duplex AGACU:AGACU is 50% A·U and 50% G·C, containing one noncomplementary base pair opposition for four normal base pairs, which is equivalent to 25 in

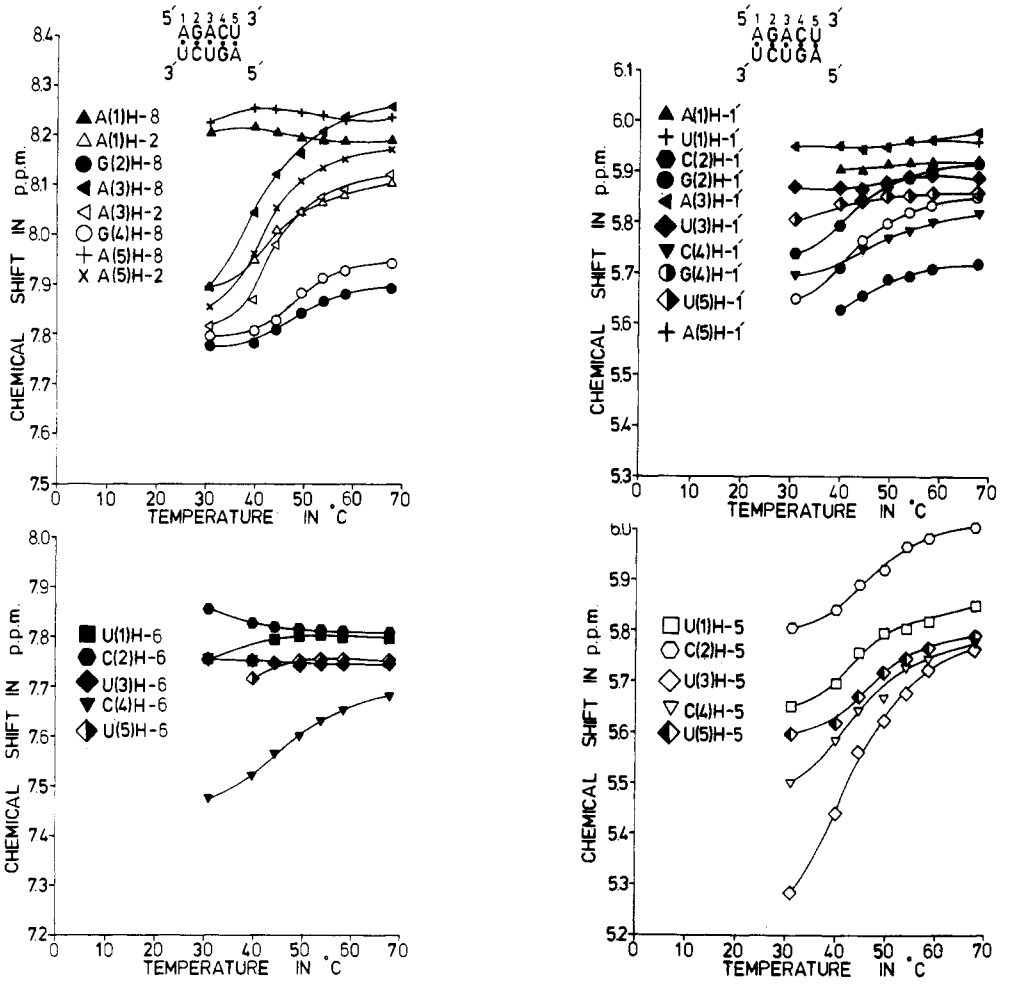
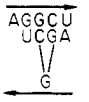


FIGURE 6: Chemical shift vs. temperature plots for duplex AGACU:AGUCU [U(1)H-1′ and U(3)H-1′ overlap].

100, reducing the T_m by 25 °C. The T_m for the duplex AGACU:AGUCU is 45 °C. Therefore, the T_m for AGACU:AGACU could be predicted to be ~20 °C or compared with AGGCU:AGCCU (T_m = 54 °C) to be ~29 °C. From the average of the five sigmoidal curves in AGACU:AGACU (Figure 2), a T_m of 26 °C was obtained, which suggests that AGACU may form a duplex with internal A·A opposition that does *not* loop out but rather stacks within the duplex.

Figure 3 shows the curves for the temperature vs. chemical shift plots for AGCCU. This sequence also exhibits many proton upfield shifts. Again several of the curves show sigmoidal behavior, indicating duplex formation. These curves are more difficult to interpret because the C(3)H-6 curves show a slight downfield shift while the C(3)H-5 curves show sigmoidal behavior and the H-1′ proton for C(3) displays a simple upfield shift. However, since cytidine is a poorer stacker than adenosine, a weaker effect on the stacking of the rest of the sequence could be expected. Very likely an equilibrium of different states exists: one may exhibit some looping out while another shows cooperative stacking within the duplex and still another is the stacked single strand. However, formation could occur with C·C opposition causing duplex instability. This duplex also contains a 50% G·C and 50% A·U base pair mixture with one noncomplementary base pair opposition and four normal base pairs. Again, a T_m between 20 and 29 °C could be expected according to the guidelines of Lomant & Fresco (1975). In fact, an average of the five curves that show sigmoidal behavior indicates a T_m of 25 °C.

Figure 4 contains the chemical shift vs. temperature plots for AGGCU. Once again, sigmoidal behavior of a significant number of curves (8 of 13) can be noted with an average T_m of 30 °C. Several explanations are possible for this sequence also. First, a duplex of a similar nature to that already described for AGACU could form. Second, a duplex of the nature described by Lomant & Fresco (1975) could occur, resulting in a duplex of the type



where one G residue is found in an extrahelical position. However, a third possible duplex with two terminal 3′-nonbonded adenosines similar to that formed from ACAUG (Neilson et al., 1980) could arise. This resultant duplex

which contains two 3′-dangling adenosines, two terminal G·U base pairs, and two internal G·C base pairs is the most likely since related duplex GGCU:GGCU with two terminal G·U base pairs (Mizuno et al., 1981) forms with a T_m of ~30 °C (D. Alkema, R. A. Bell, P. A. Hader, and T. Neilson, unpublished observations). Therefore, it is unlikely that a duplex with G·G opposition forms from AGGCU.

Figure 5 displays the chemical shift vs. temperature plot for AGUCU. Unlike the other sequences this pentaribonucleotide shows absolutely no sigmoidal behavior, only general upfield shifting of the resonances despite a possible U·U opposition.

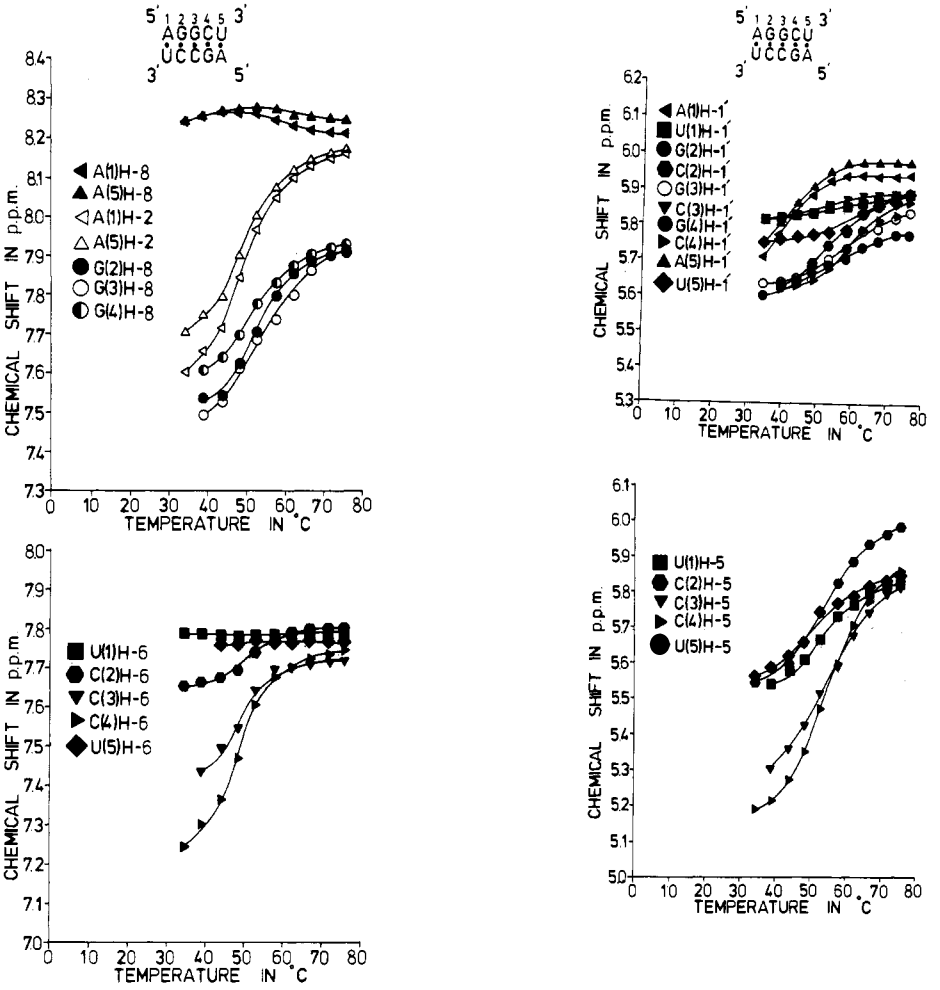


FIGURE 7: Chemical shift vs. temperature plots for duplex AGGCU:AGCCU.

Looping out of the uridine residue should result in the least disruption of base stacking because it stacks the poorest of the four major bases. Why then does this sequence show no sign of duplex formation? The answer lies in the fact that uracil is a *very* poor stacker (Everett et al., 1980). Since stacking plays an important role in duplex formation and stability (Alkema et al., 1981a,b; Turner et al., 1981), the presence of a nonbonded uridine may be sufficient to disrupt stacking of the sequence totally and thus prevent the cooperative process of duplex formation.

These results obtained from the AGXCU series of pentaribonucleotides were entirely different from those obtained in an earlier study of the series of compounds of CAXUG (Romaniuk et al., 1979). A sequence effect was probably responsible, since the G·C base pairs are immediately adjacent to the region of instability in AGXCU whereas in CAXUG this region is flanked by weaker A·U base pairs. This has been noted previously when the T_m for AGCU (34 °C) was compared to the T_m for CAUG (24 °C). In addition, the sequence AGXCU also possesses a better stacking arrangement with uracil at a terminal position where it is less disruptive than in an internal position as in CAXUG. Mixing of the appropriate pentamers results in duplexes with an additional A·U, G·C, or G·U base pair to illustrate the effects of additional internal base pairs, on duplex stability, when flanked by G·C base pairs.

An additional internal A·U base pair, as found in the duplex formed by AGACU:AGUCU, resulted in a T_m of 45 °C as determined by averaging the sigmoidal curves in Figure 6. The T_m for AGACU:AGUCU represents an 11 °C increase over

Table V: Comparison of T_m Values for AGXCU and CAXUG Series of Duplexes

duplex	T_m (°C)	duplex	T_m (°C)
AGCU UCGA	34.0	CAUG GUAC	24.0
AGGCU UCUGA	44.0	CAGUG GUUAC	23.0
AGACU UCUGA	45.0	CAAUG GUUAC	28.0
AGGCU UCCGA	54.0	CAGUG GUCAC	38.0

that found for AGCU (Neilson et al., 1980), shown in Table V. On the other hand, an internal A·U base pair in CAAUG:CAUUG only increased the T_m by 5 °C over that for CAUG (Romaniuk et al., 1979). The difference in the amount of stability conferred by an A·U base pair is primarily due to sequence. The A·U base pair now flanked by G·C base pairs rather than A·U base pairs, is significant. This result is supported by earlier work (Tinoco et al., 1973), which indicated that the free energies of $\frac{5'\text{-G-A-3'}}{3'\text{-C-U-5'}}$ and $\frac{5'\text{-A-C-3'}}{3'\text{-U-G-5'}}$ sequences are greater than for $\frac{5'\text{-A-U-3'}}{3'\text{-U-A-5'}}$ and $\frac{5'\text{-U-U-3'}}{3'\text{-A-A-5'}}$ sequences. It is noteworthy that the T_m of duplex AGACU:AGUCU is comparable to that found for the duplex AGCUA:AGCUA (T_m = 46 °C) (Neilson et al., 1980) where 3'-dangling adenosines adjacent to the duplex AGCU:AGCU stabilize to the same extent as an A·U base pair.

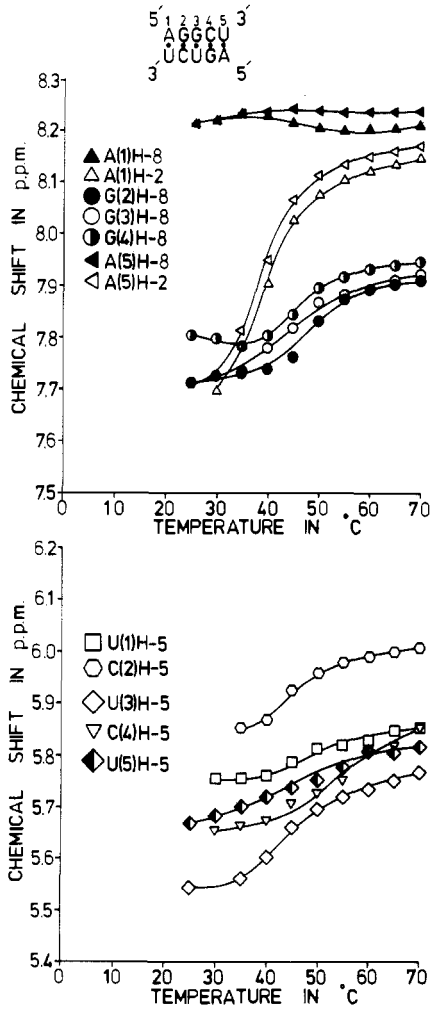


FIGURE 8: Chemical shift vs. temperature plots for duplex AGGCU:AGUCU.

Figure 7 illustrates the chemical shift vs. temperature plots for the duplex AGGCU:AGCCU. The average T_m for this duplex is 54 °C (Table V) and represents a 20 °C increase over the T_m found for reference AGCU, which demonstrates the duplex stabilizing effect of three G·C base pairs as compared with two G·C base pairs. This increase in T_m also indicates that the contribution to stability of a G·C base pair when flanked by two G·C base pairs is 9 °C more stable than that of an A·U base pair flanked by two G·C base pairs. Or, in other words, replacing an internal G·C base pair with an A·U base pair destabilizes the duplex, as previously suggested (Tinoco et al., 1973).

Sequence dependence on duplex stability was evident from consideration of inserting pairs into different cores, for example, reference tetramer duplexes AGCU:AGCU ($^{5'\text{-G-C-3'}}_{3'\text{-C-G-5'}}$ core) and CAUG:CAUG ($^{5'\text{-A-C-3'}}_{3'\text{-U-A-5'}}$ core). Insertion of a G·C pair in a $^{5'\text{-G-C-3'}}_{3'\text{-C-G-5'}}$ core showed an increase of 6 °C over a similar insertion into a corresponding $^{5'\text{-A-U-3'}}_{3'\text{-U-A-5'}}$ core (Table V), taking into account that the T_m difference between reference tetramer duplexes is 10 °C. Similar comparison of A·U pair insertion into the same cores also showed a difference of 6 °C. The T_m of a duplex containing a Watson–Crick base pair flanked by two G·C pairs is 6 °C greater than that of the corresponding duplex which contains the same base pairs flanked by two A·U pairs.

The two single strands AGGCU and AGUCU gave a complementary duplex, AGGCU:AGUCU, containing an internal G·U base pair. Averaging the T_m's from the curves in Figure 8 gave a T_m of 44 °C, which represents an increase in duplex stability of 10 °C over the reference duplex, in contrast to CAGUG:CAUUG where the G·U base pair did not enhance stability (Table V). In the AGXCU series, the G·U base pair results in an increase of duplex stability equivalent to the insertion of an additional A·U base pair.

Possible reasons for the divergence between the two series are related to sequence. G·C base pairs flanking the G·U base pair are probably more tolerant of the altered helix conformation than corresponding A·U base pairs. In addition, stacking interactions are stronger with a GGC stack in one strand and a GUC stack in the other as compared to a AGU and AUU stack in the CAXUG series. UU sequences are very poorly stacked and thus inhibit strong strand stacking within CAUUG, which in turn weakens duplex formation.

In summary, short oligoribonucleotide sequences containing perturbations from normal base pairs have the potential to form duplexes. Noncomplementary base oppositions A·A and C·C exist within duplex AGXCU (where X = A or C). The contribution to duplex stability of any base pair whether A·U, G·C, or G·U is enhanced when flanked by G·C base pairs.

References

Abelson, J. (1979) *Annu. Rev. Biochem. 48*, 1035.

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

Alkema, D., Bell, R. A., Hader, P. A., & Neilson, T. (1981b) in *Biomolecular Stereodynamics* (Sarma, R. H., Ed.) p 417, Adenine Press, New York.

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

Bell, R. A., Everett, J. R., Hughes, D. W., Alkema, D., Hader, P. A., Neilson, T., & Romaniuk, P. J. (1981) *Biopolymers 20*, 1383.

Borer, P. N., Kan, L. S., & Ts'o, P. O. P. (1975) *Biochemistry 14*, 4847.

Clarke, B. F. C. (1977) *Prog. Nucleic Acid Res. Mol. Biol. 20*, 1.

Crick, F. H. C. (1966) *J. Mol. Biol. 19*, 548.

Cross, A. D., & Crothers, D. M. (1971) *Biochemistry 10*, 4015.

Crothers, D. M., Hibers, C. W., & Shulman, R. G. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 2899.

Early, T. A., Olmsted, J., Kearns, D. R., & Lezuis, A. G. (1978) *Nucleic Acids Res. 5*, 1955.

England, T. E., & Neilson, T. (1976) *Can. J. Chem. 54*, 1714.

Everett, J. R., Hughes, D. W., Bell, R. A., Alkema, D., Neilson, T., & Romaniuk, P. J. (1980) *Biopolymers 19*, 557.

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Meregaert, J., Min Jou, W., Raeymaekers, A., Volckaert, G., Ysebaert, M., Vande Kerckhove, J., Nolf, F., & Van Montagu, M. (1975) *Nature (London) 256*, 273.

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Tserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeyamekers, A., Van den Berghe, A., Volckaert, G., & Ysebaert, M. (1976) *Nature (London) 260*, 500.

Giessner-Prettre, C., & Pullman, B. (1976) *Biochem. Biophys. Res. Commun. 70*, 578.

Giessner-Prettre, C., Pullman, B., Borer, P. N., Kan, L.-S., & Ts'o, P. O. P. (1976) *Biopolymers 15*, 2277.

Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301.

Gray, D. M., & Ratliff, R. L. (1977) *Biopolymers 16*, 1331.

Gray, D. M., Tinoco, I., Jr., & Chamberlin, M. J. (1972) *Biopolymers 11*, 1235.

Gray, D. M., Vaughn, M., Ratliff, R. L., & Hayes, F. N. (1980) *Nucleic Acids Res. 8*, 3695.

Grosjean, H., Söll, D. G., & Crothers, D. M. (1976) *J. Mol. Biol. 103*, 499.
Grosjean, H., DeHenan, S., & Crothers, D. M. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 610.
Gross, H. J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H., & Saenger, H. L. (1978) *Nature (London) 273*, 203.
Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, J. F. M., Van Boom, J. H., & Altona, C. (1979) *Nature (London) 281*, 235.
Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, J. F. M., Van Boom, J. H., & Altona, C. (1980) *Nucleic Acids Res. 8*, 169.
Hughes, D. W., Bell, R. A., England, T. E., & Neilson, T. (1978) *Can. J. Chem. 56*, 2243.
Hurd, R. E., & Reid, B. R. (1979) *Biochemistry 18*, 4017.
Johnston, P. D., & Redfield, A. G. (1978) *Nucleic Acids Res. 5*, 3913.
Johnston, P. D., & Redfield, A. G. (1981) *Biochemistry 20*, 1147.
Kan, L. S., Borer, P. N., & Ts'o, P. O. P. (1975) *Biochemistry 14*, 4864.
Krol, A., Machott, M. A., Branlaut, C., & Ebel, J. P. (1978) *Nucleic Acids Res. 5*, 4933.
Krugh, T. R., & Young, M. A. (1975) *Biochem. Biophys. Res. Commun. 62*, 1025.
Krugh, T. R., Laing, J. W., & Young, M. A. (1976) *Biochemistry 15*, 1224.
Ladner, J. E., Jack, A., Robertus, J. D., Brown, R. S., Rhodes, D., Clarke, B. F. C., & Klug, A. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4414.
Lee, C.-H., & Tinoco, I., Jr. (1980) *Biophys. Chem. 11*, 283.
Lezuis, A. G., & Domin, E. (1973) *Nature (London), New Biol. 244*, 169.
Lomant, A. J., & Fresco, J. R. (1975) *Prog. Nucleic Acid Res. Mol. Biol. 15*, 185.
Mizuno, H., Tomita, K.-I., Nakagawa, E., Ohtsaka, E., & Ikehara, M. (1981) *J. Mol. Biol. 148*, 103.
Neilson, T., & Werstiuk, E. S. (1974) *J. Am. Chem. Soc. 96*, 2295.
Neilson, T., Romaniuk, P. J., Alkema, D., Hughes, D. W., Everett, J. R., & Bell, R. A. (1980) *Nucleic Acids Res. Symp. Ser. 7*, 293.
Patel, D. J. (1976) *Biopolymers 15*, 533.
Patel, D. J. (1977) *Biopolymers 16*, 1635.
Patel, D. J. (1979) *Biopolymers 18*, 553.
Quigley, G. J., Seeman, N. C., Wang, A. H.-J., Suddath, F. L., & Rich, A. (1975) *Nucleic Acids Res. 2*, 2329.
Reid, B. R., McCollum, L., Ribeiro, N. S., Abbate, J., & Hurd, R. E. (1979) *Biochemistry 18*, 3996.
Rich, A. (1977) *Acc. Chem. Res. 10*, 388.
Robillard, G. T., Hilbers, C. W., Reid, B. R., Gaugloff, J., Dirheimer, G., & Schulman, R. G. (1976) *Biochemistry 15*, 1883.
Romaniuk, P. J., Hughes, D. W., Grêgoire, R. J., Neilson, T., & Bell, R. A. (1978a) *J. Am. Chem. Soc. 100*, 3971.
Romaniuk, P. J., Neilson, T., Hughes, D. W., & Bell, R. A. (1978b) *Can. J. Chem. 56*, 2249.
Romaniuk, P. J., Hughes, D. W., Grêgoire, R. J., Bell, R. A., & Neilson, T. (1979) *Biochemistry 18*, 5109.
Sitz, T. O., Kuo, S.-C., & Nazar, R. N. (1978) *Biochemistry 17*, 5811.
Stannard, B. S., & Fesenfeld, G. (1975) *Biopolymers 14*, 299.
Studnicka, G. M., Rahn, G. M., Cummings, T. W., & Salser, W. A. (1978) *Nucleic Acids Res. 5*, 3365.
Tinoco, I., Jr., Uhlenbeck, O. C., & Levine, M. (1971) *Nature (London) 230*, 362.
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.
Turner, D. H., Petersheim, M., Albergo, D. D., Dewey, T. G., & Freier, S. M. (1981) in *Biomolecular Stereodynamics* (Sarma, R. H., Ed.) p 429, Adenine Press, New York.
Uhlenbeck, O. C., Baller, J., & Doty, P. (1970) *Nature (London) 225*, 509.
Uhlenbeck, O. C., Martin, J. H., & Doty, P. (1971) *J. Mol. Biol. 57*, 217.
Uhlenbeck, O. C., Borer, P. N., Dengler, B., & Tinoco, I., Jr. (1973) *J. Mol. Biol. 73*, 493.
Werstiuk, E. S., & Neilson, T. (1976) *Can. J. Chem. 54*, 2689.