PERTURBATIONS OF SHORT RNA HELICES

by

© DIRK ALKEMA, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

January, 1982

PERTURBATIONS OF SHORT RNA HELICES

Doctor of Philosophy (1982)
(Biochemistry)

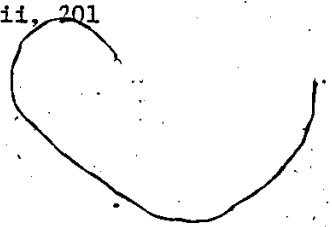
McMaster University
Hamilton, Ontario

Title: Perturbations of Short RNA Helices

Author: Dirk Alkema, B.Sc. (McMaster University)

Supervisor: Dr. Thomas Neilson

Number of Pages: xvii, 201



ABSTRACT

A variety of short oligoribonucleotide sequences were synthesized using the phosphotriester chemical synthesis developed in Neilson's laboratory. These sequences were designed to incorporate a variety of features that would aid in the study of perturbations of helix structure and stability. Variable temperature proton nuclear magnetic resonance (NMR) spectroscopy was used in this study and provides a powerful technique for the study of nucleic acid conformations and for the investigation of the effects of the mispairing and single stranded regions caused by the helix imperfections introduced.

The assignment procedure for NMR spectra was improved through the study of a series of related sequences. This study determined the effects that the addition of a nucleotide to the terminus of a sequence or the insertion of a nucleotide into the middle of a sequence would have on the chemical shifts from the rest of the sequence.

A series of self-complementary pentaribonucleotides, with a central non-base paired opposition (AGXCU, where X ≡ A, G, C or U), was studied to determine the effects of small loops on duplex stability. In contrast to earlier results, these pentamers formed stable duplexes when X ≡ A or C, although the duplex Tm's were significantly reduced. These sequences also provided the opportunity to study the sequence effects of adjacent internal G·C base pairs on duplex stability when the middle base pair was A·U, G·C or G·U. Comparison with earlier results using corresponding A·U base pair neighbours demonstrated the enhanced stabilization of G·C base pairs.

The effects of terminal non-base paired (dangling) adenosines were more closely investigated and found to contribute an average of 11°C to the thermal stability of the duplex formed. This study also demonstrated that 5'-dangling adenosines contribute less to overall stability than do 3'-dangling adenosines. However, this effect did display some sequence dependence.

The triribonucleotide GpCpA was the first trimer shown to form a stable RNA duplex (Tm = 33°C). The duplex consisted of two G·C base pairs and two 3'-dangling adenosines and had a stability equal to that of the tetramer duplex UpGpCpA with four Watson-Crick base pairs.

The influence of base stacking on duplex formation was studied and it was discovered that the direction of base stacking had a definite influence on helix stability. Stacking in the 5'→3' direction was more favourable to duplex formation than stacking in the 3'→5' direction.

Lastly, the significance of invariant adenosines at position 14 and 21, in the D-stem of tRNA, was investigated through a model study that suggested the adenosines contributed to D-stem stability.

Most of the results presented in this thesis have been published or accepted for publication in:

1. Jeremy R. Everett, Donald W. Hughes, Russell A. Bell, Dirk Alkema, Thomas Neilson and Paul J. Romaniuk. "Nearest-Neighbour and Next-Nearest-Neighbour Effects in the Proton NMR Spectra of the Oligoribonucleotides ApGpX and CpApX." (1980) Biopolymers 19, 557.

2. Thomas Neilson, Paul J. Romaniuk, Dirk Alkema, Donald W. Hughes, Jeremy R. Everett and Russell A. Bell. "The Effects of Base Sequence on the Stability of Short Ribonucleic Acid Duplexes." (1980) Nucleic

Acids Res. Sym. Series No. 7, 293.

3. Dirk Alkema, R.A. Bell, P.A. Hader and T. Neilson. "Triplet GpCpA Forms a Stable RNA Duplex." (1981) J. Am. Chem. Soc. 103, 2866.

4. Russell A. Bell, Jeremy R. Everett, Donald W. Hughes, Dirk Alkema, Paul Hader, Thomas Neilson and Paul J. Romaniuk. "Nearest-Neighbour and Next-Nearest-Neighbour Effects in the Proton NMR Spectra of the Oligoribonucleotides ApXpG, CpXpG, CpApXpUpG, ApGpXpC and ApGpXpCpU." (1981) Biopolymers 20, 1383.

5. Dirk Alkema, Russell A. Bell, Paul A. Hader and Thomas Neilson. "Short RNA duplex Stability: Contribution from non-base paired residues to the direction of stacking." (1981) in "Biomolecular Stereodynamics" R.H. Sarma, ed., Adenine Press, Elmsford, NY., p. 417.

6. Dirk Alkema, Russell A. Bell, Paul A. Hader and Thomas Neilson. "Invariant Adenosine Residues Stabilize tRNA D-stems." (1982) Accepted for publication in FEBS Letters.

7. Dirk Alkema, Paul A. Hader, Russell A. Bell and Thomas Neilson. "Effects of Flanking G·U Base Pairs on Internal Watson-Crick, G·U and Non-Bonded Base Pairs Within a Short RNA Duplex." (1982) Accepted for publication in Biochemistry.

Two additional publications will appear shortly:

1. Paul A. Hader, Thomas Neilson, Dirk Alkema, Eric C. Kofoid and M.C. Ganoza. "Sequencing of Short RNA Oligomers by Proton Nuclear Magnetic Resonance." (1982) Accepted for publication in FEBS Letters.

2. Paul A. Hader, Dirk Alkema, Russell A. Bell and Thomas Neilson. "Parameters for Proton Chemical Shift Prediction in Oligoribonucleotides." (1982) J. Chem. Comm. (in press).

ACKNOWLEDGEMENTS

My appreciation to Dr. T. Neilson for his encouragement and guidance. I would also like to thank the members of my supervisory committee, Dr. R.A. Bell and Dr. W.W.-C. Chan, for their suggestions.

I am grateful to Dr. D.W. Hughes and Dr. J.R. Everett for their contributions to the development of the NMR spectral assignment procedures. Discussions with Dr. P.J. Romaniuk and Mr. P.A. Hader were also appreciated.

I wish to acknowledge Dr. M. Smith and Dr. C. Rogers of Bruker Spectrospin in Mississauga for the 250 MH_z spectra and Dr. R. Lenkinski of the Southwestern Regional NMR facility, University of Guelph, for the use of the 400 MH_z spectrometer.

Special thanks to my wife, Sylvia, for her love and support.

TABLE OF CONTENTS

ABBREVIATIONS

A^a	adenosine
Å	angstrom (10^{-8} cm)
C^a	cytidine
CD	circular dichroism
Cu/Zn	copper-zinc couple
dA	2'-deoxyadenosine
dC	2'-deoxycytidine
dG	2'-deoxyguanosine
D_2O	deuterium oxide
DMF	N,N-dimethylformamide
DNA	deoxyribonucleic acid
G^a	guanosine
g	gauche
ΔH_1	enthalphy of single base pair formation
Hz	hertz (cycles/second)
IR	infra-red
J	coupling constant
MST	mesitylenesulfonyl 1,2,4-triazole
mRNA	messenger RNA
NMR	nuclear magnetic resonance
ORD	optical rotatory dispersion
ppm	parts per million
R_F	ratio of distance travelled by solute to that of solvent

RNA	ribonucleic acid
T	thymidine
t	trans
thp	tetrahydropyranyl
tlc	thin layer chromatography
T_m	melting temperature of a duplex at which 50% of the strands exist in the duplex form
TPS-Nl	2,4,6-triisopropylbenzenesulphonyl nitroimidazole
trac	trityloxacetyl
tRNA	transfer RNA
$tRNA_f^{met}$	formyl methionyl transfer RNA
$tRNA^{phe}$	phenylalanine transfer RNA
U [a]	uridine
UV	ultraviolet

[a] mononucleotides and oligoribonucleotides are abbreviated in the standard format (IUPAC-IUB Commission on Biochemical Nomenclature, 1970). pA represent 5'-adenylic acid, Cp represents 3'-cytidylic acid and CpA represents cytidylyl (3'-5') adenosine. Unless otherwise noted, oligoribonucleotides are numbered sequentially from the 5' terminus.

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Deciphering the structural conformations of ribo- and deoxyribonucleic acid molecules has occupied the time, energy and resources of a great many researchers. This is understandable considering these biopolymers form the biochemical basis of all life. A quote from Sarma's introduction to "Nucleic Acid Geometry and Dynamics" (1980) describes nucleic acids as "the archives in which the basic blueprints to the mechanism (of life) are preserved and xeroxed." However, despite intensive research and numerous attempts to present conformational models, a satisfactory overall structure that fits any given class of RNA or DNA has yet to be determined. Most transfer RNA sequences can be theoretically fitted to the secondary cloverleaf structure proposed by Holley et al. (1965), but require assumptions of terminal base pair stabilities, nearest-neighbour effects, as well as allowance for the contribution of non-complementary base pairs and loops to conformational stability (Ninio, 1979). A tertiary L-structure for tRNA has been interpreted from X-ray studies of yeast $tRNA^{Phe}$ (Kim et al. 1974; Robertus et al. 1974) and E. coli $tRNA_f^{met}$ (Woo et al. 1980) but is still disputed.

The molecular structure of nucleic acids is a compilation of the geometry of the individual nucleotide units and the interactions between the nucleotides. Structural features of nucleotides include: i) the furanose ring that is ribose for RNA and 2'-deoxyribose for DNA, ii) the heterocyclic bases, adenine (A), guanine (G), cytosine (C), uracil (U) and thymine (T), attached to the C-1' of the sugar ring in

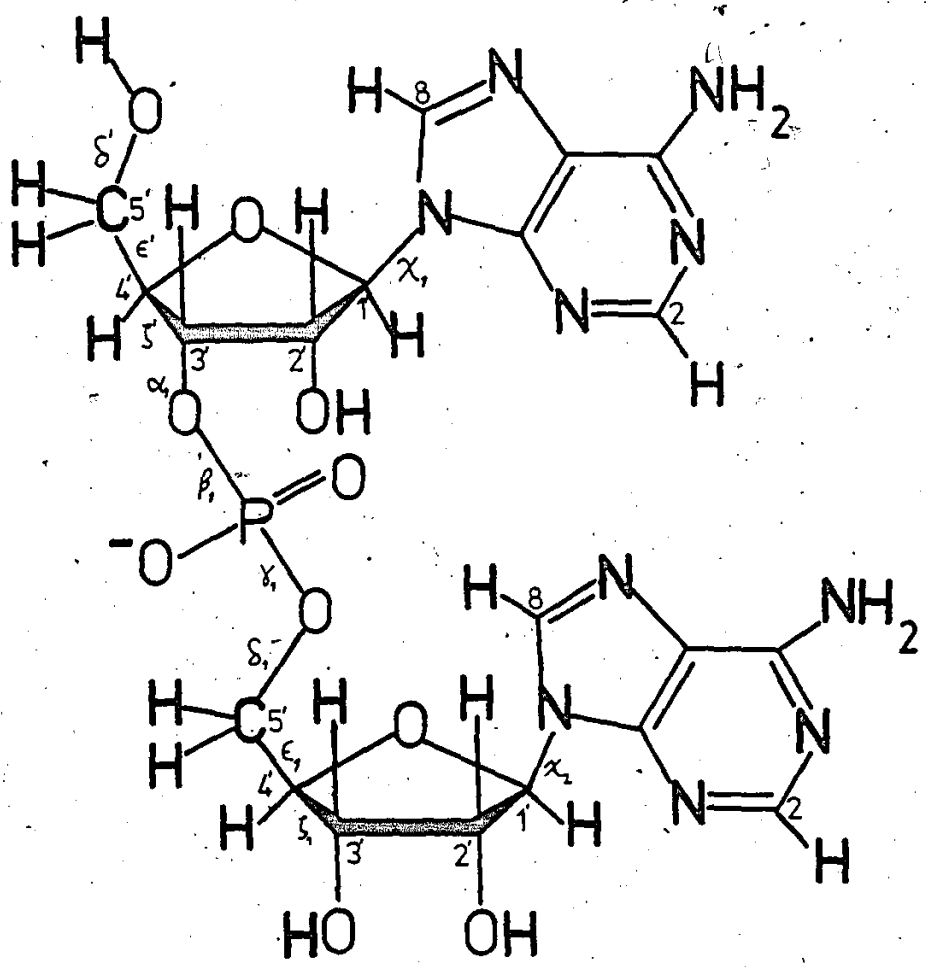


Figure 1. Structure of 3'-5' ApA, the numbering scheme and proposed IUPAC-IUB nomenclature for the various torsion angles (Sarma, 1980).

the β-configuration, and iii) the 3',5' phosphodiester linkage joining the individual nucleotide units. A complete conformational analysis of nucleic acid molecules requires the consideration of a number of bonds, within these structural features, about which rotation can occur (Figure 1). Calculation of the various torsion angles and determination of the preferred orientation or minimum energy conformation about these bonds is required. For a dinucleoside monophosphate this involves determining:

i) two sugar base torsions, i.e. glycosidic torsions, χ_1 and χ_2

ii) the mode of pucker of the two sugar rings

iii) torsion angles: α, β, γ, δ, ε_1

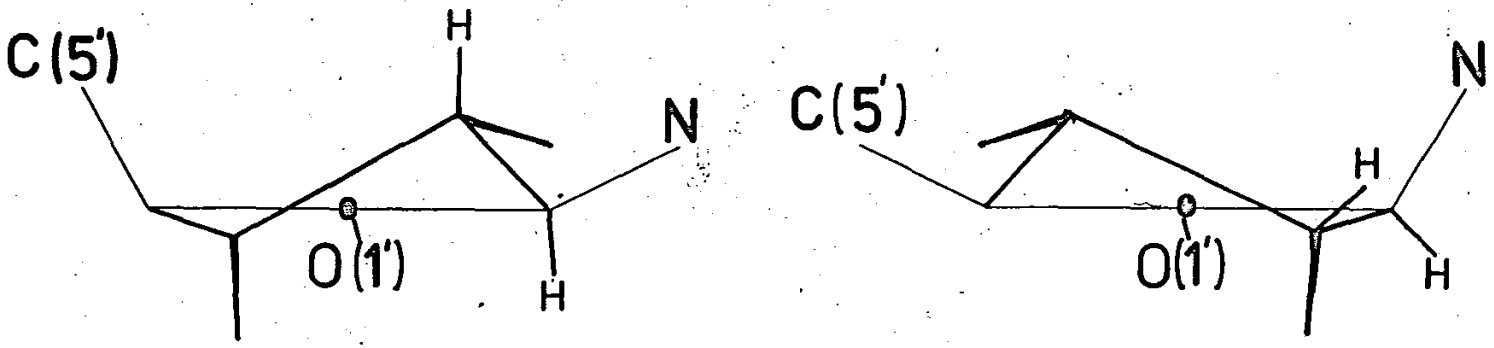
iv) torsion angles of the free exocyclic 5'-hydroxy methyl group, i.e. ε' and δ'

(Sarma, 1980)

In the case of a trinucleoside diphosphate this procedure becomes still more complex, with three glycosidic torsions, three sugar puckers and 13 other angles to consider. Calculations and variations for longer sequences soon become unwieldy.

The structural analysis is simplified by recognition of several general features. Single crystal data determined that the furanose ring exists, in the majority of cases, in a conformation where either the C-2' or C-3' atom is displaced toward the C-5' from the plane of the other furanose ring atoms (Sundaralingam,1969). When C-2' is displaced the conformation is designated as C(2')-endo and when C-3' is displaced as C(3')-endo (Figure 2) (Sarma, 1980). Figure 3 displays the Newman projections and the notation used to describe the rotations about the

Figure 2. The two major conformations of the sugar ring.



C(2')-endo $\rightleftharpoons$ C(3')-endo

$$10(10 - J_{1',2'}) = \% \text{ C(3')-endo}$$

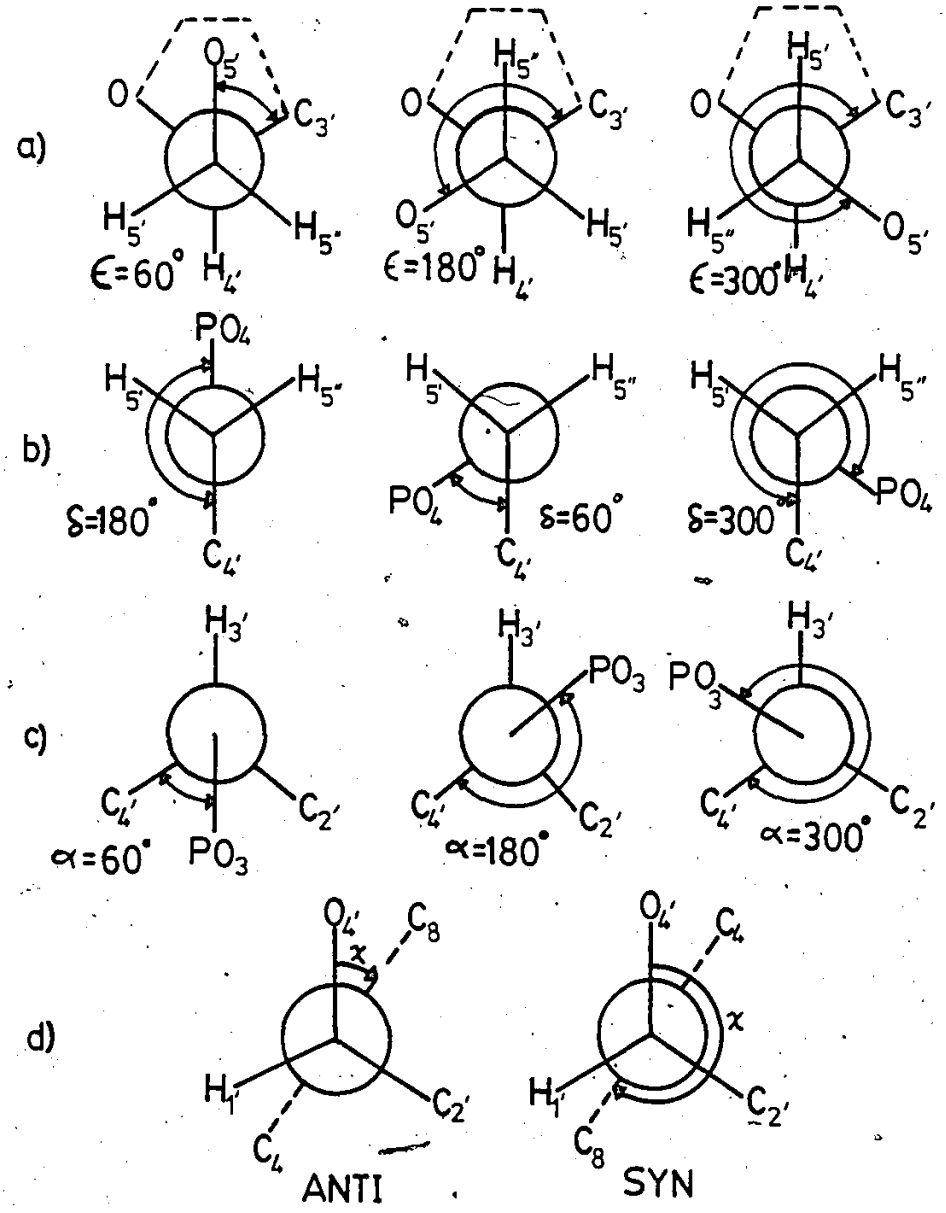


Figure 3. Schematic representation of the bond angles about the a) C(4')-C(5') (ε) bond, b) C(5')-O(5') (δ) bond, c) C(3')-O(3') (α) bond and d) the glycosidic linkage (χ) (Sarma, 1980).

C4'-C5'(ε), C5'-O5'(δ) and C3'-O3'(α) bonds and the glycosidic linkage (χ) of the various conformers. In addition, the position of the base in relation to the sugar is generally either anti (χ ≃ o ± 90°) or syn (χ ≃ 180° ± 90°).

Determination of the rotation about the various bonds is dependent upon interactions between the nucleotide units. Hydrogen bonding between complementary bases (G·C and A·U(T))(Figure 4), non-Watson-Crick base pairing (Crick, 1966) and non-complementary base oppositions in duplexes as well as vertical electronic interactions between the bases (base stacking) and hydrophobic bonding, all contribute to the preferred nucleic acid conformation.

Conformational analysis employs a number of techniques to study nucleic acids in both the solid and solution state. Solid state studies use X-ray analysis of crystals and fibres, semi-empirical energy calculations and model studies. Circular dichroism (CD), optical rotary dispersion (ORD), infra-red (IR), ultra-violet (UV), fluorescence, temperature-jump, calorimetric and nuclear magnetic resonance (NMR) techniques have all proved useful in solution state investigations.

1.1. SOLID STATE STUDIES

1.1.1. X-ray Studies

X-ray diffraction studies have played a crucial role in understanding the discrete structures formed by nucleic acid molecules by revealing the interatomic bonds and angles. X-ray diffraction of single crystals has led to the successful determination of the structures of a variety of oligonucleotides such as UpA (Sussman et al. 1972), GpC (Arnott et al. 1973), pApTpApT (Viswamitra et al. 1978) and CpGpCpGpCpG

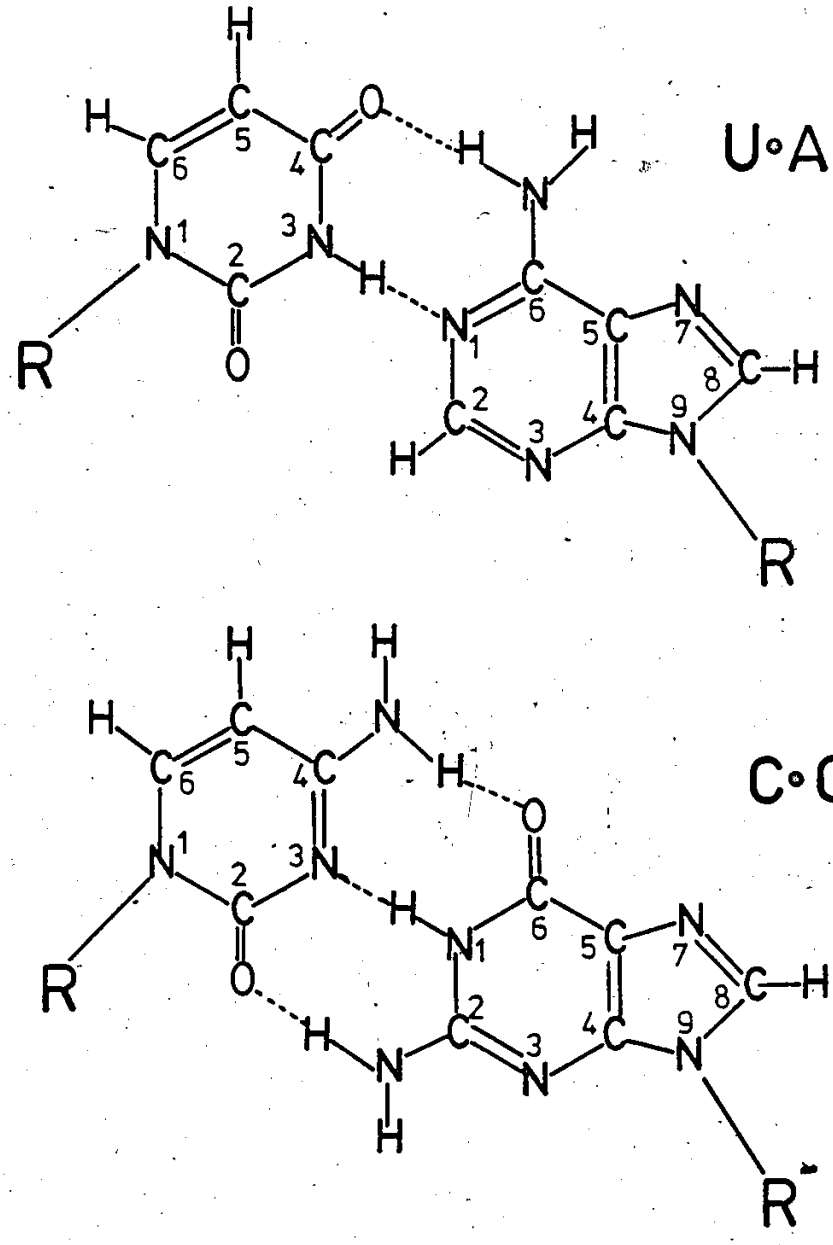


Figure 4. Representation of the U•A and C•G base pairs depicting the hydrogen bonding and the numbering system for the bases.

(Wang et al. 1979) as well as yeast $tRNA^{phe}$ (Kim et al. 1974; Robertus et al. 1974) and E. coli $tRNA_f^{met}$ (Woo et al. 1980).

Unfortunately, specific crystals crystallize in a conformation which is not necessarily the most favourable conformation in solution (Seeman, 1980a). Therefore, solution studies are required to complement crystal studies. However, solution techniques are most powerful in concert with static crystallographic structures which provide initial reference points.

Despite these limitations, X-ray analysis has provided valuable insights into the conformation of nucleic acid structures. Studies with RNA sequences indicated that the ribose ring was in the C(3')-endo conformation resulting in a double helix known as 'A'-type (Arnott et al. 1969). In the case of DNA, the sugar moiety adopts the C(2')-endo pucker distinguishing it as a 'B'-type double helix (Milman et al. 1967). Both helices are right-handed with the bases in the anti orientation and approximately 10 base pairs per turn in the B helix and 11 per turn in the A helix. Crystallographic evidence of B-type RNA helices has not been found and perhaps the 2'-hydroxyl is partly responsible for the observed C(3')-endo conformational preference (Kallenbach and Berman, 1977). These studies determined base tilt and base overlaps, important components of base stacking, are different for A- and B-type helices.

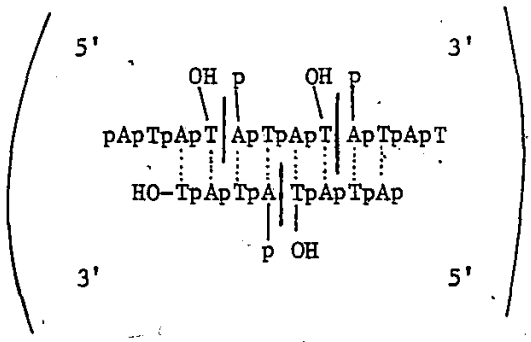
The first crystal structure of a naturally occurring dinucleoside monophosphate, UpA (Sussman et al. 1972; Rubin et al. 1972) contained two molecules incorporated within the lattice, neither of which adopted the standard Watson-Crick pairing. Apparently the acidic conditions used, protonated the N-1 of adenine preventing the expected base pairing. The

significant finding was that the conformations of each of the four nucleosides were virtually identical. The furanose rings were all in the C(3')-endo pucker, the bases were in the anti region (χ^{CN}) and the interphosphate torsion angles were very similar. In fact, the only major differences were in the ω torsion angle about the phosphodiester bond. These results supported Sundaralingam's concept of a rigid nucleotide (Sundaralingam, 1969, 1975): the various conformations of dinucleotides, trinucleotides, etc. could be interconverted by changes in ω, ω' and that the sugar pucker can exist only as C(3')-endo or C(2')-endo. Later studies, both solution and solid state, have cast doubt on this concept (Evans & Sarma, 1976; Pattabiraman et al. 1980).

The first deoxydinucleotide to undergo X-ray crystal analysis was d(pTpT) (Camerman et al. 1976). Although the backbone torsion angles were very similar to UpA, the structures were quite different because of a preferred C(2')-endo sugar conformation in d(pTpT) rather than the C(3')-endo pucker found in UpA.

The first crystallographic evidence that oligoribonucleotides form stable, Watson-Crick, right-handed, anti-parallel, double helices came from diffraction studies on the sodium salts of ApU (Rosenberg et al. 1973) and GpC (Arnott et al. 1973). Direct quantification of a number of important parameters associated with the RNA double helix were obtained. These helical parameters were slightly different owing to the different binding position of the sodium ion.

X-ray studies with longer sequences have also provided interesting results. The structure of d(pApTpApT) (Viswamitra et al. 1978) indicated the sequence formed a 'splint-like' double helix molecule.



A striking feature of this loose DNA structure was the existence of an alternating sugar conformation. The adenosines all had the C(3')-endo conformation, as found in normal RNA structures, whereas the thymidines exhibited the C(2')-endo conformation.

More recently the crystal structure of d(CpGpCpGpCpG) was published (Wang et al. 1979) revealing a left-handed helix called Z-DNA which contained 12 residues per turn. The alternating cytidines and guanosines were in different conformations with the guanosines adopting the less common "syn" conformation, with the C(3')-endo sugar pucker, while the cytidines are C(2')-endo and anti in the relationship of the base to the sugar. The 3'-terminal guanosines, however, adopted a C(2')-endo conformation as a result of an end effect. With normal Watson-Crick base pairing, the backbone torsion angles adopted conformations which resulted in a left-handed helix. This structure had no major groove and the helix axis was found in the minor groove.

It is valuable to note that methylation of cytidine, in poly(CpG), to give 5-methyl-cytidine (m^5C) resulted in the adoption of the Z-helix conformation at near physiological conditions (Behe & Felsenfeld, 1981). This result suggested the relatively high frequency of m^5C might be involved in transforming normal B-DNA to Z-DNA under the right conditions. Z-DNA has been suggested to be involved in the control of gene expression

(Behe & Fesenfeld, 1981) and has been found in restriction fragments and plasmids (Wells et al. 1981).

The sequence dependence of helix formation was studied using the self-complementary deoxy-dodecamer d(CpGpCpGpApApTpTpCpGCpG) (Dickerson & Drew, 1981). Apparently local helix parameters of twist, tilt and roll were more strongly influenced by sequence than by crystal packing. The central GAATTC segment was a B-helix flanked on either side by duplex regions having A-helix character. It was suggested that the topology of the various site sequences might be related to attack by different restriction enzymes.

The ability of X-ray diffraction studies to investigate non-Watson-Crick type base pairing was recently demonstrated by a study on r(GpGpCpU) (Mizuno et al. 1981). This RNA tetramer formed an A-RNA duplex structure with two terminal G·U base pairs and two internal G·C base pairs.

X-ray studies have also revealed the tertiary structure of tRNA, in particular yeast $tRNA^{phe}$ (Kim et al. 1974; Robertus et al. 1974) and E. coli $tRNA_f^{met}$ (Woo et al. 1980). Important facets of RNA conformational flexibility were shown. Although most of the nucleotides in this tRNA molecule adopt the C(3')-endo sugar conformation, several residues adopted the less common C(2')-endo conformation. However, all of these nucleotides were located either at sites where the chain undergoes a change in direction or where bulges occur in the polynucleotide chain (Rich et al. 1980).

Crystal studies of tRNA molecules have also confirmed the existence of G·U base pairs and tertiary base pair associations of unusual types (Ladner et al. 1975; Quigley et al. 1975).

1.1.2. Calculations of Minimum Energy Conformations

Calculations of minimum energy conformations are largely semi-empirical (Olson, 1978) but have provided important insights into the preferences of RNA sequences for an 'A' type helix while DNA duplexes are found as A or B-type helices (Broyde & Hingerty, 1978; Hingerty & Broyde, 1978). These manipulations also help to explain the relative stabilities of single strand and double strand helices (Broyde & Hingerty, 1978; Olson, 1978) and helix flexibility (Thiyagarajan & Ponnuswamy, 1978, 1979; Olson, 1980).

In general, the total potential energy is calculated as a sum of the non-bonded (E_{nb}), electrostatic (E_{es}) and torsional (E_{tor}) contributions to energy:

$$E_{total} = E_{nb} + E_{es} + E_{tor}$$

(Thiyagarajan & Ponnuswamy, 1978, 1979). The total potential energy is minimized by simultaneously varying all seven, relevant, dihedral angles, within a given nucleotide, until the conformation with the lowest potential energy is found (Thiyagarajan & Ponnuswamy, 1978, 1979). Alternatively, only a couple of the dihedral angles can be altered while maintaining other angles in fixed positions as indicated from X-ray data (Olson, 1978, 1980).

Calculations of the low energy conformers for the 'A' form of diribonucleoside monophosphates (Broyde & Hingerty, 1978) and for the 'A' and 'B' forms of deoxydiribonucleoside monophosphates (Hingerty & Broyde, 1978) revealed that they would form single strand helices (A -type for RNA and A and B -type for DNA) with sequence dependent geometries that tended to be narrower and more tightly wound than

duplexes found in X-ray studies. Additionally, the intrastrand stacking was generally greater in the calculated single strand than in the multiple strand structures. The DNA study indicated only small amounts of energy were required to promote changes needed to permit Watson-Crick base pairing (Hingerty & Broyde, 1978), whereas some of the RNA sequences required 9-12 kcal/mole to convert the single strand to the duplex (Broyde & Hingerty, 1978). These studies indicated that while the differences in the conformational angles between single stranded and double stranded helices were small, they translated into substantial differences in helix geometry.

A comprehensive study of ApA that produced 30 conformers for both the C(3')-endo and C(2')-endo sugar conformations found that while 17 C(3')-endo conformers lie within a range of 5 kcal/mole energy difference, only 7 C(2')-endo conformers lie within the same range (Thiyagarajan & Ponnuswamy, 1978). This was an indication of the flexibility of the ribonucleotide sequence in a C(3')-endo conformation as compared with the C(2')-endo form. Surprisingly, for the C(3')-endo conformation, the lowest energy conformer would promote loop formation in a duplex, whereas the normal A-RNA and Watson-Crick type conformations occurred at energy levels about 0.5 and 1.0 kcal/mole, respectively, above the lowest energy level.

For ApA with the C(2')-endo sugar pucker the best preferred conformation turned out to be an extended helical structure with no base stacking. This agrees with other studies that found RNA sequences, in an extended, non-base stacked form, were predominantly C(2')-endo (Lee et al. 1976; Ezra et al. 1977; Broyde & Hingerty, 1978).

A similar study with d(ApApA) also demonstrated that C(3')-endo was more flexible than C(2')-endo and that the lowest energy conformer would promote loop formation (Thiyagarajan & Ponnuswamy, 1979). This flexibility is essentially the result of small changes in torsion angles but readily demonstrates how severely limited rotations of a rigid backbone are magnified into the enormous flexibility of a long polynucleotide chain (Olson, 1980).

Potential energy calculations with various RNA dimers have indicated that the 'B' form of RNA subunits is energetically disfavoured, relative to the 'A' form, by at least 6 kcal/mole. This is due to the unfavourable contact between the 2'-OH of the 3'-ribose and the 5'-base in 'B' type, C(2')-endo conformations (Broyde et al. 1975a,b).

Recently, a complete, mathematical, conformational analysis of B-DNA resulted in a theoretical double helical model with flexibility limits consistent with the smoothly supercoiled DNA in chromatin (DeSantis et al. 1981a). The model proposed agreed with the reported crystal structure for d(CGCGAATTCGCG) (Wang et al. 1980).

1.2. SOLUTION STATE STUDIES

The overall aim of investigations into polynucleotide structure and factors affecting conformation is to be able to predict secondary and tertiary structures from the primary sequence. A computerized attempt to predict the structure of 5s rRNA, in cyanobacterium of Anacystic midulans from its primary structure, utilizing published values for base pairing energies, produced 30 different structures with a best possible fit of the duplex regions (Studnicka et al. 1978). The proposed secondary structures for the A protein gene (Fiers et al. 1975)

replicase gene (Fiers et al. 1976) of bacteriophage MS2 RNA and 16s RNA (Noller & Woese, 1981) were very complex and indicative of the problems faced in elucidating the secondary structures from primary sequences. These structures were proposed simply by computer maximizing the base pairing in the strands. However, many other factors which control the most stable conformation were ignored. For example, the consideration that widely separated regions within the same molecule can base pair equally, greatly increases the number of structures that can be formed (Krol et al. 1978)

Even transfer RNA, the smallest naturally occurring polynucleotide molecule, does not always conform to the accepted secondary cloverleaf structure of Holley et al.(1965), simply by maximizing base pair formation. However, the probability of tRNA cloverleaf structure, determined from the primary structure, can be improved by including the contributions from bulge loop regions, wobble and non-complementary base pairs to conformational stability, by re-evaluating the energies of terminal G·C and A·U base pairs and narrowing the gap between G·C/G·C and G·C/A·U contributions (Ninio, 1979).

Tinoco pointed out that in order to evaluate proposed secondary structures based on a known primary sequence, it is necessary to understand how the sequence of duplex regions, the placement and size of loops within a double helix, the presence of neighbouring single stranded regions and non-complementary base pairs, contribute to stability and overall secondary structure (Tinoco et al. 1971, 1973).

Solid state studies have provided part of this information but from a relatively static point of view. However, since polynucleotide

conformations are dynamic, a comprehensive understanding requires solution studies.

1.2.1. Optical Studies

Properties of the purine and pyrimidine base chromophores allows the use of UV hypochromicity, CD and ORD spectroscopy, temperature-jump and fluorescence techniques in the study of duplex formation between complementary oligonucleotides of defined sequence. Kinetic and thermodynamic parameters can be derived from the results.

Initial studies, on duplex formation, relied heavily on readily available homo-polynucleotides (for review, Ts'O, 1974). The later introduction of enzymatically and synthetically prepared sequences greatly advanced this research by providing an increased variety of sequences capable of duplex formation.

One of the first ORD studies, on the association of complementary triribonucleotides, showed that a stable duplex of the type thought to occur in codon-anticodon interactions could not form (Jaskunas et al. 1968) and therefore duplex stabilities were questioned. Also, oligoribonucleotides containing guanosine residues (i.e. GGC, GCC, AGC), self-aggregated or formed multi-strand aggregates. Subsequent CD investigations, with G-containing triribonucleotides, indicated that, in the absence of complementary bases, guanosines promoted destacking (Brahms et al. 1969). This effect was particularly pronounced whenever non-base paired guanosines were found next to uridine residues, suggesting possible sequence effects (Brahms et al. 1969). CD spectra, of sequences containing 5'-uridines with different nearest neighbours, suggested UpA and UpG sequences were essentially unstacked in single-stranded RNA

(Gray et al. 1972). Destacking sequences could conceivably influence the structures and functions of natural RNA molecules.

Results of this nature prompted closer inspection of the effects of base sequence on duplex formation and stability. An ultra-violet examination of a series of block oligoribonucleotides, of the type $(Ap)_m(Up)_{n-1}U$ (where m and n vary individually from 3 to 7) indicated the individual oligomers formed bimolecular helical complexes (when m>n) or multichain aggregates (when m<n) (Martin et al. 1971). The temperature dependence of the perfect duplexes (m=n) yielded enthalpy changes per base pair, ΔH_1, that increased from -6.4 kcal/mole base pair for $(Ap)_4(Up)_3U$ to -7.6 kcal/mole base pair for $(Ap)_7(Up)_6U$. The chain length dependence of ΔH_1 was attributed to end effects and/or variations in the stacking energies of the adenosine sequences (Martin et al. 1971).

A subsequent study (Uhlenbeck et al. 1971) demonstrated the insertion of a cytidine or guanosine between blocks of adenosine and uridine sequences, $(A)_nX(U)_n$ (where X = C or G), destabilized any helix formation. Differences in the T_ms from the various duplexes, indicated cytidine was more destabilizing than guanosine and equivalent to the removal of two A·U base pairs (T_m decreased by 7°C). It was suggested that guanosine was less destabilizing because the duplex could rearrange to accommodate guanosine opposite uridine:

i.e. ApApApApGpUpUpUpU
:::::::::
UpUpUpUpGpApApApA

(Uhlenbeck et al. 1971). Equimolar mixtures of A_nCU_n and A_nGU_n resulted in a duplex with one internal G·C base pair with a T_m that increased by 8.5°C (Uhlenbeck et al. 1971).

Follow-up studies by these workers (Porschke et al. 1973), established the stability of duplexes containing G·C base pairs was greater than a corresponding A·U duplex of the same length. Moreover, this research demonstrated helices of the same base composition, but different sequence, have quite different stabilities. Determination of the kinetics of double helix formation, from temperature-jump experiments, revealed the recombination of the strands to be second order, with rate constants dependent on chain length and nucleotide sequence (Porschke et al. 1973). In addition, nucleation of duplex formation occurred, preferentially, at the sites of G·C base pairs.

Two comprehensive studies (Borer et al. 1973, 1974) of 25 perfect duplexes, with chain lengths of 6 to 14 base pairs, determined the ΔG, ΔH and ΔS values of formation for all 10 possible Watson-Crick base paired nearest-neighbour sequences and found a striking sequence dependence of helix stability. The results were used to develop an equation that could predict the T_m of any known sequence and has been used to predict the secondary structure of RNA from primary sequence (Tinoco et al. 1973).

Recently, the effects of sequence, on CD spectra, with 18 synthetic polyribonucleotides were investigated (Gray et al.1981). This work extended and refined the findings by Borer et al. (1973 & 1974) and was able to approximate the CD spectra of all other RNA sequences by first-neighbour formulas. Relationships for the calculation of T_ms for any RNA or DNA sequence were also derived:

$$T_m(\text{DNA}) = 44.0 + 36.1\ H_G + 9.2\ H_G^2$$

$$T_m(\text{RNA}) = 42.8 + 58.4\ H_G + 11.8\ H_G^2$$

where H_G is the fraction of G·C base pairs in the sequence (Gray et al. 1981).

Unfortunately, the equations presented in these studies had a limited application to smaller sequences. Romaniuk (1979) found that the formula developed by Borer predicted a T_m of -3°C for the duplex formed by r(CAUG), whereas the actual T_m, from experimental data, was 24°C. Similarly, calculation of a T_m for duplex, AGGCU:AGCCU, using the equation for RNA developed by Gray et al. (1981), resulted in a T_m of 82°C, whereas the experimental T_m was 54°C (Section 4.2).

The use of the equations presented in both studies is probably limited because the techniques employed were not capable of evaluating individual nearest-neighbour and next-nearest-neighbour interactions. End effects constitute an additional problem.

1.2.2. Nuclear Magnetic Resonance Studies

Optical studies have significantly advanced the understanding of nucleic acid solution conformations. However, the usefulness of these techniques is limited by their ability to monitor only changes in gross structure or overall stability as an average change throughout the entire molecule, even though the change may be at one specific location. On the other hand, nuclear magnetic resonance (NMR) spectroscopy can observe changes in local micro-environments of nucleic acid molecules by following anisotropic shifts of specific resonances. In this way it is possible to focus on local modifications which alter stabilities and/or conformations.

Early NMR investigations of oligonucleotide conformations were limited to base stacking and sugar ring puckers by analyzing base proton chemical shifts and anomeric proton spin coupling constants. The complexity of the spectra, as well as the poor signal to noise ratio,

prevented complete analyses. With the advent of high frequency Fourier transform NMR systems, multi-nuclear capabilities, highly sophisticated interactive computer simulation capabilities and the development of selective deuteration techniques, hetero- and homo-decoupling, nuclear Overhauser effects and solvent suppression techniques, many of the problems associated with spectral analysis were resolved.

NMR has developed into a powerful tool for the structural elucidation of nucleic acid molecules because of its ability to monitor the different atoms involved in determining the molecular structure. Phosphorus (^{31}P) NMR has been used to examine the structural aspects of the phosphate backbone of nucleic acids (Patel & Canuel, 1976; Patel, 1980; Lerner & Kearns, 1981). Changes in the ω ω' phosphodiester bonds from the gauche-gauche (helical stacked) to the gauche-trans (coil structure) were reflected by upfield shifts of the ^{31}P resonances (Patel & Canuel, 1976). ^{13}C-NMR spectra, of ^{13}C-enriched sequences, are used to study secondary and tertiary structures of tRNA (Schmidt et al. 1980), as well as for observing the conformational dynamics of DNA, tRNA and other polynucleotide sequences (Bolton & James, 1980), by following the motion of the carbon atoms. Studies of this nature revealed that furanoside carbons were more mobile than hetero-base carbons and that DNA conformations were far more diverse than RNA conformations. This difference between DNA and RNA was thought to be a manifestation of the conformational mobility of their respective sugar rings (Bolton & James, 1980).

While ^{31}P and ^{13}C NMR studies provide valuable structural data about nucleic acids, by far the most widely used NMR, in nucleic acids,

is proton NMR (^{1}H-NMR). ^{1}H-NMR can observe base pairing in duplexes by monitoring the exchangeable ring N-H protons involved in the hydrogen bonding of base pair formation. In addition, ^{1}H-NMR can examine the conformation of nucleic acid furanosides under various conditions by measuring the various proton coupling constants, but especially the $J_{1'-2'}$ coupling constant. Changes in the chemical shifts of the non-exchangeable base protons are indicative of changes in the orientation of the nucleotide units within a helix.

When the temperature increases the ring nitrogen protons involved in base pair hydrogen bonding, exchange with deuterium atoms from D_2O and H-bonding is disrupted. The N-H resonances disappear with increasing temperature and can be assigned (Kearns, 1971, 1976; Kan et al. 1975). These exchangeable protons are normally found in the very low-field region (11-15 parts per million) of the spectrum (Kearns, 1971, 1976; Hurd & Reid, 1979) because the ring N-H resonances are significantly deshielded by H-bonding (Katz & Penman, 1966), providing a relatively easy means for observing base pair formation. More important, each base pair contains only one ring N-H hydrogen bond and therefore the number of resonances found in this low-field region of the spectra corresponds to the number of stable base pairs formed in solution (Reid et al. 1979).

Monitoring the exchange rates of these N-H bonds was used extensively to explore the secondary and tertiary structures of tRNA. The stabilities of the different duplex regions of tRNA were found to vary significantly, with the D-stem being the least stable (Hilbers & Shulman, 1974; Kearns, 1976). ^{1}H-NMR experiments with tRNA molecules indicated there were 19-20 secondary base pairs and 6-7 tertiary base

pairs per molecule (Reid et al. 1979). The use of nuclear Overhauser effects (NOE) (Johnston & Redfield, 1978, 1981; Hurd & Reid, 1979), paramagnetic ion effects (Hurd et al. 1979) and saturation recovery techniques (Johnston & Redfield, 1977; Johnston et al. 1979; Hurd & Reid, 1980) has led to the identification of all possible base pairs formed in several tRNAs. ^{1}H-NMR probes into the effects of the tertiary base pairs and their location in the tRNA molecule (Johnston & Redfield, 1977, 1978, 1981; Reid et al. 1979; Hurd & Reid, 1979a,b), has confirmed the three dimensional tRNA structure, proposed from crystal studies (Kim et al. 1974; Rich & Schimmel, 1977). A review by Reid (1981) summarizes NMR investigations on tRNA.

The non-exchangeable protons in nucleic acids include the AH-8, AH-2, GH-8, CH-6, UH-6, CH-5 and UH-5, on the base, as well as the various CH ribose protons, but especially the anomeric carbon H-1' (Figure 4). The non-exchangeable protons produce well defined signals at all temperatures, unlike the N-H protons, making it possible to observe structural changes above and below the melting temperature of the helix. The temperature dependence of the non-exchangeable resonances has been used to examine the melting behaviour of a number of duplexes (Arter et al. 1974; Borer et al. 1975; Krugh et al. 1976; Romaniuk et al. 1979). Analysis of these resonances was instrumental in resolving the conformational anomalies of a number of single strand helices including r(CCA) (Cheng et al. 1978), r(AAA) (Evans & Sarma, 1976), CAXUG (where X = A, G, C, U) (Romaniuk et al. 1979), AUCCA (Stone et al. 1981); the structures of 15 dinucleoside monophosphates (Lee et al. 1976; Ezra et al. 1977), as well as contributions of bulge loops (Romaniuk et al. 1979;

Lee & Tinoco, 1980), terminal dangling residues (Romaniuk et al. 1979; Neilson et al. 1980) and base sequence (Romaniuk et al. 1979; Bubienko et al. 1981), to helix stability. Changes in the chemical shifts of these nonexchangeable resonances have been used to monitor the effects of oligomer concentration (Krugh & Young, 1975; Cheng et al. 1978; Romaniuk et al. 1979; Bubienko et al. 1981), pH (Ezra et al. 1976), salt concentration (Borer et al. 1975) and metal ions (Ezra et al. 1976) on duplex formation and stability.

Measuring the various furanoside ring proton coupling constants reveals much about the structural conformation of the ribose and its effects on the helix. The temperature dependent $J_{1'-2'}$ coupling constant has been used to explore the helix conformation in both single (Lee et al. 1976; Ezra et al. 1977) and double strands (Bubienko et al. 1981). The large $J_{1'-2'}$ value (5-6 Hz) at high temperatures (70-80°C) decreases as the temperature is lowered, indicative of a shift in ring conformation from the C(2')-endo (unstacked, extended coil) to the C(3')-endo (stacked, compact helix) sugar pucker (Lee et al. 1976; Ezra et al. 1977).

Changes in the chemical shift of the non-exchangeable proton resonances also reflect the influence of the phosphodiester linkage on conformation (Kondo et al. 1970; Ezra et al. 1977) and have been analyzed to help resolve such diverse concepts as nucleotide rigidity (Evans & Sarma, 1979) and flexibility (Early & Kearns, 1979; Neumann et al. 1979).

Clearly, nuclear magnetic resonance analyses play an important role in elucidating nucleic acid structure. NMR can also investigate interactions during drug intercalation with nucleic acids (Jones & Kearns, 1975; Patel, 1976; Patel and Canuel, 1977, 1979; Kastrup et al.

1978). The interactions of a DNA helix-destabilizing protein, the gene-5-protein of E. coli phages (M13f1 and fd), with an octadeoxyribo-adenylic acid sequence (dA8) have also been explored by NMR (Alma et al. 1981). The dynamic properties of various DNA restriction fragments have been similarly examined by NMR (Early et al. 1980; Haasnoot et al. 1980; Early et al. 1981a,b).

1.3. STUDIES ON LOOP FORMATION

The secondary structure of naturally occurring RNA molecules consists of approximately 50% duplex regions, that stabilize the overall structure, separated by single stranded regions that can form loops. These loops appear in three different forms.

1.3.1. Bulge Loops

Bulge loops consist of non-base paired residues that may adopt an extrahelical or intrahelical conformation within one of the duplex strands. A review of bulge loop studies, with homopolymer-copolymer mixtures (Lomant & Fresco, 1975), indicated that bulge loops preferentially adopted the extrahelical conformation regardless of the type of non-bonded opposition. The extrahelical conformation was favoured entropically and permitted greater solvation due to increased H-bonding with the solvent molecules. In addition, the extrahelical position required the least backbone distortion (Lomant & Fresco, 1975). The extent of duplex destabilization was dependent on the relative stacking tendency of the looped-out residue (decreasing in the order, A>G>C>>U); i.e. the ability to stack with nearest neighbours in the duplex strand (Lomant & Fresco, 1975). Most destabilization occurred when the stacking interactions were interrupted. However, the least destabilization

occurred if the loop extended stacking. It was noted that destabilization by an internal mismatch increased with length, but was proportionately the largest for a single residue (Lomant & Fresco, 1975). One derived value on the effect of a single 'looped-out' base on duplex stability was an addition of +2.5 kcal/mole to ΔG (Fink & Crothers, 1972).

NMR analysis of the complex formed by r(GUG) and r(CC), in the presence of ethidium bromide, suggested a stable RNA mini-helix with two G·C base pairs, the uridine residue bulged outside the helix and ethidium bromide intercalated between the G·C pairs (Lee & Tinoco, 1978). The complex strongly supports the potential for formation of extrahelical loops in small duplexes and indicates a model for frameshift mutations. Later studies (Lee & Tinoco,1980) of 13 triribonucleoside diphosphates, all with the sequence purine-pyrimidine-purine, and unable to duplex, showed these sequences could form single strand helices with the pyrimidine bulged-out and the two purines stacked on each other.

These studies suggested that helix stability was determined by maximizing base stacking interactions. In the case of bulge loops this requires that the residue adopt an extrahelical conformation (Lomant & Fresco, 1975).

1.3.2. Internal Loops

The conformational considerations for bulge loops apply equally well to internal loops (Lomant & Fresco, 1975). Although most of the information on the conformation of internal loops comes from homopolymer-co-polymer mixing experiments (Lomant & Fresco, 1975), they could not quantitatively provide the thermodynamics of loop formation. This information was obtained from studies with enzymatically prepared sequences.

One such study showed that when cytidine was inserted into the self-complementary duplex r(AnUn), the stability of the resulting bimolecular helical complex was reduced. The effect of the two opposing non-base paired cytidines on the T_m of the duplex was equivalent to the loss of two A·U base pairs (Uhlenbeck et al. 1971).

Oligomer r((A)$_4$G(C)$_n$(U)$_4$), formed imperfect dimer helices, when n > 4, with the internal loop becoming more destabilizing as the loop size increased (Gralla & Crothers, 1973b). Comparison with the study by Uhlenbeck and co-workers (1971) revealed the free energy of loop closure was less for a G·C base pair than that for an A·U base pair, suggesting the destabilizing effect of a loop could be sequence dependent (Gralla & Crothers, 1973b).

NMR investigations of internal loops have provided some interesting results. For example, the trimer, r(CUG), formed a stable RNA mini-helix composed of two C·G base pairs, with the internal opposing uridines looped out, in the presence of ethidium bromide (Lee & Tinoco, 1978). NMR studies with the deoxy dodecamer, d(ATCCTA(T)$_n$TAGGAT), (Cornelius et al. 1979; Haasnoot et al. 1980) produced interesting and contrasting results. When n = 1, a dimer structure formed,but the opposing thymidines appeared to form a wobble base pair rather than adopting the expected extrahelical conformation (Cornelius et al. 1979). Although this internal T·T wobble pair significantly lowered the duplex stability, a normal B-DNA structure was maintained. Also,when n = 3, 4 or 5, a hairpin structure formed, while for n = 2 both structures were found (Haasnoot et al. 1980). Another interesting point from the DNA study was the effect of the non-bonded residues in the loop, on duplex

stability, appeared to be strictly local. No influence on the stability of distant duplex regions was observed (Haasnoot et al. 1980).

A proton NMR examination of the pentaribonucleotide series, CAXUG (where X = A, G, C or U), each of which could theoretically form a duplex with an internal mismatched base pair, showed no duplex formation occurred over the temperature range 10°-70°C (Romaniuk et al. 1979). Instead, the chemical shift versus temperature plots indicated that stacked single strand helices formed without looped-out bases (Romaniuk et al. 1979).

1.3.3. Hairpin Loops

Hairpin loop formation was examined in a thermodynamic study of oligoribonucleotide series, $A_6C_mU_6$ (m = 4, 5, 6), using UV and CD (Uhlenbeck et al. 1973). Data indicated the most stable loop structure occurred when m = 6. The smaller loop structures were strained and might require base pair disruption to form a stable loop structure, whereas with the larger loops (m = 6) the probability of loop closure diminished (Uhlenbeck et al. 1973). Hairpin loop stability might also be dependent on the sequence of the loop, since there was a loss in the enthalpy of formation, thought to be a result of destacking of the loop bases (Uhlenbeck et al. 1973).

Kinetic studies of hairpin loop formation, with the sequence $A_4G(C)_nU_4$, demonstrated that nucleation of loop closure (intramolecular, n = 5, 6) was 10^3 fold faster than intermolecular (n = 4) nucleation (Gralla & Crothers, 1973a). In this case, the loop was closed by G·C base pairs, rather than A·U base pairs, resulting in a lower free energy of loop closure (Gralla & Crothers, 1973a).

1.4. Sequence Dependence of Duplex Stability

The effects of base sequence must also be considered in determining helix stability and structure. Consideration of the number of H-bonds in the two normal Watson-Crick base pairs, A·U with two H-bonds and G·C with three H-bonds, suggests that G·C base pairs will contribute more to duplex stability than A·U base pairs. These contributions to duplex stability were confirmed by comparison of the T_ms of the homopolymers poly(A·U) and poly(G·C), which are 38°C and 97°C, respectively (Gray et al. 1981).

Inserting a G·C base pair into a duplex of 10 A·U base pairs, enhanced the T_m by 8.5°C and two G·C base pairs doubled the effect (Uhlenbeck et al. 1971). A subsequent study (Porschke et al. 1973) showed that the stability of duplexes containing G·C base pairs was significantly higher than that of A·U duplexes of the same length. Further observation showed that sequences with the same base ratios and length had different stabilities. The enthalpies (ΔHs) of formation for these sequences were the same, indicating helix stability was entropy dependent and that differences were due to the effects of single strand stacking (Porschke et al. 1973).

Tinoco et al. (1973) found two adjacent G·C base pairs were considerably more stable than two G·C base pairs separated by an A·U base pair in a sequence of the same length. Comparison of the NMR studies with tetramer, r(CAUG) (Romaniuk et al. 1978), which exhibited a T_m of 24°C, and tetramer, r(AGCU) (Neilson et al. 1980), with a T_m of 34°C, supports this finding. Borer et al. (1974) pointed out the more subtle effects of sequence when it was demonstrated that $\overset{\rightarrow}{\underset{\leftarrow}{GG}}$ sequences were more

stable than $\overset{\rightarrow}{\underset{\leftarrow}{GC}}$ which were in turn more stable than $\overset{\rightarrow}{\underset{\leftarrow}{CG}}$ sequences. This interesting aspect of sequence dependent helical stabilities is inspected more closely in following sections.

Investigations have determined that nucleation, the initiation of duplex formation, occurs preferentially where G·C base pairs are located (Kan et al. 1975; Borer et al. 1975) and on melting, A·U(T) base pairs dissociate before G·C base pairs (Kan et al. 1975; Haasnoot et al. 1980; Early et al. 1981). All these results are consistent with the observation that the duplex stability is proportional to the percentage of G·C base pairs present (Krugh et al. 1976; Gray et al. 1981).

1.4.1. Effects of non-Watson-Crick Base Pairs

The wobble hypothesis, proposed by Crick (1966) to explain the degeneracy of the genetic code, suggested a number of non-Watson-Crick type base pairs could exist. The most important of these is the wobble G·U base pair, The existence and biological importance of G·U base pairs have now been discussed in a number of papers.

Early studies on the association of the codon GUG with the complementary anticodon, CAU, of f-met-tRNA (Uhlenbeck et al. 1970) as well as CD studies of poly(G·T) (Lezius & Domin, 1973) and poly(G·U) (Gray & Ratcliff, 1977) suggested the existence of G·U base pairs. However, it was the use of NMR that firmly established G·U base pairs with two hydrogen bonds occurred in both short ribonucleotides (Romaniuk et al. 1979) and tRNA molecules (Robillard et al. 1976; Hurd & Reid, 1979b), and G·T base pairs in poly(G·T) (Early et al. 1978). NMR also confirmed that G·U and G·T base pairs adopted the wobble conformation proposed by Crick (1966) (Early et al. 1978; Reid et al. 1979).

Clarke (1977) proposed that a G·U base pair could be accommodated by an RNA double helix without causing gross helical distortion. However, it would be a point of potential weakness because a shift is required, in the glycosyl torsion angles from the angles usually associated with Watson-Crick base pairing, for two hydrogen bonds to be made (Mizuno & Sundaralingam, 1978). The NMR study of the short RNA duplex, CAGUG:CAUUG, which contains an internal G·U base pair, confirmed this proposal (Romaniuk et al. 1979). The resultant duplex had a T_m of 24°C, with stability equivalent to that of the self-complementary duplex r(CAUG) even though there was an extra base pair. The G·U base pair did not add any apparent stability to the duplex. Earlier work had indicated that the stability of a terminal G·U base pair was equivalent to a terminal A·U base pair (Uhlenbeck et al. 1970). However, comparison of the duplex, CAGUG:CAUUG, with the duplex formed by CAAUG:CAUUG, with an internal A·U base pair, indicated the internal G·U base pair contributed less to helix stability than an internal A·U base pair (Romaniuk et al. 1979). However, it should be noted that the G·U base pair in this sequence was flanked on either side by A·U base pairs which are less stable than G·C base pairs and, therefore, probably less tolerant of the unusual conformation of the G·U base pairs. This is a problem that must be examined more closely since it may be that the strength of the G·U base pair is sequence dependent.

Next to G·U base pairs, A·C base pairs were thought to be the most stable (Lomant & Fresco, 1975; Grosjean et al. 1978; Ninio, 1979). Accurate predictions of nucleic acid structures require further investigations of the contributions to helical stability from non-

complementary oppositions such as T·T (Cornelius et al. 1979) and others, as found in the duplex regions of tRNA (Ninio, 1979).

1.5. Effects of Dangling Bases on Duplex Stability

The effects of single stranded regions, found adjacent to double stranded regions, on duplex stability, has also received some attention. Investigations of this phenomenon have used synthetically prepared sequences that will form duplexes with one or more non-base paired (dangling) terminal residues:

i.e. $\overset{\rightarrow}{\underset{\leftarrow}{\text{CAUGA}\atop\text{AGUAC}}}$ 3'-dangling adenosines

The positive contribution of a dangling residue to duplex stability was suggested in an investigation of A·U block polymers, formed by $r(A_m U_n)$ that discovered the T_m of the duplex was increased when $m > n$ and that the increase was approximately additive for one to three dangling adenosines (Martin et al. 1971).

Kinetic studies, employing temperature-jump methods, demonstrated that duplexes formed by two tRNA molecules, complementary anticodons, were stronger by a factor of 10^6, than expected for a three base pair duplex (Grosjean et al. 1976, 1978). Partial fragmentation of these duplexes concluded that a 10^3-fold portion of this increased stability could be attributed to the additional base stacking of the single strand regions adjacent to the double helix (Grosjean et al. 1976).

Examination of the duplexes formed by r(CAUGA) and r(CAUGU) (Romaniuk et al. 1978b) determined that a single 3'-dangling base significantly stabilized the duplex. The effect of a dangling adenosine was equivalent to a terminal A·U base pair, while the effect of a dangling uridine was about half that of adenosine. The stabilizing

influence of dangling bases was attributed to the increased stacking interactions that resulted throughout the duplex (Romaniuk et al. 1978b).

A recent thermodynamic determination of the effects of both 3'- and 5'-dangling adenosines on the duplex formed by r(CCGG), indicated the extra adenosine stacks increased the T_m by at least 10°C (Turner et al. 1981). The magnitude of the ΔH values obtained suggested base stacking was a major contributor to duplex stability. In addition, the magnitude of the enthalpy was different for 5'-dangling and 3'-dangling adenosines, intimating a possible sequence effect on the stacking interactions (Turner et al. 1981).

1.6. Effects of Base Stacking on Duplex Stability

Throughout the investigations on factors affecting the structural conformations of nucleic acids, outlined in this thesis, reference has been made to the importance of base stacking interactions in determining helix stability. Base stacking interactions essentially involve vertical electronic interactions between any two parallel bases (Figure 5) and includes; dipole-dipole, dipole-induced dipole and induced dipole-induced dipole forces. Base stacking is dependent on the interplanar distances between bases, however, the best geometrical overlap between the bases need not necessarily mean the most favourable stacking interaction has been found (Gupta & Sasisekharan, 1978a). Minimum energy calculations of a variety of deoxy-diribonucleotides (Gupta & Sasisekharan, 1978b; Hingerty & Broyde, 1978) and diribonucleotides (Broyde & Hingerty, 1978) indicated base stacking effects were also dependent on helix geometry and sequence.

The important role of base stacking in nucleic acid conformation

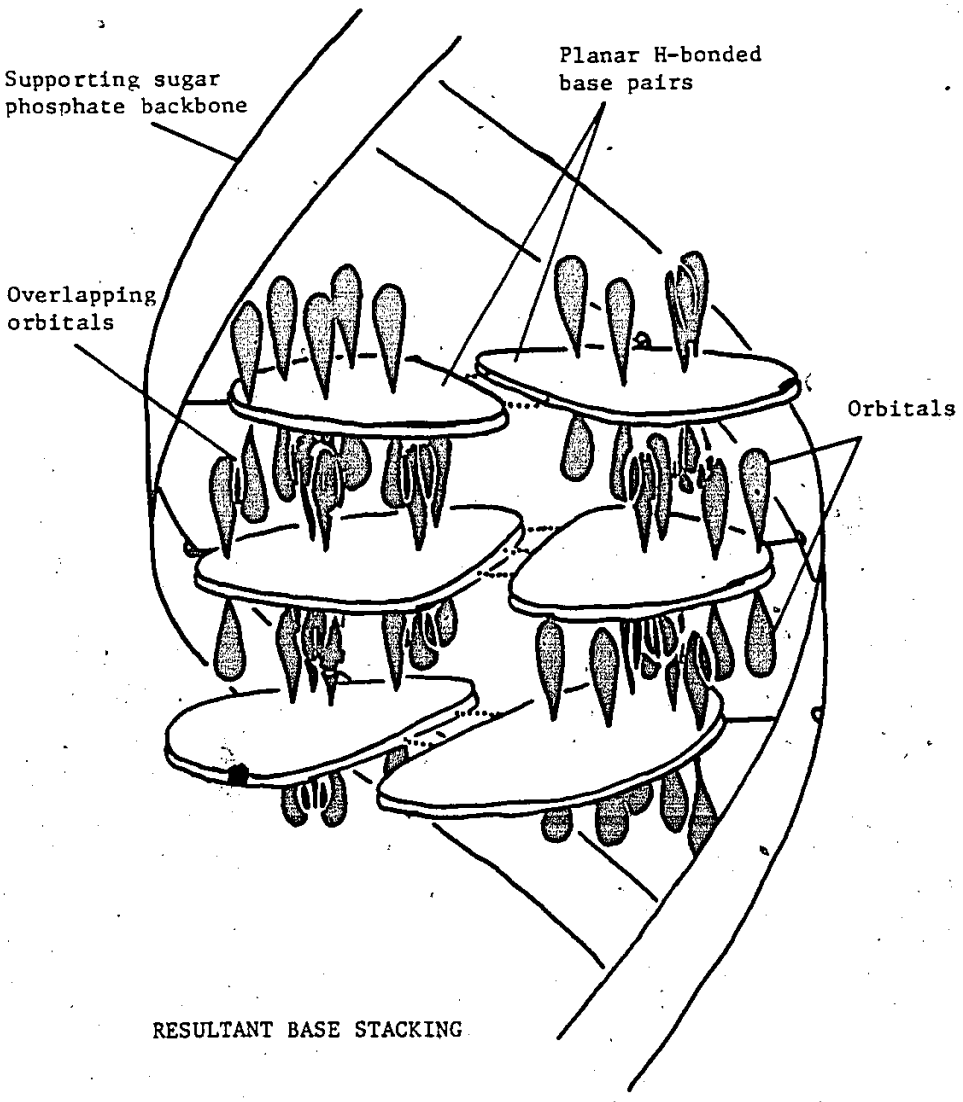


Figure 5. Diagramatic representation of the vertical electronic interactions, between parallel bases, inherent in base stacking.

and biological function has been suggested by several studies on the anticodon arm of tRNA. Removal of the Y-base, found adjacent to the anticodon sequence, changed the coding properties of yeast $tRNA^{phe}$ (Ghosh & Ghosh, 1970). It was suggested that the removal of the modified base altered the stacking pattern and the anticodon loop configuration.

As a result of the codon-anticodon interaction, an allosteric rearrangement of the tertiary structure of $tRNA^{phe}$ occurred, exposing the TψCG sequence for binding with the CGAA sequence of the 5s rRNA in the 50s subunit (Swartz et al. 1976). Further investigation revealed this rearrangement only occurred upon binding of the appropriate codon (Swartz & Gassen, 1977). Apparently, in free tRNA, there was a complete base overlap within the anticodon stack, comparable to an oligo A stack. When double strand formation occurs between the codon and the anticodon, the characteristic anticodon helix may become extended. It was thought this rearrangement of the tertiary structure was caused by the change in the anticodon conformation (Swartz & Gassen, 1977). Semi-empirical energy calculations subsequently determined that the stacked single strand was more compact than the duplex (Broyde & Hingerty, 1978; Thiyagarajan & Ponnuswamy, 1978, 1979). In fact it was determined that energy would be required to destack the single strand sufficiently to allow duplex formation. These results support the suggestion that the anticodon configuration changed when it duplexed with the respective anticodon (Swartz & Gassen, 1977).

An investigation on the influence of codon context on the genetic code, found the efficiency of a suppressor tRNA, in translating the stop codon UAG, was increased ten-fold if the 3'-neighbour cytidine was

replaced with adenosine (Bossi & Roth, 1980). The authors felt that an increase in the stacking energy of the codon-anticodon interaction, due to the adjacent adenosine, accounted for the increase in suppressor efficiency.

These studies and proposals strongly suggest that a determination of the relative contributions, to conformation and stability, from base stacking is required. It is also important to learn what affects base stacking.

Semi-empirical calculations are possible to determine the contributions of base stacking. In the absence of hydrogen bonding and base stacking, the calculations predict extended, randomly coiled conformations (Olson, 1978), in agreement with NMR studies (Evans & Sarma, 1976). Calculation of the minimum energy conformers of stacked helices determined the bases were approximately parallel and separated by a mean perpendicular distance of 3.5 ± 0.2 Å (Olson, 1978) and corresponds to their van der Waal's thickness, pointing to the importance of non-specific van der Waal's forces in stacking interactions.

Calculations involving changes in the torsion angle χ, indicated that changing the base orientation from syn to anti, dramatically increased the degree of base stacking and conformational stability (Olson, 1978). The minimum energy conformers for the nucleotide sequences occurred when there was a high anti conformation of the base (Thiyagarajan & Ponnuswamy, 1978; Olson, 1978, 1980).

Minimum energy calculations for d(AAA) indicated that the 5'-phosphate group of the nucleoside residue had a definite effect on the base orientation, keeping it in the normal high anti region and, therefore,

affecting the base stacking property (Thiyagarajan and Ponnuswamy, 1979).

CD studies with a number of triribonucleotides (Brahms et al. 1967; Gray et al. 1972) as well as homo- and copolymers (Gray et al. 1972; Lomant & Fresco, 1972) showed that not only were stacking effects sequence dependent, but also that the four major bases have different stacking abilities (A>G>C>U). The results also indicated that uridines tended to promote destacking of the sequence (Brahms et al. 1967; Gray et al. 1972). Energy calculations of a variety of DNA (Hingerty & Broyde, 1978) and RNA (Broyde & Hingerty, 1978) dimers found the stacking energies were sequence dependent.

In an attempt to extract the thermodynamic parameters of helix formation (Borer et al. 1974) it was shown that helix stability had a striking dependence on sequence and that the largest contribution to helix stability probably came from the vertical stacking of the bases. This thorough UV investigation, of 19 oligoribonucleotides, showed that $\overset{\rightarrow}{\underset{\leftarrow}{G\cdot C}}$ duplexes were more stable than $\overset{\rightarrow}{\underset{\leftarrow}{C\cdot G}}$ which were more stable than $\overset{\rightarrow}{\underset{\leftarrow}{G\cdot C}}$. Also duplexes of alternating A·U sequences $\overset{\rightarrow}{\underset{\leftarrow}{A\cdot U}}$ were more stable than duplexes formed by a strand of adenosines base paired to a strand of uridines $\overset{\rightarrow}{\underset{\leftarrow}{A\cdot A}}$. In addition, duplexes formed by sequences with blocks of G·C base pairs and blocks of A·U base pairs were more stable than duplexes of alternating A·U and G·C base pairs. These results were attributed to stacking patterns of the sequences (Borer et al. 1974).

A study with the sequences $r(A_2GCU_2)$ and $r(A_2CGU_2)$ found that A·U-G·C stacking interactions were more favourable to duplex formation than A·U-C·G stacking interactions (Ravetch et al. 1974)

Investigations of polynucleotides that form stacked single stranded helices, have estimated the contributions of intrastrand stacking to helix stability. Poly rA (Stannard & Felsenfeld, 1975) and poly rC (Porschke, 1976; Frier et al. 1981) undergo a transition, at neutral pH, from a random coil conformation, at high temperature, to a helical, stacked conformation, at low temperature. The enthalpy changes, for poly rA stacking, range from -3 to -13 kcal/mole stack (for review see Turner et al. 1981) and when compared to the value of -7 kcal/mole base pair associated with the fully stacked double helix, poly rA:poly rU (Swurkuusk et al. 1977), can either suggest that base-base stacking contributions to duplex stability could be minor or dominant.

Recent reports have proposed that single strand stacking accounts for more than half the enthalpy associated with duplex formation (Frier et al. 1981; Turner et al. 1981). This agrees with minimum energy calculations that indicated stacking in the single strand was greater than in the duplex where stacking is constrained by base pair formation (Hingerty & Broyde, 1978; Broyde & Hingerty, 1978,1979) but that stacking is still the dominant factor in duplex stability (Desantis et al. 1981b).

1.6.1. Direction of Base Stacking

Studies with tRNA anticodon loops have also indicated that the direction of base stacking was also important. Carl Woese (1970) proposed that the molecular mechanics of translation depended on conformational changes in the tRNA. He determined the direction of stacking, in the anticodon, was energetically favoured in the 3' to 5' direction when the tRNA was charged and paired to the codon. When the tRNA was not charged, the favoured direction for the stacking interactions was 5' to 3' (Woese, 1970).

Urbanke & Mass (1978) observed a conformational transition in the anticodon loop with temperature. The alteration was attributed to a change in the stacking pattern, of the anticodon loop, from 5'-stacked at high temperatures, to 3'-stacked at low temperatures.

Stacking interactions have also been implicated in the elongation cycle of protein translation. Lake (1977) proposed translocation of the correct aminoacyl tRNA to the A-site from the recognition site (R-site), on the small ribosome, involved a change in the stacking interactions of the anticodon loop. The details include recognition at the R-site by an aminoacyl tRNA, with its anticodon in the 5'-stacked conformation. Movement from the R-site to the A-site required a switch, in the stacking pattern of the anticodon, from the 5'-stacked conformation to the 3'-stacked conformation (Lake, 1977).

Preliminary results from Turner's group (1981) indicated that sequences with a dangling 3'- or 5'- adenosine increased duplex stabilities due to stacking interactions but that the 3'-dangling adenosine contributed more than the 5'-dangling to stability. This was an indication that stacking contributions from one end of the duplex had a greater effect than from the other end.

Further investigations into the direction of base stacking and the other factors affecting base stacking are required for complete understanding of this phenomenon and its contributions to double helical stability and conformation.

2. SYNTHESIS OF OLIGORIBONUCLEOTIDES

Currently a number of methods, both chemical and enzymatic, are available for the synthesis of oligonucleotide sequences. Each method has certain drawbacks and, therefore, some thought must be given as to which method is of the most value to a particular project. The strategies developed for oligonucleotide synthesis include:

i) Phosphodiester synthesis

ii) Phosphotriester synthesis

iii) Enzymic synthesis

iv) Combined chemical/enzymatic synthesis

The chemical approaches to synthesis have proved to be extremely successful in the preparation of both deoxyribose and ribose sequences. Chemical synthesis also provides large scale amounts of any desired sequence. However, this technique is limited by the length of the sequence that can be efficiently produced. Although sequences of up to 20 bases have been made, this method is far more reliable for the preparation of sequences of 10 bases or less.

While the enzymatic methods are far more successful in the preparation of longer sequences, the quantities produced are much less. The amounts produced are sufficient for biological studies but are largely insufficient for the structural studies reported in this thesis. For this reason all the sequences investigated were chemically prepared.

The two approaches to the chemical synthesis of oligonucleotides differ in the nature of the phosphate bridges of the blocked oligomers.

In the phosphodiester method the phosphate bridges are monoanionic, whereas in the phosphotriester method the phosphate ester is neutral. Although the phosphodiester method has been successfully employed in the preparation of oligonucleotides, the monoanionic nature of the phosphate ester can result in possible side reactions, during chain extension, leading to branching or cleavage of the chain. Also the phosphodiester linkage is sensitive to acid and base so that care must be taken when working up reactions.

The phosphotriester synthesis protects the phosphodiester linkage and thus solves these problems. The drawback of the phosphotriester approach is one of time. The phosphorylation and coupling reactions are more lengthy and thus slow down the sequence building. However, the increased yields and purity of the compounds more than compensate for the loss of time. The phosphotriester chemical synthetic procedure developed in Neilson's laboratory has proved especially successful in the production of relatively large quantities of pure, biologically active, short oligoribonucleotides of any desired sequence. This is the method of choice for the research reported.

2.1. Chemical Synthesis of Oligoribonucleotides by the Phosphotriester Method

The phosphotriester approach used was developed before initiation of this research and details of the blocking groups and the procedures employed have been fully reported elsewhere (Neilson, 1969; Neilson & Werstiuk, 1971a, 1971b; Werstiuk & Neilson, 1972; Neilson et al. 1973; Werstiuk & Neilson, 1973; Neilson & Werstiuk, 1974; Neilson et al. 1975; Werstiuk & Neilson, 1976; England & Neilson, 1976;

England, 1976; England & Neilson, 1977; Gregoire & Neilson, 1978). However, a brief description is provided to familiarize the reader with the strategy used to synthesize the sequences of interest.

In order to relate the synthetically prepared sequences to naturally occurring RNA molecules it is essential that proper 3'-5' internucleotide phosphodiester linkages are formed without altering the glycosidic bond or causing loss of natural conformation or biological activity. To ensure the formation of the correct phosphodiester linkage other reactive centres on the nucleosides must be chemically blocked. Then, after the internucleotide linkages are formed, these blocking groups must be removed without disrupting the nucleosides or the phosphodiester bonds.

2.1.1. Blocking of Functional Groups

Functional groups that require blocking at different times during synthesis include the primary 5'-hydroxyl, the secondary 3'- and 2'-hydroxyl, the primary amino groups and the phosphodiester linkage. The most important prerequisites for a successful blocking group are the easy introduction onto the nucleoside and selective removal under conditions that leave other blocking groups intact. It is also essential that the intermediates of blocked nucleosides can be readily identified, in a routine manner, to ensure the blocking group is attached at the correct position. NMR analysis provides a simple procedure for this identification, (Reese, 1978).

2.1.1.1. Protection of the Terminal 5'-Hydroxyl

Sequences are usually built in the 5'→3' direction with a phosphate on the 3'-hydroxyl reacting with the free 5'-hydroxyl on the

incoming nucleoside. Short sequences (2-5 bases in length) are normally synthesized in a stepwise fashion and thus require that only the 5'-hydroxyl of starting nucleoside be blocked. However, with longer sequences (6 and up) a block synthesis is often employed to reduce the number of steps and increase yields. For example, in block synthesis, two trimers are joined together to yield a hexanucleotide. The incoming block must have a free 5'-hydroxyl but all other reactive groups must be protected. Therefore the blocking group chosen to protect the 5'-hydroxyl plays a pivotal role. It is necessary to be able to remove the 5' blocking group selectively, leaving all of the other protecting groups intact. The triphenylmethoxyacetyl (trac) group, developed by Neilson and co-workers (Werstiuk & Neilson, 1972), is ideal since it can be selectively introduced to the 5'-hydroxyl protected nucleoside derivatives, and is removed under mild base conditions (1% methanolic ammonia) which do not deprotect the N-benzoyl groups. An additional feature of this blocking group is the readily identifiable yellow spots it produces on t.l.c. plates when sprayed with ceric sulphate. This makes monitoring of phosphorylation and coupling reactions easier.

2.1.1.2. Blocking the 2'-Hydroxyl

Proper blocking of the 2'-hydroxyl is important in order to prevent 2'-5' phosphodiester linkages from occurring. The use of base labile acyl groups is unfavourable because of rapid migration between the 3' and 2'-hydroxyls in a pyridine solution with only traces of water present, as obtained during work-up of reaction mixtures (Reese & Trentham, 1965; Griffin et al. 1966). In addition, the equilibrium favours the 3'-isomer over the 2'-isomer. Also, studies on diribo-

nucleoside monophosphates demonstrated that 0.1 M HCl at 80°C causes up to 50%, acid catalyzed, migration of the 3'-phosphodiester linkage to the 2'-hydroxyl as well as chain cleavage (Brown et al. 1956). Both types of isomerization must be avoided because it is not a simple procedure to separate the desired product from those containing 2'-5' linkages.

These restrictions have made the tetrahydropyranyl group the most appropriate 2'-hydroxyl blocking group (Smrt & Sorm, 1962; Smith et al. 1962). This group is easily removed in three to four hours upon treatment with 0.01 M HCl at room temperature, conditions which do not lead to phosphodiester isomerization. The thp group can be specifically attached to the 2'-hydroxyl by blocking of other functional groups and specific removal leaving the 2'-hydroxyl as the only reactive site (Neilson et al. 1973; Gregoire & Neilson, 1978). Although the point of attachment on the thp ring is asymmetric, resulting in a diastereomeric mixture of the modified nucleoside, the two diastereomers are easily resolved by silica gel chromatography and can be crystallized. In addition, both diastereomers can be used in coupling reactions to yield the proper linkage.

The bulky nature of the thp group is sufficient to prevent 3'-3' phosphodiester formation without the need for separate 3'-hydroxyl blocking on the incoming nucleoside (Neilson, 1969).

2.1.1.3. Blocking of the Primary Amino Groups

Adenosine, guanosine and cytodine all have primary amino groups on the base that must be protected to prevent the formation of phosphoramidates via nucleophilic attack on the phosphorylating agent during

chain extension (Schaller et al.1963). These primary amino groups are protected by base labile benzoyl groups (Lohrmann et al. 1966). This benzoyl group is stable to neutral, acidic and mildly basic pH conditions used to selectively deprotect other functions. The group is removed by ammonolysis (7% methanolic ammonia).

2.1.1.4. The Phosphotriester Blocking Group

It is this blocking group that distinguishes the phosphotriester synthesis from the phosphodiester synthesis. A variety of blocking groups have been investigated for this most important position (Michelson & Todd, 1955; Letsinger & Mahadevan, 1965 & 1966; Letsinger & Ogilvie, 1967 & 1969; Eckstein & Rizk, 1967a & b) as well as a number of deblocking procedures (Van Boom et al. 1974; Ogilvie et al. 1976; Adanuak et al. 1977; Reese, 1978). The versatile phosphotriester synthesis of oligoribonucleotides developed in Neilson's laboratory uses 2,2,2-trichloroethyl as the phosphate protecting group (Eckstein & Rizk, 1967a & b) with good success (Neilson, 1969; Werstiuk & Neilson, 1973, 1976; Neilson & Werstiuk, 1974). Carefully prepared copper-zinc couple completely deprotects the trichloroethyl groups, in anhydrous dimethyl formamide, by O-alkyl bond cleavage, from sequences as long as a nonaribonucleotide, in reasonable yields (Neilson & Werstiuk, 1974; Werstiuk & Neilson, 1976).

Therefore the basic features of the phosphotriester synthesis used to prepare the sequences in this report are:

i) terminal 5'-hydroxyl is blocked by a triphenylmethoxyacetyl group which can be specifically removed by mild base for block synthesis.

ii) 2,2,2-trichloroethyl phosphate blocking group, removed by carefully prepared copper-zinc couple via O-alkyl bond cleavage with no chain cleavage.

iii) the tetrahydropryanyl group blocks the 2'-hydroxyl and is removed by treatment at pH 2.0 at room temperature with no appreciable phosphate migration. The bulky nature of the thp group prevents 3'-3' internucleotide bonds without 3'-hydroxyl protection (Neilson, 1969).

iv) primary amino groups are protected by benzoyl, removed by ammonolysis.

v) phosphorylations and couplings are affected using the pyridinium salt of 2,2,2-trichloroethyl phosphate and mesitylene-1,2,4-triazolide as an activating agent.

After complete synthesis of a desired sequence is complete, a three step deblocking procedure is used to obtain the free oligoribonucleotides (England & Neilson, 1976).

i) treatment for one hour at 60°C in DMF with Cu/Zn couple to remove the trichloroethyl groups.

ii) treatment for two days with 50% methanolic ammonia to remove the trac and benzoyl groups, followed by purification by Whatman 1 paper chromatography.

iii) treatment for two days at pH 2.0 and room temperature to remove the thp groups followed by purification by paper chromatography on Whatman 40.

The versatility of this procedure is attested to by the number and variety of sequences that have been prepared for this study as well

as others (England & Neilson, 1976; Werstiuk & Neilson, 1976; Ganoza et al. 1978; Romaniuk et al. 1979; Neilson et al. 1980b). The sequence integrity is checked by NMR and has been confirmed enzymatically (Ganoza et al. 1978). Another important feature of this synthesis is the confirmed biological activity of the free oligoribonucleotides (Ganoza et al. 1978; Neilson et al. 1980b). In addition, this synthetic method is the only one that is capable of providing satisfactory coupling to guanosine residues in sequences longer than a trimer for oligoribonucleotides (Ganoza et al. 1978).

3. EXPERIMENTAL

3. Experimental Procedures

3.1. Materials and Reagents

All reagents and solvents were commercial reagent grade and used without further purification except for the following: mesitylenesulphonyl chloride (Aldrich) was recrystallized from hot petroleum ether (30° - 60°C); pyridine, dimethyl formamide and p-dioxane (Baker Analyzed reagents) were dried over Fischer 4Å molecular sieves prior to use.

It is essential that all moisture be excluded from condensation reactions. After evaporation in vacuo of the pyridine solution, normal pressure is restored with dry nitrogen.

The emulsions frequently obtained during methylene chloride extractions of aqueous pyridine solutions were broken by addition of saturated sodium chloride solution (1-2 ml).

Protection of nucleosides and coupling reaction mixtures were followed by thin layer chromatography using prescored Silica Gel G plates (Analtech) run in 10% methanol in methylene chloride. Detection was accomplished by spraying with 1% $H_4Ce(SO_4)_4$ in 10% sulphuric acid and heating to 150°-200°C. Compounds containing trityl groups appeared as yellow spots and those without appeared as brown spots.

Silica gel (40-140 mesh, Baker Analyzed) used in column chromatography was dried at 80°C over NaOH for two days prior to use. Columns were prepared in methylene chloride and eluted using a step gradient of methanol in methylene chloride.

The deprotected sequences were purified by descending paper chromatography on Whatman #1 and #40 papers. The solvent system was ethanol/NH_4OAc (1 M, pH 7.3), 1:1 (v/v). A short wave UV lamp was used to detect deblocked oligonucleotides.

3.1.1. Protected Nucleosides and Coupling Reagents

The nucleoside derivatives were synthesized and characterized by published procedures: HOU_tOH (Griffin et al. 1968); $HOC_t^{bz}OH$ (Neilson & Werstiuk, 1971a); Trac U_tOH (Werstiuk & Neilson, 1972); Trac $A_t^{bz}OH$ and Trac $C_t^{bz}OH$ (Werstiuk & Neilson, 1973); $HOG_t^{bz}OH$ and $HOA_t^{bz}OH$ (Gregoire & Neilson, 1978).

The coupling reagents also were synthesized and characterized by established procedures: 2,2,2'-trichloroethyl phosphate (England & Neilson, 1977); MST (Katagiri et al. 1974); TPS-NI (Van Boom et al. 1977).

3.2. Preparation of Oligoribonucleotides

Tables 1 and 2 provide the complete details for the synthesis of the sequences used in this study. All were made using published techniques (England & Neilson, 1976; Werstiuk & Neilson, 1976). The stepwise synthesis of fully protected ApGpApCpU is detailed below to illustrate the general procedures.(Figure 6).

3.2.1. N^6-benzoyl-2'-O-tetrahydropyranyl-5'-O-triphenylmethoxyacetyladenylyl[3'(2,2,2-trichloroethyl)-5'] N^2-benzoyl-2'-O-tetrahydropyranylguanosine (Trac $A_t^{bz}pG_t^{bz}OH$).

A sample of 2,2,2-trichloroethylphosphate (England & Neilson, 1976) (1220 mg, 5.3 mmole, 2 equiv.) was evaporated in vacuo from anhydrous pyridine (3 x 50 ml) to convert it to the pyridinium salt. To activate the phosphate, MST (2660 mg, 10.6 mmole, 4 equiv.) was added

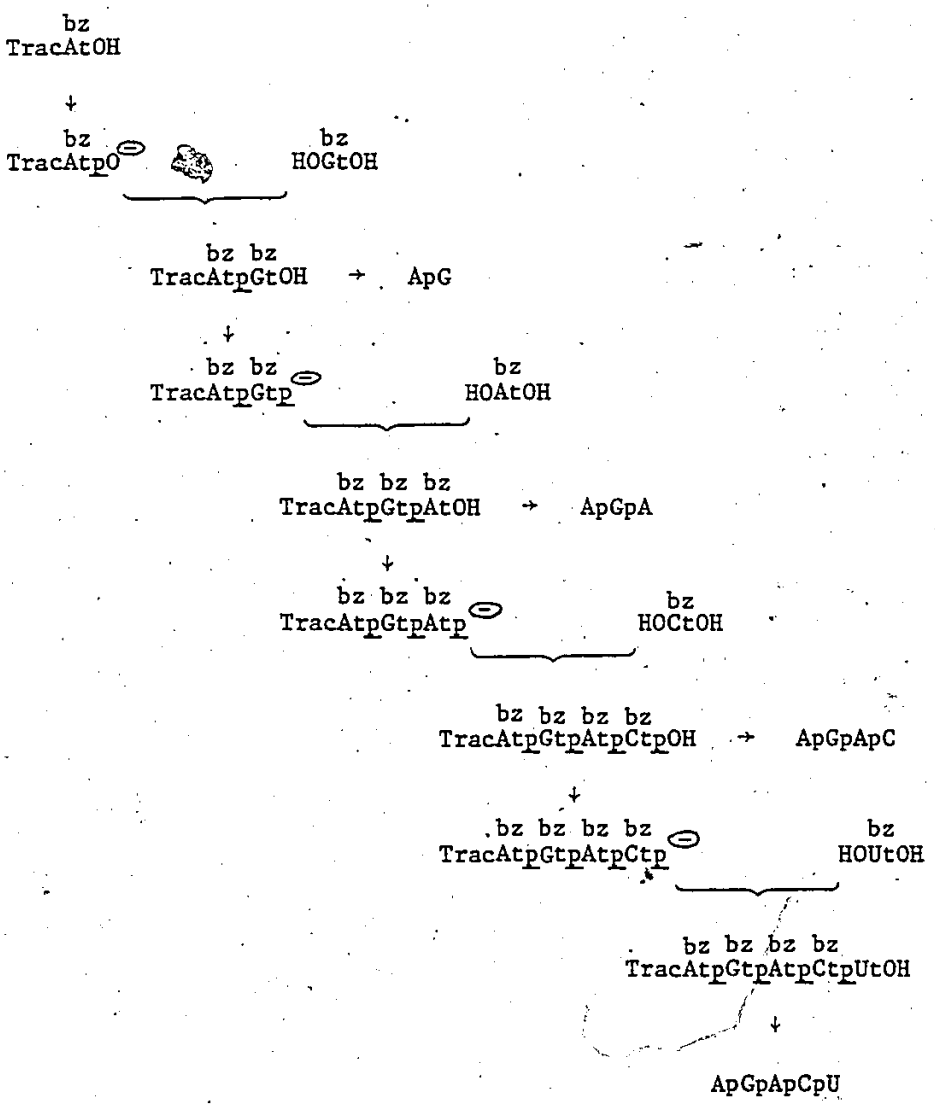


Figure 6. Schematic diagram for the stepwise chemical synthesis of the pentaribonucleotide ApGpApCpU by the phosphotriester method. In addition to the abbreviations recommended by IUPAC-IUB (1970) the following are also used: Trac, triphenylmethoxyacetyl; t, tetrahydropyranyl; p$O^{\ominus}$, 3'-O-(2,2,2-trichloroethyl) phosphate; p between two characters, 3',5'-(2,2,2-trichloroethyl) phosphotriester.

TABLE 1

SUMMARY OF THE STEPWISE PREPARATION OF PROTECTED OLIGORIBONUCLEOTIDES[a]

Reactants						Products		
	Quantity			Quantity			Quantity	
Compd	mg	mmole	Compd	mg	mmole	Compd	mg	% Yield
A	2000	2.65	G	1500	3.18	AG	2200	59
AG	500	0.35	A	191	0.42	AGA	370	51
AGA	250	0.12	C	43	0.17	AGAC	208	64
AGAC	160	0.06	U	22	0.07	AGACU	113	59
AG	500	0.35	C	181	0.42	AGC	470	65
AGC	350	0.17	C	109	0.25	AGCC	292	64
AGCC	220	0.08	U	33	0.10	AGCCU	160	61
AG	500	0.35	U	138	0.42	AGU	340	50
AGU	250	0.13	C	69	0.16	AGUC	213	65
AGUC	213	0.08	U	33	0.10	AGUCU	127	50
AG	700	0.49	G	278	0.59	AGG	480	47
AGG	200	0.07	C	62	0.14	AGGC	150	58
AGGC	110	0.04	U	16	0.05	AGGCU	61	46
AGC	200	0.10	U	40	0.12	AGCU[b]	125	48
AGCU	120	0.05	A	21	0.05	AGCUA[b]	63	43
A	1980	2.62	U	1040	3.17	AU	1400	42
AU	1400	1.10	G	620	1.32	AUG	1520	71
C	4860	6.65	U	2617	7.98	CU	3778	45
CU	300	0.24	G	135	0.29	CUG	293	59
CU	1780	1.42	A	759	1.67	CUA	1900	71
CU	1780	1.42	U	547	1.67	CUU	2030	81
CUU	1000	0.56	A	306	0.67	CUUA	867	63

[a] Column 1 contains the 5'-trityloxyacetyl reactants and A is Trac ^{bz}AtOH. Column 4 contains the incoming nucleoside all of which are protected at the 2'-OH group with a tetrahydropyranyl residue and all except U are protected at their NH_2 groups with a benzoyl residue. Column 7 contains the trityloxyacetyl products and for example, AG stands for Trac-^{bz}A-p-^{bz}GOH. Two equivalents of pyridinium mono 2,2,2-trichloroethyl phosphate activated by 4 equivalents of mesitylene sulphonyl 1,2,4-triazolide (MST) in anhydrous pyridine are used in each phosphorylation step. The coupling step to the incoming nucleoside derivative is driven by 1.2 equivalents of MST.

[b] Compounds prepared by Mr. René Grégoire.

TABLE 2

SUMMARY OF THE STEPWISE PREPARATION OF

PROTECTED OLIGORIBONUCLEOTIDES

Reactants						Products		
	Quantity			Quantity			Quantity	
Compd	mg	mmole	Compd	mg	mmole	Compd	mg	% Yield
U	550	0.88	G	430	0.88	UG	560	49
UG	500	0.40	C	175	0.41	UGC	480	63
UGC	350	0.19	A	95	0.21	UGCA	360	77
UGCA	300	0.12	A	60	0.13	UGCAA[b]	180	47
G	2500	3.24	C	1677	3.89	GC	1860	41
GC	1860	1.32	A	721	1.58	GCA	1862	69
GCA	400	0.20	A	106	0.23	GCAA	270	51
C	690	0.94	G	532	1.13	CG	445	34
CG	415	0.30	A	164	0.36	CGA	222	37
A	438	0.58	C	250	0.58	AC	608	65
AC	512	0.37	G	209	0.44	ACG	508	67
A	3170	4.20	A	2293	5.04	AA	4987	84
AA	4937	3.51	G	1984	4.21	AAG	3880	54
U	1000	1.59	C	823	1.91	UC	1300	65

[a] For synthetic details see footnote to Table 1

[b] Compounds prepared by Dr. P.J. Romaniuk

Synthesis of CAUG, CAUGA, CAUGU, CAXUG, CAX and CXG (where X = A, G, C or U) has been previously reported in the Ph.D. thesis of Paul Romaniuk (1979).

to the pyridine solution (25 ml) and the solution was warmed (40°C, 1 hr) under dry nitrogen. This solution of activated phosphate was added to a pyridine solution (ca. 20 ml) of Trac $A_t^{bz}OH$ (2000 mg, 2.65 mmole, 1 equiv.) and the total volume reduced in vacuo ca. 25 ml, sealed under N_2 and let stand at room temperature in the dark. Tlc indicated the phosphorylation was ca. 80% complete after one day (R_f 0.80 → 0.30). Additional MST (50 mg, 0.20 mmole) was added and the reaction was complete after two days. The reaction was quenched by addition of ice (ca. 5 gm). After 30 minutes, the reaction mixture was poured into ice water (ca. 100 ml) and extracted with methylene chloride (5 x 50 ml). The combined organic extracts were washed with water (1 x 50 ml) and evaporated in vacuo to a light yellow oil.

Trac $A_t^{bz}pO^-$ was evaporated from anhydrous pyridine (3 x 50 ml, final volume ca. 25 ml) and MST (800 mg, 3.18 mmole, 1.2 equiv.)added. The solution was warmed gently (35°C) for one hour. A pyridine solution (ca. 20 ml) of high R_f isomer of $HOG_t^{bz}OH$ (1500 mg, 3.18 mmole, 1.2 equiv) was then added, the final volume reduced in vacuo ca. 25 ml and the reaction mixture sealed under dry nitrogen. The reaction mixture was then stored at room temperature, in the dark, for two days. Tlc indicated the reaction was 95% complete (R_f 0.30 - 0.70). The reaction was quenched with ice, poured into ice water (ca. 100 ml) and extracted into methylene chloride (5 x 50 ml). The combined extracts were washed with water (1 x 50 ml) and evaporated in vacuo to dryness. The last traces of pyridine were removed by co-distillation with toluene (3 x 15 ml) to give a yellow foam. The foam was purified by silica gel (25 gm) chromatography. Pure Trac $A_t^{bz}pG_t^{bz}OH$ eluted with 3% methanol in methylene chloride (2200 mg, 1.32 mmole, 59%).

3.2.2. N^6-benzoyl-2'-O-tetrahydropyranyl-5'-O-triphenylmethoxyacetyl-adenylyl[3'(2,2,2-trichloroethyl)-5'] N^2-benzoyl-2'-O-tetrahydropyranylguanylyl [3'(2,2,2-trichloroethyl)-5'] N^6-benzoyl-2'-O-tetrahydropyranyladenosine (Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}OH$).

2,2,2-trichloroethylphosphate (161 mg, 0.70 mmole, 2 equiv.) was converted to its pyridinium salt by repeated evaporation *in vacuo* from anhydrous pyridine (3 x 20 ml) and activated by MST (351 mg, 1.40 mmole, 4 equiv.). After warming for 1 hr., the activated solution was added to a pyridine solution of Trac $A_t^{bz}pG_t^{bz}OH$ (500 mg, 0.35 mmole, 1.2 equiv.) and the volume reduced *in vacuo ca.* 15 ml. The reaction stood at room temperature under dry nitrogen. After 2 days, the reaction was complete (R_f 0.70 - 0.35). Ice (3 gm) was added, the reaction poured into ice water (*ca.* 50 ml) and repeatedly extracted with methylene chloride (4 x 30 ml). The combined extracts were washed with water (1 x 50 ml) and evaporated *in vacuo* to a yellow oil.

The Trac $A_t^{bz}pG_t^{bz}pO^-$ was then evaporated from anhydrous pyridine (3 x 20 ml, final volume *ca.* 10 ml) and activated with MST (105 mg, 0.42 mmole, 1.2 equiv.). After 1 hr., a pyridine solution of the high R_f isomer of $HOA_t^{bz}OH$ (455 mg, 0.43 mmole, 1.2 equiv.) was added, the final volume reduced *in vacuo ca.* 15 ml and the reaction sealed under dry nitrogen and stored in the dark. After two days the reaction was 70% complete (R_f 0.03 - 0.60) and some MST (30 mg, 0.12 mmole) was added to push the reaction to completion. Next day the reaction was complete and quenched with ice (ca. 3 gm), poured into ice water after 30 minutes and repeatedly extracted with methylene chloride (4 x 30 ml). The combined extracts were washed with water (1 x 50 ml) and evaporated to a

yellow foam. Purification on a 15 gm silica gel column yielded pure Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}OH$ (370 mg, 0.16 mmole, 51%) upon elution with 3% methanol-methylene chloride.

3.2.3. N^6-benzoyl-2'-O-tetrahydropyranyl-5'-O-triphenylmethoxyacetyl-adenylyl [3'-(2,2,2-trichloroethyl)-5'] N^2-benzoyl-2'-O-tetrahydropyranylguanylyl[3'-(2,2,2-trichloroethyl)-5'] N^6-benzoyl-2'-O-tetrahydropyranyladenylyl[3'-(2,2,2-trichloroethyl)-5']-N^4-benzoyl-2'-O-tetrahydropyranylcytidine (Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}pC_t^{bz}OH$).

A sample of 2,2,2-trichloroethylphosphate (56 mg, 0.24 mmole, 2 equiv.) was converted to its pyridinium salt by repeated evaporation from anhydrous pyridine (3 x 10 ml) and activated by MST (122 mg, 0.48 mmole, 4 equiv.). After 1 hr, the activated solution was added to a pyridine solution of Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}OH$ (250 mg, 0.12 mmole, 1 equiv.) and the volume reduced in vacuo ca. 10 ml. The reaction stood at room temperature under dry nitrogen. After 3 days the reaction was complete (R_f 0.60 → 0.30) and quenched by addition of ice (ca. 2 gm). The reaction was poured into ice water (ca. 25 ml) and extracted with methylene chloride (5 x 25 ml). The combined extracts were washed with water (1 x 25 ml) and evaporated in vacuo to a yellow oil.

Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}pO^-$ was evaporated from anhydrous pyridine (3 x 15 ml, final volume ca. 10 ml) and activated with MST (36 mg, 0.14 mmole, 1.2 equiv.). After 1 hr, a pyridine solution of the high R_f isomer of $HOC_t^{bz}OH$ (63 mg, 0.14 mmole, 1.2 equiv.) was added, the final volume reduced in vacuo ca. 10 ml and the reaction sealed under dry N_2. After 3 days the reaction was complete by tlc (R_f 0.30 → 0.50) and the reaction quenched by addition of ice (ca. 2 gm). The reaction mixture was then poured into ice water (ca. 20 ml) and extracted with methylene

chloride (5 x 20 ml). The combined extracts were washed with water (1 x 25 ml) and evaporated *in vacuo* to a yellow foam. Purification on a 10 gm silica gel column yielded Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}pC_t^{bz}OH$ (208 mg, 0.077 mmole, 64%) upon elution with 4% methanol-methylene chloride.

3.2.4. *N^6-benzoyl-2'-O-tetrahydropyranyl-5'-O-triphenylmethoxyacetyl-adenylyl[3'-(2,2,2-trichloroethyl)-5'-] N^2-benzoyl-2'-tetrahydropyranylguanylyl[3'-(2,2,2-trichloroethyl)-5') N^6-benzoyl-2'-O-tetrahydropyranyladenylyl[3'-(2,2,2-trichloroethyl)-5'] N^4-benzoyl-2'-O-tetrahydropyranylcytidylyl[3'-(2,2,2-trichloroethyl)-5']-2'-O-tetrahydropyranyluridine* (Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}pC_t^{bz}pU_t^{bz}OH$)

A sample of 2,2,2-trichloroethylphosphate (26 mg, 0.11 mmole, 2 equiv.) was converted to its pyridinium salt by repeated evaporation *in vacuo* from anhydrous pyridine (3 x 10 ml) and activated by MST (56 mg, 0.22 mmole, 4 equiv.). After 1 hr, the activated solution was added to a pyridine solution of Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}pC_t^{bz}OH$ (160 mg, 0.056 mmole, 1 equiv.) and the volume reduced *in vacuo ca.* 10 ml. The reaction stood at room temperature under dry nitrogen. After 3 days, the reaction was *ca.* 80% complete (R_f 0.50 → 0.30) and MST (10 mg) was added. Two days later the reaction was complete and ice, *ca.* 2 gm, was added, the reaction poured into ice water (*ca.* 20 ml) and extracted with methylene chloride (5 x 25 ml). The combined extracts were washed with water (1 x 30 ml) and evaporated *in vacuo* to dryness.

Trac $A_t^{bz}pG_t^{bz}pA_t^{bz}pC_t^{bz}pO^-$ was evaporated from anhydrous pyridine (3 x 15 ml, final volume *ca.* 10 ml) and activated with MST (20 mg, 0.078 mmole, 1.4 equiv.). After 1 hr, a pyridine solution of the high R_f isomer of HOU_tOH (22 mg, 0.067 mmole, 1.2 equiv.) was added, the

final volume reduced ca. 10 ml and the reaction sealed under nitrogen and stored in the dark. After two days tlc indicated the reaction was 80% complete (R_f 0.30 → 0.48) and MST (10 mg) was added. After two more days tlc indicated the reaction was complete. Ice (ca. 2 gm) was added, 30 mins. later the reaction mixture was poured into ice water (ca. 20 ml) and extracted into methylene chloride (5 x 25 ml). The combined extracts were washed with water (1 x 30 ml) and evaporated in vacuo to a yellow foam. Purification on a 10 gm silica gel column yielded pure Trac A_t^{bz}-$pG_t^{bz}pA_t^{bz}pC_t^{bz}pU_t^{bz}OH$ (113 mg, 0.035 mmole, 60%) upon elution with 5% methanol-methylene chloride.

3.3. Block Synthesis

While the stepwise procedure had been very successful in the production of short nucleotide sequences, the yields after each addition were often less than 70%. As a result, final yields for sequences of 6 or more bases were quite small. Occasionally then, it was necessary to use the procedure of block synthesis to prepare longer sequences (Werstiuk & Neilson, 1972). Figure 7 gives a scheme for the block synthesis of $\vec{AAGCUA}$.

To synthesize $\vec{AAGCUA}$ by block synthesis first required the stepwise synthesis of two trimers, Trac $A_t^{bz}pA_t^{bz}pG_t^{bz}OH$ and Trac $C_t^{bz}pU_t^{bz}pA_t^{bz}OH$. Trac ApApG-OH was phosphorylated in the usual manner, as described above. However, in order to afford a linkage with the second trimer a free 5'-hydroxyl is required. It is, therefore, necessary to first remove the 5'-O-triphenylmethoxyacetyl (Trac) blocking group from TracCpUpA-OH, selectively. This is accomplished by dissolving the sequence to be deprotected in a prepared solution of 1% ammonia in methanol (1 ml of solution per 10 mg of sequence).

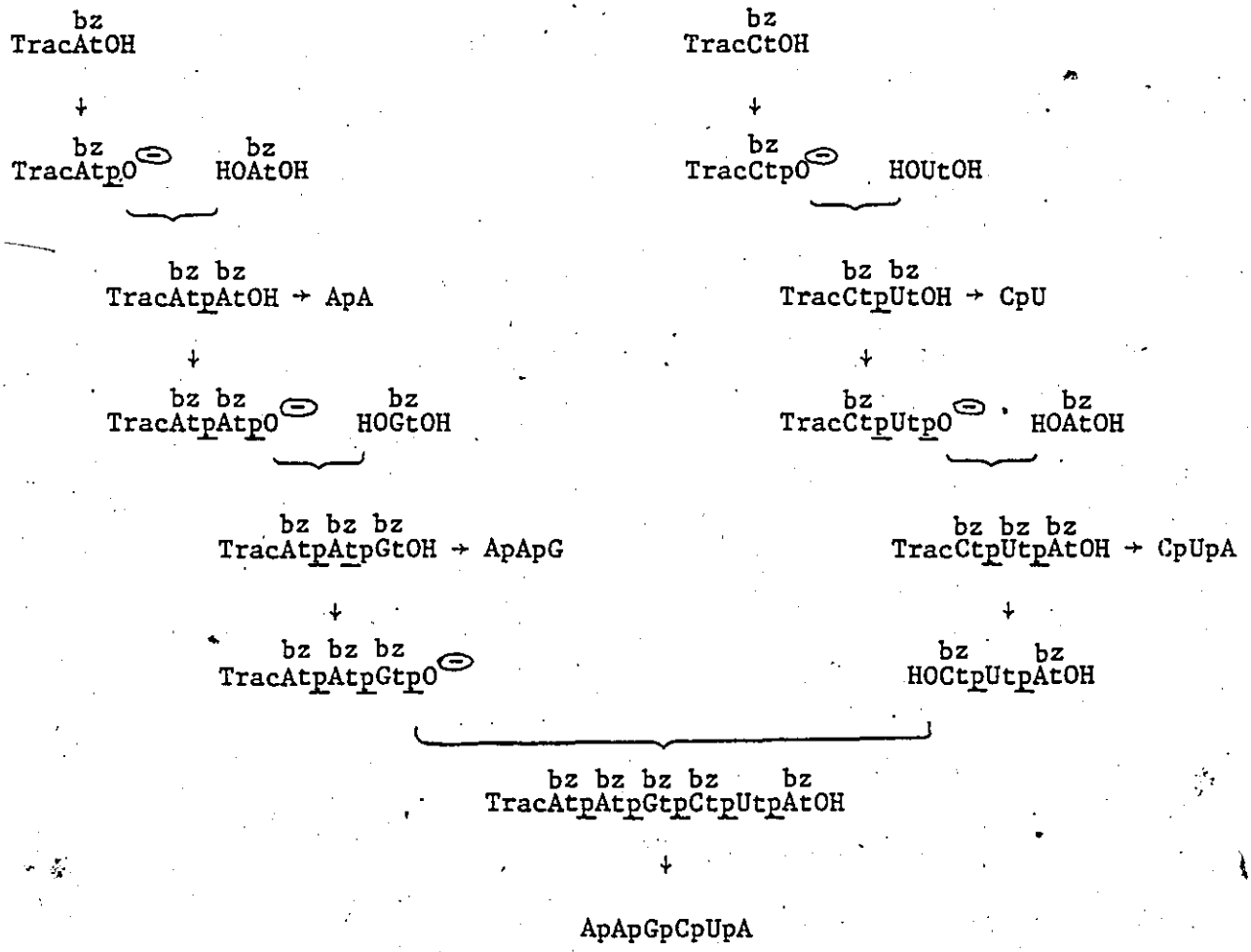


Figure 7. Schematic representation for the chemical block synthesis of hexaribonucleotide ApApGpCpUpA. Abbreviations used as in Figure 6.

TABLE 3

SUMMARY OF THE BLOCK SYNTHESIS

OF OLIGORIBONUCLEOTIDES[a]

Reactants						Products		
	Quantity			Quantity			Quantity	
Compd	mg	mmole	Compd	mg	mmole	Compd	mg	% Yield
AAG	649	0.31	CU	359	0.38	AAGCU	718	71
AAG	322	0.16	CUA	300	0.19	AAGCUA	380	63
AAG	284	0.14	CUU	243	0.17	AAGCUU	454	65
AAG	311	0.15	CUUA	384	0.18	AAGCUUA	455	57
AC	150	0.11	AUG	186	0.11	ACAUG[b]	155	44
UC	183	0.15	AUG	262	0.16	UCAUG[b]	180	40

[a] Footnotes as in Table 1

[b] Compounds prepared by Ms. E.S. Quittkat

The reaction was followed by tlc. As the reaction proceeds, the yellow spot associated with Trac groups moves to the solvent front and a brown spot with a lower R_f appears (R_f 0.70 → 0.40). The brown spot indicates removal of the Trac group. Care must be taken to avoid leaving the reaction too long or removal of N-benzoyl blocking groups will also occur (R_f 0.30). The reaction is quenched by evaporation in vacuo and HO-C_t^{bz}pU_t^{bz}pA_t^{bz}OH purified by silica gel chromatography. This trimer now has a free 5'-hydroxyl and can be reacted with phosphorylated TracApApG-OH in the same manner as with the stepwise synthesis coupling reactions, to yield fully protected TracApApGpCpUpA-OH. In this way larger yields of longer sequences can be obtained. Table 3 provides the preparative data for the sequences made using the block synthesis procedure.

3.4. Deprotection of Short Oligoribonucleotides

All protected oligoribonucleotides were deblocked in the same manner (England & Neilson, 1976). The protected oligonucleotide (10 or 20 mg) was dissolved in 0.5 ml of dry dimethylformamide and Cu/Zn couple (ca. 10-20 mg) was added. This reaction was then stirred for 1-2 hours at 50°C, and monitored by tlc. The reaction was judged to be complete when the spot on the tlc did not move off the origin (R_f 0.0). Methanolic ammonia [methanol:cNH_3, 1:1 (v/v)] was added, the reaction vessel tightly sealed, and left to stir for two days at room temperature. The Cu/Zn couple was filtered and washed with 1 N ammonia. Chelex-100 (NH_4^+ form) was added to the combined filtrate and washes and stirred for 1 hr. The chelex was then filtered off and washed with 1 N ammonia. The combined filtrate and washings were then

evaporated *in vacuo* to dryness and the partially deprotected oligomer was purified by descending paper chromatography on Whatman #1 paper in ethanol: 1 M NH_4OAc (pH 7.3) (1:1, v/v). The desired band (R_f 0.8 - 0.9) was cut out, desalted by soaking in ethanol (1 hr) and then anhydrous diethyl ether (15 min). The bands were eluted with glass distilled water. The pH of the eluate was adjusted to 2.0 with 2 N HCl and left to stand for two days at room temperature. The solution was neutralized with ammonia, evaporated to dryness and purified using descending paper chromatography on Whatman #40 paper with the same solvent system as before. The desired band was cut out and eluted with glass distilled water.

All glassware used in this procedure was acid washed to prevent possible ribonuclease digestion of the free oligoribonucleotides. Data for the deprotection of the oligomers used in this study are presented in Table 4. The sequences were characterized by their pmr spectra at 70°C.

3.5. Methodology for ^{1}H-NMR Characterization of Sequence

The various ^{1}H-NMR spectra required for this study were obtained from Bruker WP-80, WH-90, WM-250 and WH-400 MHz spectrometers, all operating in the Fourier Transform mode and equipped with quadrature detection. Probe temperatures were maintained to within ±1°C by Bruker variable temperature units and calibrated by thermocouple measurements. 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 *ca.* 7.2) and 1.0 M sodium chloride. Sample concentrations were 10^{-3} - 10^{-2}M. t-Butanol-OD was used as an internal reference (1.231 ppm)

TABLE 4

CHROMATOGRAPHIC DATA AND DEPROTECTION YIELDS

Compound[a]	R_f[b]	Yield in %[c]	Compound[a]	R_f[b]	Yield in %[c]
ApG	0.40	60	UpGpC	0.49	57
ApU	0.59	71	ApGpApC	0.19	79
ApC	0.52	65	ApGpGpC	0.21	77
ApA	0.46	66	ApGpCpC	0.20	62
GpC	0.50	23	ApGpUpC	0.20	49
CpU	0.67	70	ApGpCpU	0.31	33
CpG	0.40	23	CpUpUpA	0.32	51
UpG	0.47	39	UpGpCpA	0.24	37
UpC	0.62	27	GpCpApA	0.35	38
ApUpG	0.43	69	ApGpApCpU	0.15	62
ApCpG	0.42	54	ApGpGpCpU	0.17	32
ApApG	0.29	49	ApGpCpCpU	0.17	77
ApGpA	0.29	73	ApGpUpCpU	0.16	81
ApGpC	0.36	60	ApGpCpUpA	0.12	34
ApGpG	0.32	26	ApApGpCpU	0.12	41
ApGpU	0.34	85	UpGpCpApA	0.14	40
GpCpA	0.35	45	ApCpApUpG	0.13	45
CpGpA	0.31	46	UpCpApUpG	0.16	30
CpUpG	0.45	66	ApApGpCpUpU	0.10	40
CpUpU	0.53	81	ApApGpCpUpA	0.08	41
CpUpA	0.46	84	ApApGpCpUpUpA	0.07	45

[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 U.V. spectrophotometric data assuming a 90% hypochromicity factor.

and the chemical shifts are reported relative to sodium 2,2-dimethyl-2-silapentane-5-sulphanate (DSS). The field/frequency lock was provided by the deuterium signal of D_2O.

For the purpose of checking sequence integrity the samples were run at 70°C to ensure no secondary structure interactions occurred. In order to study duplex formation the spectra were taken at intervals over the temperature range 70°-10°C. Graphing of chemical shift versus temperature yields curves that can be interpreted. Upfield shifts are indicative of general strand stacking whereas downfield shifts usually mean looping out has occurred. Curves displaying sigmoidal behaviour are interpreted as being indicative of base pairing and thus duplex formation.

The base ratios of the oligoribonucleotide sequences are readily available from the distinct aromatic signals at low field. Sequence integrity is checked by the procedure of incremental analysis (Borer et al. 1975) which involves checking each intermediate in the stepwise synthesis of a sequence. Since the sequences are built in the 5' direction only, there is little chance of an incorrect sequence being made. NMR analysis also provides a method for checking the purity of the compound.

4. RESULTS AND DISCUSSION

4.1. Development of NMR Strategy

(Everett et al. 1980; Bell et al. 1981).

Proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy has become an extremely useful technique for conformational studies of nucleic acids in aqueous solutions. ^{1}H-NMR spectroscopy has been applied to mononucleotides (Blackburn et al. 1970; Imoto et al. 1976; Davies, 1978), dinucleoside monophosphates (Evans et al. 1975; Krugh et al. 1976; Lee et al. 1976; Ezra et al. 1977), higher oligomers (Borer et al. 1975; Patel, 1976; Romaniuk et al. 1978, 1979; Early et al. 1981) and complete natural nucleic acid molecules (Reid & Hurd, 1977; Kearns, 1977; Johnston & Redfield, 1978, 1981). Proton chemical shifts and homonuclear (^{1}H - ^{1}H) and heteronuclear (^{1}H - ^{31}P) coupling constants are the most frequently used parameters. However, nuclear Overhauser effects (NOE), relaxation times and lanthanide-induced shifts have all been used to aid in the conformational analysis of nucleic acids.

A chief advantage of ^{1}H-NMR spectroscopy over other solution techniques such as circular dichroism (CD) and ultra-violet (UV) spectroscopy is the ability to monitor changes in local micro-environments by recording molecular events at specific atomic sites in the molecule, rather than monitoring the molecule as a whole. More simply, several signals are usually obtained from each nucleotide residue in the molecule yielding spectra with a high information content. However, this very feature can also become a severe disadvantage with oligomers

containing more than three nucleotide units, since the complexity of the spectra makes them difficult to interpret. The usefulness of NMR spectroscopy is greatly diminished if unambiguous assignments cannot be made. In the past, studies of longer sequences have been complicated by this problem (Patel, 1976).

Since this thesis deals primarily with sequences at the tetramer level or longer it was fundamental to the thrust of this thesis that the unambiguous assignment of the spectra for any sequence could be achieved. To this end, a strategy for spectral assignment was developed.

Initially, two sets of triribonucleoside diphosphate sequences, $\vec{A}\vec{G}X$ and $\vec{C}\vec{A}X$ (where X = A, G, C or U), were studied to determine the nearest and next-nearest-neighbour effects in the proton NMR spectra of a nucleotide added to the dimer $\vec{A}G$ or $\vec{C}A$ (Everett et al. 1980). This study was followed by a work that determined the effects of a base inserted into the middle of a sequence on the chemical shifts of the sequence. This work studied the sequences $\vec{C}G$, $C\vec{X}G$, $C\vec{A}UG$, $CA\vec{X}UG$, $\vec{A}G$, $A\vec{X}G$, $\vec{A}GC$, $A\vec{G}XC$ and $AG\vec{X}CU$ (Bell et al. 1981). The position of a nucleotide within an oligoribonucleotide sequence was found to have effects on the chemical shifts of protons on residues at the nearest and next-nearest-neighbour positions. The effects could be rationalized in terms of relative shielding abilities of the bases.

4.1.1. Results

Table 5 lists the ^{1}H-NMR chemical shifts of all the nonexchangeable base protons and the ribose H-1' protons in both the $Ap\vec{G}pX$ and the $Cp\vec{A}pX$ series at 71°C ± 2°C. From the data in this table two sets of shielding parameters were obtained; the first set, listed in Table 6, shows the

shift changes induced in the protons of the dimer ApG on the addition of a 3'-terminal nucleotide pX; the second set, listed in Table 7, shows the corresponding data for the CpApX series of compounds. Table 9 lists the ^{1}H-NMR chemical shifts of all the nonexchangeable base protons and ribose H-1' protons in both the CXG and AXG series of compounds at 71°C ± 2°C. Tables 10 and 11 list the corresponding data for the CAXUG and AGXC series of compounds, respectively. Table 12 lists the assignments for the AGXCU series. From each of Tables 9-12 a set of shielding parameters was obtained corresponding to the effect of inserting a nucleotide X in the parent molecule. These shielding data are given in Tables 13-16.

4.1.2. Discussion

In the assignment of resonances several, general, well-known principles are used as guidelines in the assignment of a particular resonance to a certain class (Borer et al. 1975).

i) The chemical shifts of the protons in the oligonucleotides at high temperature closely resemble those of the mononucleotides.

ii) The H-5 and H-6 doublets from the pyrimidines have temperature-invariant coupling constants with $J_{5,6}$ values of 7.6 Hz for C and 8.1 Hz for U. The H-5 resonances and and the H-6 resonances are separated by about 1.8 ppm with the H-5 resonances located upfield in the region of the H-1' resonances.

iii) The H-1' resonances have temperature-sensitive coupling constants, $J_{1',2'}$, which are less than 6 Hz and are further reduced at low temperatures.

iv) The H-2 and H-8 singlets can be unambiguously distinguished on the basis of their spin-lattice relaxation time, T_1. The T_1 (5-7 sec.) of H-2 is longer than the T_1 (1-3 sec.) of H-8 (Ts'o et al. 1973). Therefore, if the 90° pulse rate is increased in the FT mode for spectrum aquisition, the signal intensity of the H-2 resonance decreases faster than the other resonances which all have much shorter T_1 values.

The procedure of incremental analysis (Borer et al. 1975) was developed to aid in the assignment of longer sequences where a nucleotide residue is likely to be repeated. This procedure compares the spectra of a sequence related series of oligoribonucleotides, where each member of the series is incremented one nucleotide unit from the one before. The shortest sequence is assigned by standard methods. For example, $\vec{A}G$, $\vec{A}G\vec{A}$, $\vec{A}G\vec{A}C$ and $AG\vec{A}CU$ are the sequences made in preparing the pentamer $AG\vec{A}CU$. Each intermediate can be studied separately by ^{1}H-NMR and used to assign the subsequent intermediate by comparison. Changes in the spectra from one intermediate to the next longer one are due to the appearance of the resonances from the additional nucleotide and its shielding effects on the previously present protons. Figure 8 displays this type of incremental assignment and is based on a similar representation by Borer (Borer et al. 1975).

However, in the number of sequences dealt with in this work, guidelines had to be further modified and refined to ensure unambiguous assignment of similar resonances and are explained below. The resonances of the nonexchangeable base and ribose H-1' protons could be divided into two groups: a low field group (8.4 - 7.5 ppm) containing the AH-8, AH-2,

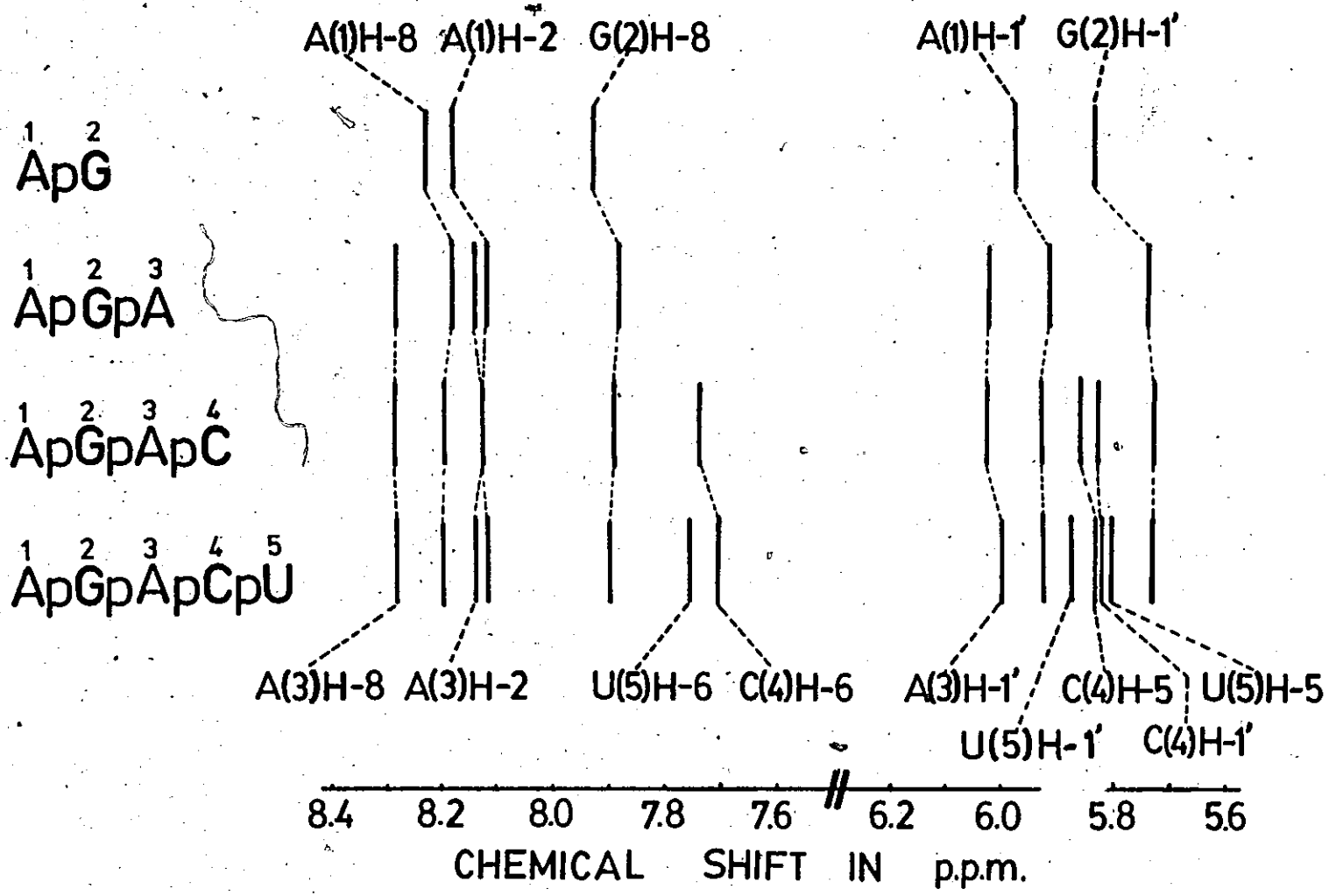


Figure 8. Incremental assignment Scheme for all the base and H-1' resonances of ApGpApCpU.

GH-8, CH-6 and UH-6 resonances, and a high field group (6.1 - 5.3 ppm) containing the CH-5, UH-5 and the ribose H-1' signals. Within these two groups several 'rules' of assignment (at 71 ± 2°C) held for all the compounds in this study. For the low field group, the signals from low to high field were in the order: AH-8, AH-2, GH-8 and U, CH-6. Within the higher field group the AH-1' signal was the lowest field signal of the whole group.

As stated earlier, the coupling constants were also useful in the analysis of the spectra. The ribose H-1' resonances had splittings (largely due to coupling with H-2') which were in general temperature sensitive and largest at high temperatures. The pyrimidine H-5 and H-6 protons displayed temperature invariant $J_{5,6}$ couplings of ∿7.6 Hz (C) and ∿8.2 Hz (U). Although the H-6 resonances were relatively easy to distinguish at low field, the H-5 resonances often overlapped the H-1' resonances at high field and thus were sometimes difficult to sort out. In such cases, pulsed homonuclear decoupling experiments were used to correlate coupled H-5 and H-6 protons in ambiguous situations. In this procedure the H-6 resonance is irradiated causing the collapse of the coupling and making it easier to pinpoint.

The assignment of the dimers $\vec{A}G$ and $\vec{C}A$ could be easily made on the basis of the above guidelines and agreed with those of Sarma and coworkers (Lee et al. 1976; Ezra et al. 1977). Many of the resonances of the trimers (Table 5) could be assigned by comparison with the resonances for $\vec{A}G$ and $\vec{C}A$ as in the method of incremental analysis of Borer, together with the guidelines outlined above. However, the assignment of some signals in the spectra was quite difficult and

occasionally required reference to compounds outside the $\vec{A}\vec{G}X$ and $\vec{C}\vec{A}X$ series. The examples below illustrate the type of arguments used. The 71°C spectrum of $\vec{C}\vec{A}U$ was assigned on the basis of the above guidelines with exception of the CH-1' and UH-1' protons which could not be differentiated by the preceeding arguments alone. The signal at 5.765 ppm (Table 5) was assigned to CH-1' on the basis that it should be similar to the value in $\vec{C}A$ because the addition of a 3'-uridine residue, with its low ring current intensity (Geissner-Prettre et al. 1976), would have little effect on the C or A shifts. Resonances with very similar chemical shifts (5.765, 5.772 and 5.764 ppm) were found in the 71°C spectra of $\vec{C}\vec{A}UG$, $CA\vec{U}GU$ and $CA\vec{U}GA$, respectively, all of which were assigned to the CH-1' (Romaniuk et al. 1978a & b). A more difficult case was presented by $\vec{A}\vec{G}G$. The assignment of all the adenosine resonances was facile but it was difficult to distinguish between the protons of G(2) and G(3). The only method of assignment, at that time, was a comparison with the spectra of $\vec{C}\vec{G}G$ and $\vec{C}\vec{C}GG$ which were already assigned (Romaniuk, 1979). In $\vec{C}\vec{G}G$ and $\vec{C}\vec{C}GG$ the terminal GH-8 protons resonated at 7.968 and 7.948 ppm respectively, at 71°C. Thus, the signal at 7.935 in the spectrum of $\vec{A}\vec{G}G$ (Table 5) was assigned to G(3)H-8. The G(3)H-1' in $\vec{A}\vec{G}G$ was assigned similarly. The remaining difficult cases, presented by $\vec{C}\vec{A}C$, $\vec{C}\vec{A}A$ and $\vec{A}\vec{G}A$ were also assigned with the aid of model compounds outside the $\vec{C}\vec{A}X$ and $\vec{A}\vec{G}X$ series.

The use of incremental analysis, when applied to longer sequences should be used with caution, for example, the assignment of the guanosine resonances in $\vec{A}\vec{G}G$.

TABLE 5

THE CHEMICAL SHIFTS, δ, IN ppm OF THE NON-EXCHANGEABLE BASE PROTONS AND RIBOSE H-1' PROTONS IN THE NUCLEOTIDE SERIES CpApX AND ApGpX AT 71 ± 2°C

PROTON	NUCLEOTIDE SEQUENCE[a]									
	CpA	CpApA	CpApG	CpApC	CpApU	ApG	ApGpA	ApGpG	ApGpC	ApGpU
A(1)H-8						8.238	8.191	8.212	8.242	8.240
A(2)H-8	8.377	8.277	8.303	8.369	8.372					
A(3)H-8		8.294					8.286			
A(1)H-2						8.186	8.126	8.167	8.180	8.185
A(2)H-2	8.260	8.128	8.189	8.222	8.245					
A(3)H-2		8.186					8.149			
A(1)H-1'						5.967	5.922	5.946	5.974	5.971
A(2)H-1'	6.093	5.952	6.003	6.071	6.083					
A(3)H-1'		6.032					6.033			
G(2)H-8						7.942	7.891	7.906	7.926	7.944
G(3)H-8			7.926					7.935		
G(2)H-1'						5.842	5.736	5.785	5.812	5.841
G(3)H-1'			5.832					5.836		
C(1)H-6	7.660	7.620	7.639	7.679	7.666					
C(3)H-6				7.740					7.744	
C(1)H-5	5.958	5.865	5.897	5.894	5.910					
C(3)H-5				5.869					5.876	
C(1)H-1'	5.779	5.743	5.752	5.768	5.765					
C(3)H-1'				5.850					5.886	
U(3)H-6					7.739					7.753
U(3)H-5					5.743					5.774
U(3)H-1'					5.849					5.867

[a]All sequences are written and numbered starting from the 5'-end to the 3'-end.

It will be noticed that the initial assignment of all sequences was carried out at high temperatures (71°C). The data were recorded at this temperature to avoid any secondary interactions and multistrand aggregation (Ts'o et al. 1969; Martin et al. 1971).

4.1.2.1. Shielding Trends

The ^{1}H chemical shifts data obtained for series $\vec{A}GX$ and $\vec{C}AX$ (Table 5) were analyzed in terms of the effect produced on the dimer ($\vec{A}G$ or $\vec{C}G$) protons when a 3'-terminal nucleotide, pX, was added to the dimer. Table 6 contains the shieldings produced in the protons of $\vec{A}G$ when a 3'-terminal nucleotide, pX, was added. Table 7 contains the corresponding data for the $\vec{C}AX$ series. The results showed an almost unbroken trend in the shielding for protons of the two dimers. The addition of the third nucleotide, pX, to the 3'-end of the dimers produced an X-dependent shielding of the dimer protons which decreased in the order A>G>C>U. In certain cases the anisotropic effect of pU and pC were such that they produced a deshielding effect on some of the protons; e.g. the C(1)H-6 proton in $\vec{C}AU$. The shielding trend observed was expected on the basis of the ring-current intensities of the bases (Geissner-Prettre et al. 1976). These same trends were present in the dimerization data of Sarma and coworkers (Lee et al. 1976; Ezra et al. 1977) who have discussed their work in terms of purines and pyrimidines.

The data (Tables 6 & 7) show the shielding trend operates at the nearest-neighbour and the next-nearest neighbour positions. Both Kroon (Kroon et al. 1974) and Geissner-Prettre (Geissner-Prettre, et al. 1976) have predicted a maximal shielding of 0.2 ppm for a proton with an adenosine next-nearest-neighbour (ca. 6.8 Å separation of the base

TABLE 6

SHIELDINGS[a], $\Delta\delta_S$, IN ppm OBSERVED FOR THE PROTONS IN ApG UPON ADDITION OF A 3'-TERMINAL NUCLEOTIDE pX[b]

	NUCLEOTIDE X			
PROTON	A	G	C	U
A(1)H-8	-0.047	-0.026	+0.004	+0.002
A(1)H-2	-0.060	-0.019	-0.006	-0.001
A(1)H-1'	-0.045	-0.021	+0.007	+0.004
G(2)H-8	-0.051	-0.036	-0.016	+0.002
G(2)H-1'	-0.106	-0.057	-0.030	-0.001

[a]Negative values indicate proton shielding, positive values deshielding.

[b]All shieldings are calculated for 71 ± 2°C.

TABLE 7

SHIELDINGS[a], $\Delta\delta_S$, IN ppm OBSERVED FOR THE PROTONS IN CpA UPON ADDITION OF A 3'-TERMINAL NUCLEOTIDE pX[b]

	NUCLEOTIDE X			
PROTON	A	G	C	U
C(1)H-6	-0.040	-0.021	+0.019	+0.006
C(1)H-5	-0.093	-0.061	-0.064	-0.048
C(1)H-1'	-0.036	-0.027	-0.011	-0.014
A(2)H-8	-0.100	-0.074	-0.008	-0.005
A(2)H-2	-0.132	-0.071	-0.038	-0.015
A(2)H-1'	-0.141	-0.090	-0.022	-0.010

[a]Negative values indicate proton shielding, positive values deshielding.

[b]All shieldings calculated for 71 ± 2°C.

planes). The maximum next-nearest-neighbour shielding observed in this work was 0.093 ppm for the C(1)H-5 proton in $\vec{CAA}$. Next-nearest-neighbour shieldings are predicted to drop off slowly as the position of the affected nucleus is moved (in a plane at 6.8 Å separation) away from the adenine ring centres (Geissner-Prettre et al. 1976). The observed maximum shielding was lower than the theoretical maximum value probably because, at 71°C, the trimers were largely unstacked and had considerable internal flexibility. Thus, any given proton may not, under these circumstances, experience the full shielding of the adenine rings. In addition, a variety of different conformations are presumably present, including bulge-loop conformations, which may contribute to the observed next-nearest-neighbour effects by allowing the next-nearest-neighbour bases to come into close proximity. Under these conditions of high conformational flexibility, the NMR shifts are an average of all shielding and deshielding effects, and it would be very unlikely that a given proton would show the full theoretical shielding effect of a particular base ring.

Using the tables of shielding effects of the four major bases, it is possible to predict the effect the addition of a particular base to the end of a sequence will have on the chemical shifts of other protons in the sequence. For instance, the chemical shifts of the G(3) protons in $\vec{AGGA}$ were predicted from a knowledge of the shifts of the protons in $\vec{AGG}$ together with the data from Table 6 which showed the effects of adding a 3'-terminal nucleotide, pA, to $\vec{AG}$. Table 8 lists the predicted and observed chemical shifts for the G(3) protons $\vec{AGGA}$. In this case, agreement was so good that a distinction between the two sets of G protons could be made on the basis of the shift parameters alone.

TABLE 8

PREDICTED AND OBSERVED CHEMICAL SHIFT OF THE G(3) PROTONS IN AGGA AND CGGA AT 71°C

PROTON	PREDICTED SHIFT (ppm)	OBSERVED SHIFT (ppm)
AGGA		
G(3)H-8	7.884	7.883 or 7.874
G(3)H-1'	5.730	5.726
CGGA		
G(3)H-8	7.917	7.913
G(3)H-1'	5.767	5.751

Table 8 also lists the predicted and observed chemical shifts for the G(3) protons in $C\vec{G}GA$.

The study with $A\vec{G}X$ and $C\vec{A}X$ gave an indication of the shielding effects due to the addition of a nucleotide, incrementally, to a sequence. These incremental shielding effects are, to a first approximation, due only to the nature of the added nucleotide. The second part of this study consisted of determining the substitutive shielding effects which occur when a nucleotide such as Ap is inserted into an oligoribonucleotide such as XpY ..., e.g.:

$$XpY \rightarrow XpApY$$

Substitutive shielding effects are manifested as changes in the chemical shifts of the protons of the X and Y residues. Consider the protons on the Xp nucleotide. The changes in chemical shift of its protons are due to changing its nearest-neighbour from Y to A and adding a next-nearest-neighbour pY nucleotide.

The sequences studied were $\vec{A}G$, $A\vec{X}G$, $\vec{C}G$, $C\vec{X}G$, $C\vec{A}UG$, $CA\vec{X}UG$, $A\vec{G}C$, $A\vec{G}XC$, $A\vec{G}CU$ and $AG\vec{X}CU$ (where X = A, G, C or U). The spectra obtained for these sequences were analyzed according to the guidelines described previously and are found in Tables 9-12. Although the sequences $C\vec{A}UG$ and the $CA\vec{X}UG$ series had been previously assigned and reported elsewhere (Romaniuk et al. 1978a,b, 1979a,b). they were essential to this study as well. The actual discussions of assignment of the other sequences are presented in other sections.

The chemical shift data in Tables 9-12 were analyzed in terms of the shielding effects produced on the lower-field unexchangeable protons when a nucleotide unit X was inserted into the strand. Table 13 shows

TABLE 9

THE CHEMICAL SHIFTS, δ, IN ppm OF THE NON-EXCHANGEABLE BASE PROTONS AND THE RIBOSE H-1' PROTONS IN THE SERIES ApXpG AND CpXpG[a] AT 71 ± 2°C

PROTON	NUCLEOTIDE SEQUENCE									
	ApG	ApApG	ApGpG	ApCpG	ApUpG	CpG	CpApG	CpGpG	CpCpG	CpUpG
A(1)H-8	8.238	8.191	8.212	8.264	8.276					
A(1)H-2	8.186	8.091	8.167	8.170	8.198					
A(1)H-1'	5.967	5.894	5.946	6.008	6.023					
A(2)H-8		8.243					8.303			
A(2)H-2		8.142					8,189			
A(2)H-1'		5.963					6.003			
G(2)H-8	7.942		7.906			7.999		7.942		
G(2)H-1'	5.842		5.785			5.891		5.803		
G(3)H-8		7.901	7.935	7.952	7.963		7.926	7.968	7.984	7.977
G(3)H-1'		5.800	5.836	5.834	5.847		5,832	5.873	5.884*2	5.888
C(1)H-6						7.686	7.639	7.662	7.750*1	7.766*3
C(1)H-5						5.978	5.897	5.946	7.976	5.995
C(1)H-1'						5.817	5.752	5.811	5.819	5.865
C(2)H-6				7.688					7.750*1	
C(2)H-5				5.850					5.954	
C(2)H-1'				5.848					5.884*2	
U(2)H-6					7.730					7.766*3
U(2)H-5					5.771					5.831
U(2)H-1'					5.838					5.865

[a]All sequences are written and numbered starting from the 5'end.

[b]*i - overlap of two resonances i.

TABLE 10

THE CHEMICAL SHIFTS, δ, IN ppm OF THE NON-EXCHANGEABLE BASE PROTONS AND THE RIBOSE H-1' PROTONS IN THE SERIES CpApXpUpG at 71 ± 2°C

PROTON	NUCLEOTIDE SEQUENCES				
	CpApUpG	CpApApUpG	CpApGpUpG	CpApCpUpG	CpApUpUpG
C(1)H-6	7.661	7.623	7.641	7.672	7.659
C(1)H-5	5.912	5.865	5.891	5.889	5.917
C(1)H-1'	5.772	5.750	5.759	5.776	5.770
A(2)H-8	8.349	8.255†	8.294	8.352	8.359
A(2)H-2	8.196	8.079	8.160	8.183	8.206
A(2)H-1'	6.037	5.928	5.975	6.042	6.062
A(3)H-8		8.268†			
A(3)H-2		8.134			
A(3)H-1'		5.974			
G(3)H-8			7.906		
G(3)H-1'			5.845		
C(3)H-6				7.709	
C(3)H-5				5.850	
C(3)H-1'				5.824	
U(3)H-6					7.716†
U(3)H-5					5.764
U(3)H-1'					5.842*
U(4)H-6	7.692	7.654	7.703	7.709	7.731†
U(4)H-5	5.739	5.695	5.767	5.798	5.834
U(4)H-1'	5.814	5.776	5.790	5.824	5.842*
G(5)H-8	7.958	7.942	7.958	7.978	7.978
G(5)H-1'	5.843	5.830	5.845	5.871	5.871

†Assignments may be reversed.

*Coincidence.

TABLE 11

THE CHEMICAL SHIFTS, δ, IN ppm OF THE NON-EXCHANGEABLE BASE PROTONS AND THE RIBOSE H-1' PROTONS IN THE SERIES ApGpXpC AT 71 ± 2°C

PROTON	NUCLEOTIDE SEQUENCE				
	ApGpC	ApGpApC	ApGpGpC	ApGpCpC	ApGpUpC
A(1)H-8	8.242	8.198	8.206	8.247	8.240
A(1)H-2	8.180	8.133*	8.150	8.168	8.185
A(1)H-1'	5.974	5.934	5.939	5.979	5.973
G(2)H-8	7.926	7.904	7.916	7.927	7.950
G(2)H-1'	5.812	5.730	5.770	5.785	5.826
A(3)H-8		8.289			
A(3)H-2		8.133*			
A(3)H-1'		6.026			
G(3)H-8			7.929		
G(3)H-1'			5.816		
C(3)H-6	7.744			7.722	
C(3)H-5	5.876			5.823	
C(3)H-1'	5.886			5.873	
U(3)H-6					7.748
U(3)H-5					5.776
U(3)H-1'					5.876
C(4)H-6		7.735	7.750	7.791	7.787
C(4)H-5		5.864	5.871	5.978	6.009
C(4)H-1'		5.834	5.888	5.886	5.901

*Coincidence.

TABLE 12

CHEMICAL SHIFTS[a] OF THE OLIGORIBONUCLEOTIDES, ApGpXpCpU, 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	7.737	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, using t-butyl alcohol-d as an internal reference, and are accurate to ±0.005 ppm.

[b] pD 7.0.

the shielding experienced by the protons of the dimers $\vec{A}pG$ and $\vec{C}pG$ when a nucleotide X was inserted generating the trimers $Ap\vec{X}pG$ and $Cp\vec{X}pG$. Consider the $Ap\vec{X}pG$ series. All the A(1) protons now had a nearest-neighbour A instead of a G, which, as ring-current calculations (Geissner-Prettre et al. 1976) and experiments have shown, leads to greater shielding. (In terms of shielding ability, the order of the bases is A>G>C>U). The shielding was enhanced by the presence of the next-nearest G. The contribution of the next-nearest G was discerned by considering X = G. The shielding effects observed for the A(1) protons were mainly due to the next-nearest-neighbour G effect because with $\vec{A}pG \rightarrow Ap\vec{G}pG$, any additional effects on A(1) must be due to G(3). When X = C, more complicated effects arose. The A(1) protons now had a nearest-neighbour C which, when compared to G, shielded less (Table 13)(Geissner-Prettre et al. 1976); thus deshielding of the A(1) resonances was expected. However, the next-nearest-neighbour G shielding effect (observed when X = G) acted in the opposite direction. The resultant effect was dependent on nucleoside geometry, with the A(1)H-8 and A(1)H-1' protons being deshielded and the A(1)H-2 proton still being shielded. When X = U, the next-nearest-neighbour effect of the G no longer compensated for the very poor nearest-neighbour shielding ability (Table 13) of uridine (Geissner-Prettre et al. 1976), and consequently all the A(1) protons were deshielded in AUG relative to AG. A smaller rationalization of the shielding effect observed for the G(3) protons in AXG and also C(1) and G(3) protons in CXG was made.

Table 14 shows the changes in chemical shift of the C(1), A(2), U(4) and G(5) protons on going from $C\vec{A}UG$ to $CA\vec{X}UG$. That the substitutive

TABLE 13

A TABLE OF SHIELDINGS[a] Δδ IN ppm OBSERVED FOR THE PROTONS IN ApG AND CpG UPON INSERTION OF AN INTERNAL NUCLEOTIDE Xp AT 71 ± 2°C

PROTON	NUCLEOTIDE SEQUENCE							
	ApApG	ApGpG	ApCpG	ApUpG	CpApG	CpGpG	CpCpG	CpUpG
A(1)H-8	-0.047	-0.026	+0.026	+0.038				
A(1)H-2	-0.095	-0.019	-0.016	+0.010				
A(1)H-1'	-0.073	-0.021	+0.041	+0.056				
G(3)H-8	-0.041	-0.007	+0.010	+0.021				
G(3)H-1'	-0.042	-0.006	-0.008	-0.001				
C(1)H-6					-0.047	-0.024	+0.064	+0.080
C(1)H-5					-0.081	-0.032	-0.002	+0.017
C(1)H-1'					-0.065	-0.006	+0.002	+0.048
G(3)H-8					-0.073	-0.031	-0.015	-0.022
G(3)H-1'					-0.059	-0.012	-0.007	-0.003

[a]Negative values indicate proton shielding, positive values deshielding.

TABLE 14

A TABLE OF SHIELDINGS[a] Δδ IN ppm OBSERVED FOR THE PROTONS IN CpApUpG UPON INSERTION OF A NUCLEOTIDE X INTO THE MIDDLE OF THE SEQUENCE[b]

PROTON	NUCLEOTIDE SEQUENCE			
	CpApApUpG	CpApGpUpG	CpApCpUpG	CpApUpUpG
C(1)H-6	-0.038	-0.020	+0.011	-0.002
C(1)H-5	-0.047	-0.021	-0.023	+0.005
C(1)H-1'	-0.022	-0.013	+0.004	-0.002
A(2)H-8	-0.094	-0.055	+0.003	+0.010
A(2)H-2	-0.117	-0.036	-0.013	+0.010
A(2)H-1'	-0.109	-0.062	+0.005	+0.025
U(4)H-6	-0.038	+0.011	+0.017	+0.039
U(4)H-5	-0.044	+0.028	+0.059	+0.095
U(4)H-1'	-0.039	-0.024	+0.010	+0.028
G(5)H-8	-0.016	0.000	+0.020	+0.020
G(5)H-1'	-0.013	+0.002	+0.028	+0.028

[a] negative values indicate proton shielding

[b] at 71 ± 2°C.

shielding effect operated at the next-nearest-neighbour level could be seen by observing the shielding of all the C(1) protons when X = A. This resulted from the much stronger next-nearest-neighbour shielding effect of an A on the C compared with a U on the C. For the A(2) and U(4) protons the substitutive shielding effects diminished and became deshielding as X was changed from A to G to C to U. This paralleled the result for the $C\vec{X}G$ and $A\vec{X}G$ series. The small shieldings observed for the G(5) protons when X = A were possible evidence for the next-next-nearest-neighbour effect by the A on the terminal G. The effect was small, as expected.

Table 15 shows the changes in chemical shift of the A(1), G(2) and C(4) protons on going from $A\vec{G}C$ to $A\vec{G}XC$. Again, the substitutive shielding effects predictably fell off and became deshielding as X was changed from A to G to C to U. The deshielding produced on the C(4) protons for X = U was very pronounced, since in this case a nearest-neighbour G and next-nearest-neighbour A were replaced by a nearest-neighbour U, next-nearest-neighbour G and a next-next-nearest-neighbour A. Since the latter effect was probably very small and the other two effects were deshielding, especially the effect involving the difference between a nearest-neighbour U and a nearest-neighbour G, it was not surprising to observe large substitutive deshielding effects.

Table 16 shows the experimental chemical shift changes caused by going from $A\vec{G}CU$ to $AG\vec{X}CU$ for the A(1), G(2), C(4) and U(5) protons. These are very similar to the changes observed going from $A\vec{G}C$ to $A\vec{G}XC$ (Table 15).

TABLE 15

A TABLE OF SHIELDINGS[a,b] IN ppm OBSERVED FOR THE PROTONS IN ApGpC UPON INSERTION OF A NUCLEOTIDE X[c]

PROTONS	NUCLEOTIDE SEQUENCE			
	ApGpApC	ApGpGpC	ApGpCpC	ApGpUpC
A(1)H-8	-0.044	-0.036	+0.005	-0.002
A(1)H-2	-0.047	-0.030	-0.012	+0.005
A(1)H-1'	-0.040	-0.035	+0.005	-0.001
G(2)H-8	-0.022	-0.010	+0.001	+0.024
G(2)H-1'	-0.082	-0.042	-0.027	+0.014
C(4)H-6	-0.009	+0.006	+0.047	+0.043
C(4)H-5	-0.012	-0.005	+0.102	+0.133
C(4)H-1'	-0.052	-0.002	0.000	+0.015

[a] negative values indicate proton shielding

[b] at 71 ± 2°C.

[c] i.e. ApGpC → ApGpXpC

TABLE 16

A TABLE OF SHIELDINGS[a,b] $\Delta\delta$ IN ppm OBSERVED FOR THE PROTONS IN ApGpCpU UPON INSERTION OF A NUCLEOTIDE X INTO THE MIDDLE OF THE SEQUENCE[c]

PROTON	NUCLEOTIDE SEQUENCE			
	ApGpApCpU	ApGpGpCpU	ApGpCpCpU	ApGpUpCpU
A(1)H-8	-0.052	-0.038	-0.003	-0.011
A(1)H-2	-0.055	-0.012	-0.006	+0.009
A(1)H-1'	-0.047	-0.036	-0.001	-0.005
G(2)H-8	-0.026	-0.015	-0.003	+0.024
G(2)H-1'	-0.061	-0.011	-0.008	-0.011
C(4)H-6	-0.026	+0.015	+0.072	+0.063
C(4)H-5	-0.023	+0.002	+0.129	+0.164
C(4)H-1'	-0.071	+0.007	+0.001	+0.021
U(5)H-6	-0.011	+0.004	+0.028	+0.025
U(5)H-5	-0.020	+0.002	+0.022	+0.034
U(5)H-1'	-0.018	0.000	+0.001	+0.035

[a] negative values indicate shielding

[b] all values for 71 ± 2°C

[c] i.e. ApGpCpU → ApGpXpCpU

In order to test the generality of the shielding parameters obtained in the present study, the parameters derived for $\vec{A}\vec{G}C \rightarrow \vec{A}\vec{G}XC$ (Table 15), together with the chemical shift data for the protons in $\vec{A}\vec{G}CU$ (Table 12) were used to predict the chemical shifts of the A(1), G(2) and C(4) protons in $\vec{A}\vec{G}\vec{X}CU$ (Table 17). Thus the shift for the A(1)H-8 proton in $\vec{A}\vec{G}\vec{A}CU$ was calculated using the A(1)H-8 chemical shift in $\vec{A}\vec{G}CU$ and the deshielding parameter derived for $\vec{A}\vec{G}C \rightarrow \vec{A}\vec{G}AC$ (i.e. 8.248 - 0.044 = 8.204 ppm) and compared to that found experimentally for A(1)H-8, 8.196 ppm. Most of the predicted values were within 0.02 ppm of the observed chemical shifts, which was considered good agreement. The assumption that $\vec{A}\vec{G}C \rightarrow \vec{A}\vec{G}XC$ modeled the change $\vec{A}\vec{G}CU \rightarrow \vec{A}\vec{G}\vec{X}CU$ appeared justified.

The use of these substitutive shielding effects as well as the incremental shielding effects enable the researcher to predict changes in chemical shifts due to an additional nucleotide residue in the sequence and thus are a tremendous aid in the unambiguous assignment of resonances in the spectra. With the constantly increasing library of completely assigned sequences, prediction of chemical shifts for any given sequence should become possible.

Additional aids to the assignment of signals were obtained by studying the temperature dependence of the various protons, over the temperature range 10°C - 70°C on the sequences $\vec{A}\vec{G}X$ and $C\vec{A}X$. Table 18 gives the chemical shift difference $\Delta\delta T$ (ppm) between 71°C and 10°C for all the nonexchangeable base and ribose H-1' protons. Figure 9 a-d shows the full variable temperature behaviour of these protons in the two respective compounds, $\vec{A}\vec{G}X$ and $C\vec{A}X$. For some protons the plots of

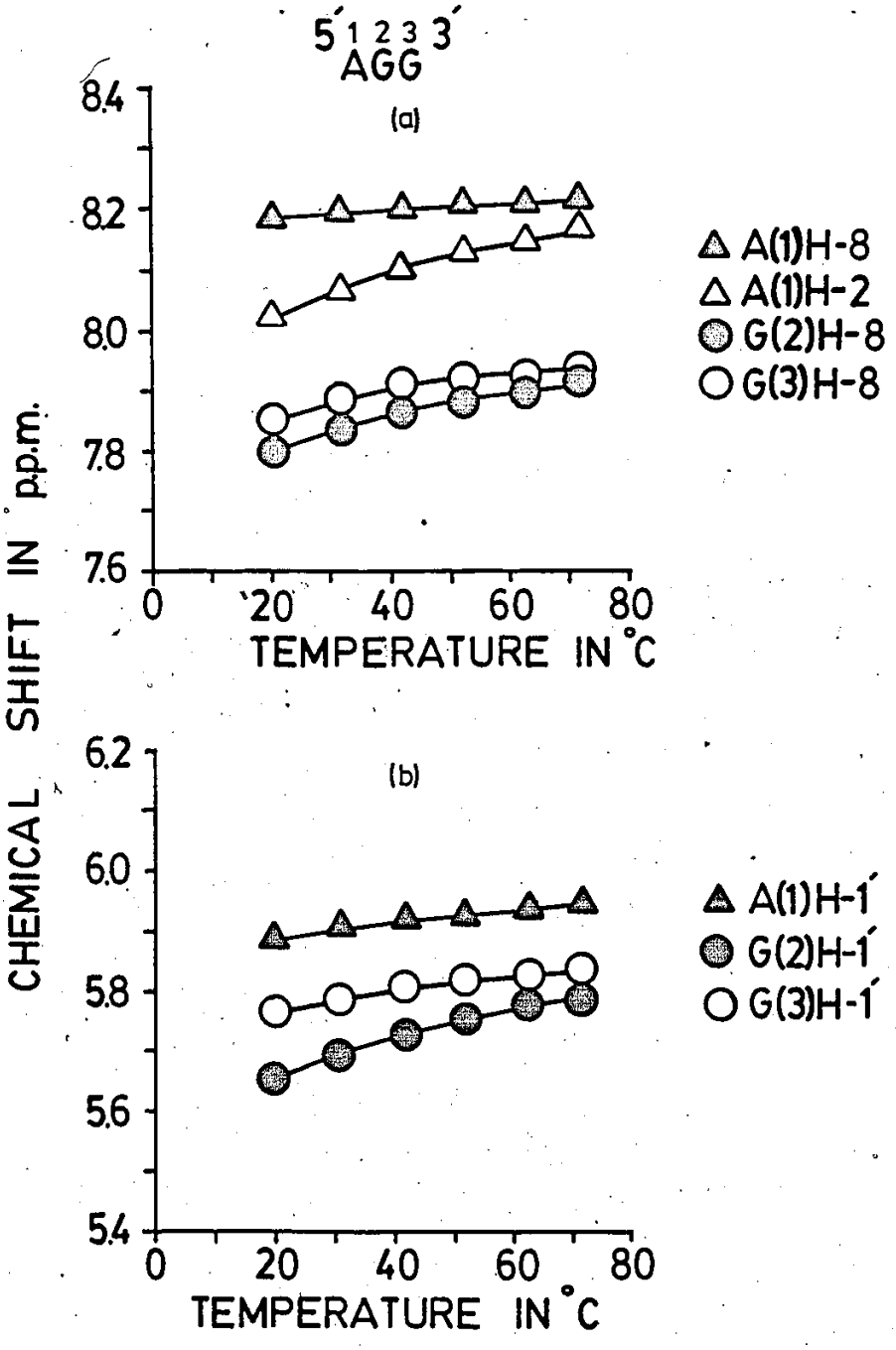


Figure 9(a&b). Variable temperature results for the aromatic and anomeric protons in ApGpG.

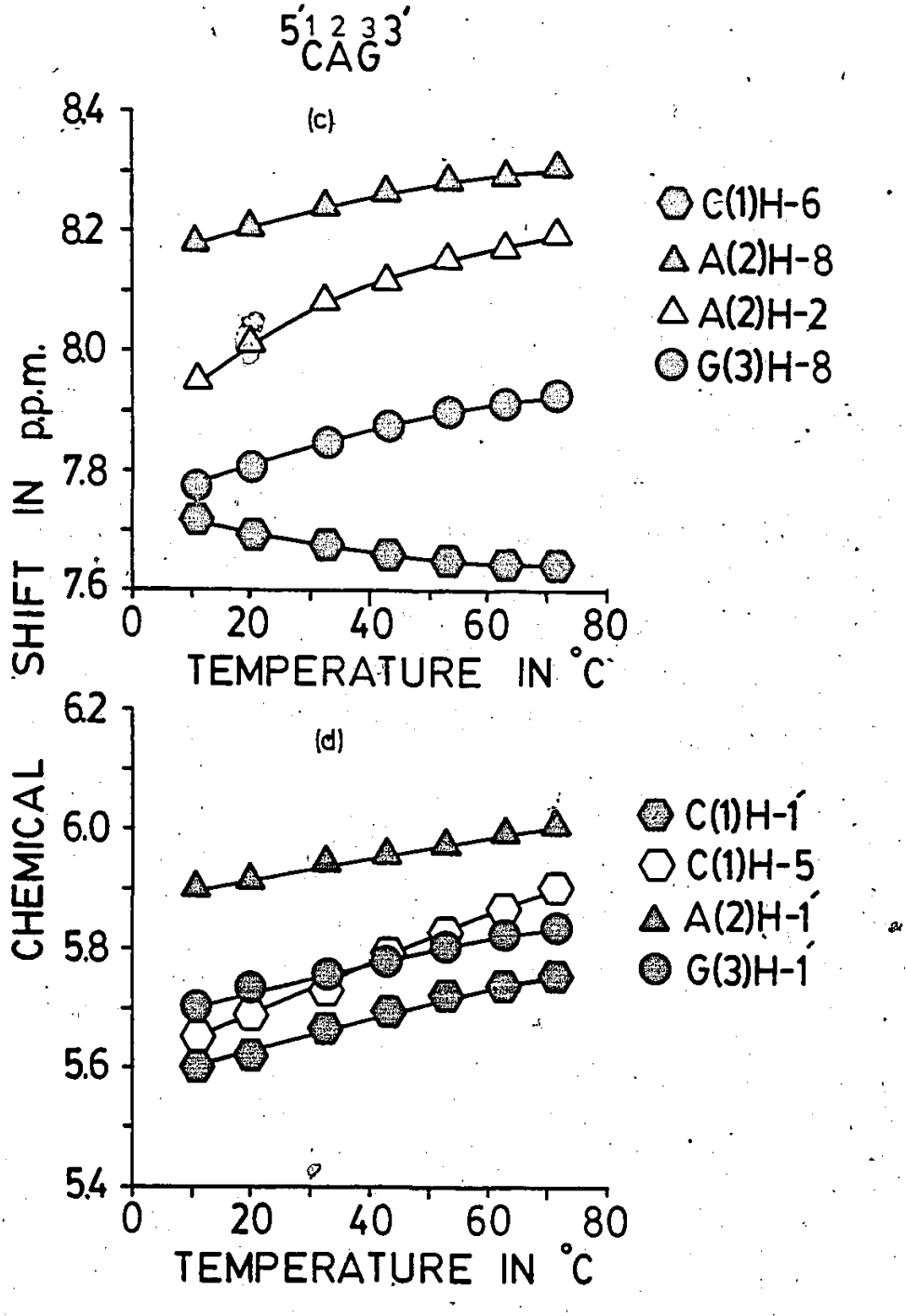


Figure 9(c&d). Variable temperature results for the aromatic and anomeric protons in CpApG.

TABLE 17

OBSERVED AND CALCULATED CHEMICAL SHIFTS δ IN ppm FOR THE NON-EXCHANGEABLE BASE PROTONS AND THE RIBOSE H-1' PROTONS IN THE SERIES ApGpXpCpU[a]

PROTON	NUCLEOTIDE SEQUENCE							
	ApGpApCpU		ApGpGpCpU		ApGpCpCpU		ApGpUpCpU	
	Observed	Calculated	Observed	Calculated	Observed	Calculated	Observed	Calculated
A(1)H-8	8.196	8.204	8.210	8.212	8.245	8.253	8.237	8.246
A(1)H-2	8.121	8.129	8.164	8.146	8.170	8.164	8.185	8.181
A(1)H-1'	5.931	5.938	5.942	5.943	5.977	5.983	5.973	5.977
G(2)H-8	7.903	7.907	7.914	7.919	7.926	7.930	7.953	7.953
G(3)H-1'	5.731	5.710	5.839	5.750	5.784	5.765	5.878	5.806
C(4)H-6	7.711	7.728	7.752	7.743	7.809	7.784	7.800	7.780
C(4)H-5	5.827	5.838	5.852	5.845	5.979	5.952	6.014	5.983
C(4)H-1'	5.821	5.840	5.919	5.890	5.893	5.892	5.924	5.907

[a] At 71 ± 2°C

TABLE 18

CHEMICAL SHIFT DIFFERENCES, $\Delta\delta_T$[a], IN ppm BETWEEN ∿71 AND ∿11°C FOR THE NON-EXCHANGEABLE BASE PROTONS AND THE RIBOSE H-1' PROTONS

PROTON	NUCLEOTIDE SEQUENCE									
	CpA[b]	CpApA	CpApG	CpApC	CpApU[b]	ApG	ApGpA	ApGpG[b]	ApGpC	ApGpU
A(1)H-8						-0.048	-0.062	-0.029	-0.017	-0.026
A(2)H-8	+0.007	-0.081	-0.123	-0.075	-0.052					
A(3)H-8		-0.137					-0.124			
A(1)H-2						-0.176	-0.241	-0.144	-0.189	-0.183
A(2)H-2	-0.054	-0.186	-0.237	-0.094	-0.049					
A(3)H-2		-0.094					-0.153			
A(1)H-1'						-0.101	-0.077	-0.059	-0.078	-0.060
A(2)H-1'	-0.015	-0.079	-0.102	-0.067	-0.037					
A(3)H-1'		-0.120					-0.089			
G(2)H-8						-0.085	-0.120	-0.111	-0.225	-0.173
G(3)H-8			-0.154					-0.084		
G(2)H-1'						-0.096	-0.155	-0.132	-0.223	-0.175
G(3)H-1'			-0.134					-0.072		
C(1)H-6	+0.021	+0.046	+0.080	+0.110	+0.093					
C(3)H-6				-0.123					-0.083	
C(1)H-5	-0.123	-0.293	-0.249	-0.231	-0.183					
C(3)H-5				-0.301					-0.313	
C(1)H-1'	-0.105	-0.139	-0.155	-0.148	-0.123					
C(3)H-1'				-0.199					-0.108	
U(3)H-6					-0.079					-0.054
U(3)H-5					-0.228					-0.252
U(3)H-1'										-0.069

[a] negative values indicate proton shielding, positive values deshielding

[b] for this compound the low temperature limit was ca. 18°C.

chemical shift versus temperature were straight lines, however, most of the plots gently curved to high field with decreasing temperature. Occasionally a plot went gradually to lower field with decreasing temperature, however, this was generally a sequence effect. From the variable temperature results some general conclusions, useful in the assignment procedure, could be made:

i) The adenosine H-2's were usually the most sensitive to temperature change in comparison to the H-8's.

ii) The pyrimidine H-5's were more sensitive than the H-6's.

iii) If a uridine or cytidine was present in a molecule, its shielding usually exhibited the greatest shielding of any proton in the molecule.

The temperature dependence of chemical shifts of the protons in these compounds probably arises from the temperature dependence of the stacking-destacking equilibria (Kan et al. 1973), although solvent effects may also be significant. As the temperature is lowered and the trimers become more stacked, suitably oriented protons will experience the shielding influence of the base rings to a greater extent and move to higher field, as is observed for most of the protons.

The evaluation of the shielding effects, from the four major bases in a nearest- and next-nearest-neighbour position, on the chemical shifts of the protons on the other bases in a given sequence, coupled with the incremental analysis procedure (Borer et al. 1975), has proven to be invaluable in the assignment of the large number of sequences used in this study. These methods have proven to be both reliable (reproducibility of assignments in similar sequences) and very

efficient. The versatility of this assignment procedure will be demonstrated in the following sections.

4.2. Effects of Flanking G·C Base Pairs on Internal Watson-Crick, G·U and Non-Bonded Base Pairs within a Short RNA Duplex

(Alkema et al. 1981a, accepted by Biochemistry)

In an earlier study (Romaniuk et al. 1979) a series of pentaribonucleotides, $\vec{CAXUG}$ (where X = A, G, C or U), was studied to determine whether or not duplex formation occurred with a sequence containing a non-bonded internal base pair. The non-bonded base pair was flanked by two A·U base pairs and was found to totally disrupt duplex formation. Instead, the pentaribonucleotides, $\vec{CAXUG}$, preferred to adopt a stacked single-stranded conformation.

However, some early optical studies with longer sequences containing internal non-bonded cytidines found RNA double helices with bulge loops (Uhlenbeck et al. 1971; Gralla & Crothers, 1973). More recently a stable DNA duplex with a single non-bonded T-T opposition was studied (Haasnoot et al. 1979, 1980). Also a ^{1}H-NMR study of the complex formed between ethidium bromide and $\vec{CUG}$ (Lee & Tinoco Jr. 1978) revealed the formation of a small duplex where the mismatched bases adopted an extrahelical conformation. Although the $(\vec{CUG})_2$ minihelix only formed in the presence of ethidium bromide, it did suggest that non-bonded base oppositions could exist in small duplexes.

Since non-bonded base oppositions are a common feature of RNA secondary structures (Lomant & Fresco, 1975; Clarke, 1977), additional research into the possible formation of duplexes with non-bonded bases was considered valid and important. To this end a series of pentaribo-

nucleotides AGXCU (where X = A, G. C or U) was synthesized. Theoretically this sequence can form a duplex containing a 'looped-out' region:

$$\begin{matrix} & & X & & \\ AG & \cdot & \cdot & CU \\ UC & \cdot & \cdot & GA \\ & & X & & \end{matrix} \quad (X = A, G, C, U)$$

Such a duplex would be an ideal model to probe the effect small non-bonding regions have on the formation of adjacent base pairs. In addition, because in this sequence the non-bonded region is flanked by G·C base pairs, as opposed to A·U base pairs in CAXUG, comparison with the study on CAXUG (Romaniuk et al. 1979) would provide information on the effect that the composition of adjacent base pairs has on non-bonded regions. Comparison of the series of pentamers, AGXCU, with that of CAXUG, also allows for the investigation of sequence related effects to loop closure.

The four pentaribonucleotides, AGACU, AGUCU, AGGCU and AGCCU, were also 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). This study can also be compared to that with the CAXUG series (Romaniuk et al. 1979) and would be helpful in establishing rules of sequence stability. The study with an internal G·U base pair is significant in that in the duplex formed by AGGCU:AGUGU the G·U base pair is flanked by G·C pairs which may have a significant effect on its stability.

4.2.1. Results

Each of the four pentaribonucleotides was studied by ^{1}H-NMR at a

concentration of 6.3 mM, as were each of the mixing experiments. ^{1}H-NMR was used both to check sequence integrity and to observe changes in conformation as a function of temperature.

The low field nonexchangeable proton chemical shift assignments for the four pentamers were determined by two methods: comparison with NMR data on AGCU (Neilson et al. 1980, Table 12, Section 4.1) 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 ppm and 8.134 ppm, as well as a doublet at 5.998 ppm (cf. AGCU) 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 neighbouring residues, when compared to AGCU, was explained by the strong ring-current shielding effect of adenosine (Geissner-Prettre et al. 1976).

The two GH-8 resonances in AGGCU were difficult to distinguish since they occurred at 7.928 ppm and 7.914 ppm. The GH-8 resonance in AGCU occurs at 7.928 ppm (Neilson et al. 1980), however, the G(2)H-8 of AGGCU was assigned to the higher field resonance (7.914 ppm) because the G(3) residue would result in some shielding of its neighbours (Bell et al. 1981). The G(2)H-1' and G(3)H-1' signals were more readily distinguishable at 5.780 ppm and 5.839 ppm respectively. In the spectrum of AGUCU the additional resonances were readily assigned to the U(3) protons at 7.752 ppm (H-6), 5.878 ppm (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 ppm, 5.799 ppm and 5.844 ppm while the C(4)H-6, H-5 and H-1' protons occurred at 7.809 ppm, 5.979 ppm and 5.892 ppm. Simple comparison of this sequence with AGCU was insufficient, so the method of incremental analysis, developed by Borer (Borer et al. 1975) was used to confirm the assignments (Table 19).

TABLE 19

INCREMENTAL ANALYSIS OF AGCCU IN D_2O[a] AT 70°C

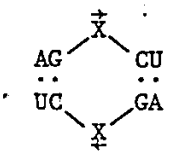
PROTON	AG	AGC	AGCC	AGCCU
		CHEMICAL SHIFTS[b]		
A(1)H-8	8.238	8.238	8.247	8.245
A(1)H-2	8.186	8.170	8.168	8.170
G(2)H-8	7.942	7.934	7.927	7.926
C(3)H-6		7.748	7.722	7.713
C(4)H-6			7.791	7.809
U(5)H-6				7.796
A(1)H-1'	5.967	5.969	5.979	5.977
G(2)H-1'	5.842	5.809	5.785	5.780
C(3)H-1'		5.889	5.873	5.844
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] pH 7.0, concentrations 9-12 mM

[b] Chemical shifts are in ppm relative to DSS using t-butyl alcohol-OD as an internal reference and are accurate to 40 ± 0.005 ppm.

4.2.1.1. Effect of a Non-Complementary Base Opposition on the Stability of a Short Duplex

Theoretically each AGXCU 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 10-13 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), the AGXCU series of compounds display significant upfield shifts. Some proton curves display sigmoidal behaviour which could indicate base pairing.

4.2.1.2. Studies on Duplex Formation in Complementary Sets

Before observing the formation and stability of an internal G·U base pair, the effects of other internal Watson-Crick base pairs on the T_m of AGCU were studied. 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° (averaged from the 15 sigmoidal curves) in Figure 14. This represents a T_m 11° 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 using the complementary set, AGGCU:AGCCU. Figure 15 shows the results of this experiment. Averaging

the 18 curves displaying sigmoidal character produced a T_m of 54°C, a full 20° higher than the parent tetramer duplex, $\vec{A}GCU$.

4.2.1.3. Formation of a G·U Wobble Base Pair within a Double Helix

By mixing the single strand pentamers, $AG\vec{G}CU$ and $AG\vec{U}CU$ to make the resultant duplex, it was possible to observe the formation of an internal G·U base pair. The sigmoidal behaviour of the chemical shift versus temperature plots, given in Figure 16, clearly indicates that duplex $AG\vec{G}CU$:$AG\vec{U}CU$ forms, and averaging the 10 curves gave a T_m of 44°C. Comparison with duplex $A\vec{G}ACU$:$AG\vec{U}CU$ (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 earlier studies with series $CA\vec{X}UG$ (Romaniuk et al. 1979) where a G·U base pair neither stabilized nor destabilized the duplex formed by $CA\vec{G}UG$:$CA\vec{U}UG$, which had a T_m of 24°C similar to that of the parent tetramer $C\vec{A}UG$ (T_m = 24°C).

4.2.2. 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 centre of instability. In Romaniuk's study the internal G·U base pair in the duplex $CA\vec{U}UG$:$CA\vec{G}UG$ was flanked by two Watson-Crick A·U base pairs. The present study is 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, $\vec{A}GCU$, is higher than that for duplex,

$\vec{C}A\vec{U}G$, by ∿10°C, the possibility of duplex formation by the single strands, $A\vec{G}\vec{X}\vec{C}U$ where X = A, G, C or U is conceivable.

Four pentaribonucleotide sequences, $A\vec{G}\vec{A}\vec{C}U$, $A\vec{G}\vec{U}\vec{C}U$, $A\vec{G}\vec{G}\vec{C}U$ and $A\vec{G}\vec{C}\vec{C}U$, were prepared. Each of these sequences has the potential to form a duplex of the type, $\begin{matrix} AG\vec{X}CU \\ UC_{\vec{X}}GA \end{matrix}$, with internal non-complementary base opposition and can be compared to the reference duplex $A\vec{G}CU:A\vec{G}CU$ (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) non-complementary base oppositions or (2) an additional base pair on duplex stability when flanked by G·C base pairs.

In studying the $A\vec{G}\vec{X}CU$ series some unexpected results were obtained. In Figure 10, a significant number of the chemical shift vs. temperature plots for $A\vec{G}\vec{A}CU$ 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). However, ^{1}H-NMR evidence of looping out of non-complementary base pairs is a downfield shift of the chemical shift with decreasing temperature similar to the effect on the GH-8 resonance for the 'looped out' G of AGA_n single stranded oligoribonucleotides (Shum, 1977). Therefore, either $A\vec{G}\vec{A}CU$ was forming a highly stacked single stranded structure, such as the rod-like 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 centre of the duplex. However, the upfield shifts

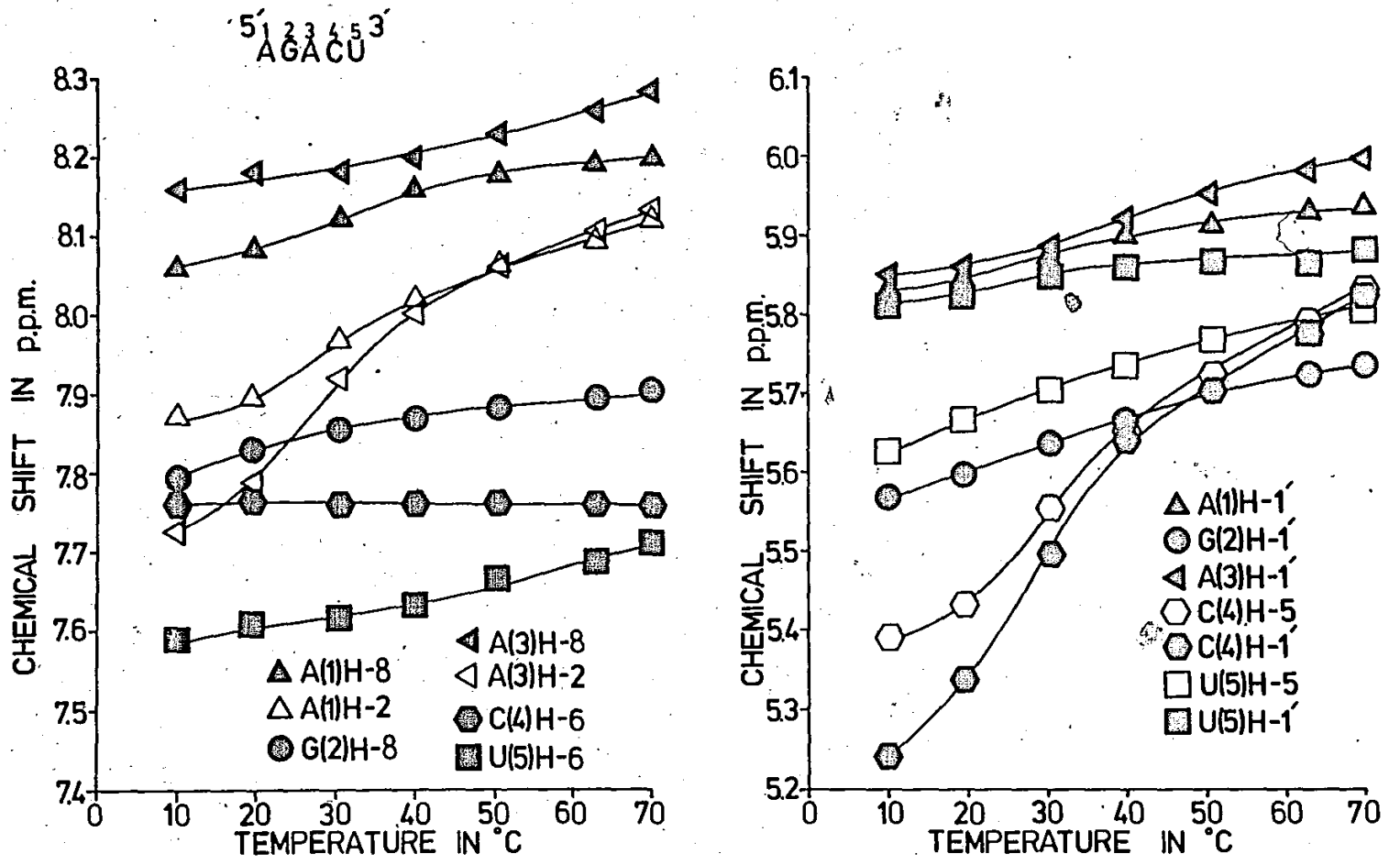


Figure 10. Variable temperature results for AGACU.

(deshielding) 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 behaviour and these bases could base pair. The A(3)H-2 also displays sigmoidal behaviour but cannot form a base pair. Adenine, however, stacks better than any other base and is perhaps drawn cooperatively into a stacked conformation allowing the rest of the sequence to duplex. Related studies (Alkema et al. 1981b) have shown that terminal non-paired adenosines can in fact determine stacking direction and duplex stability. Therefore, in duplex AGACU:AGACU it is likely that the central adenosine, instead of looping out of the sequence and thus disrupting the stacking and base pairing of its neighbours, 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 non-complementary base opposition. In a duplex containing 50% A·U and 50% G·U base pairs, one non-complementary base opposition for every 100 normal base pairs reduced the T_m by ∿1°C. Duplex AGACU:AGACU is 50% AU and 50% GC containing one non-complementary base pair opposition for four normal base pairs, which is equivalent to 25 in 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°) to be ∿29°C. From the average of the five sigmoidal curves in AGACU:AGACU (Figure 10) a T_m of 26°C was obtained which suggests that

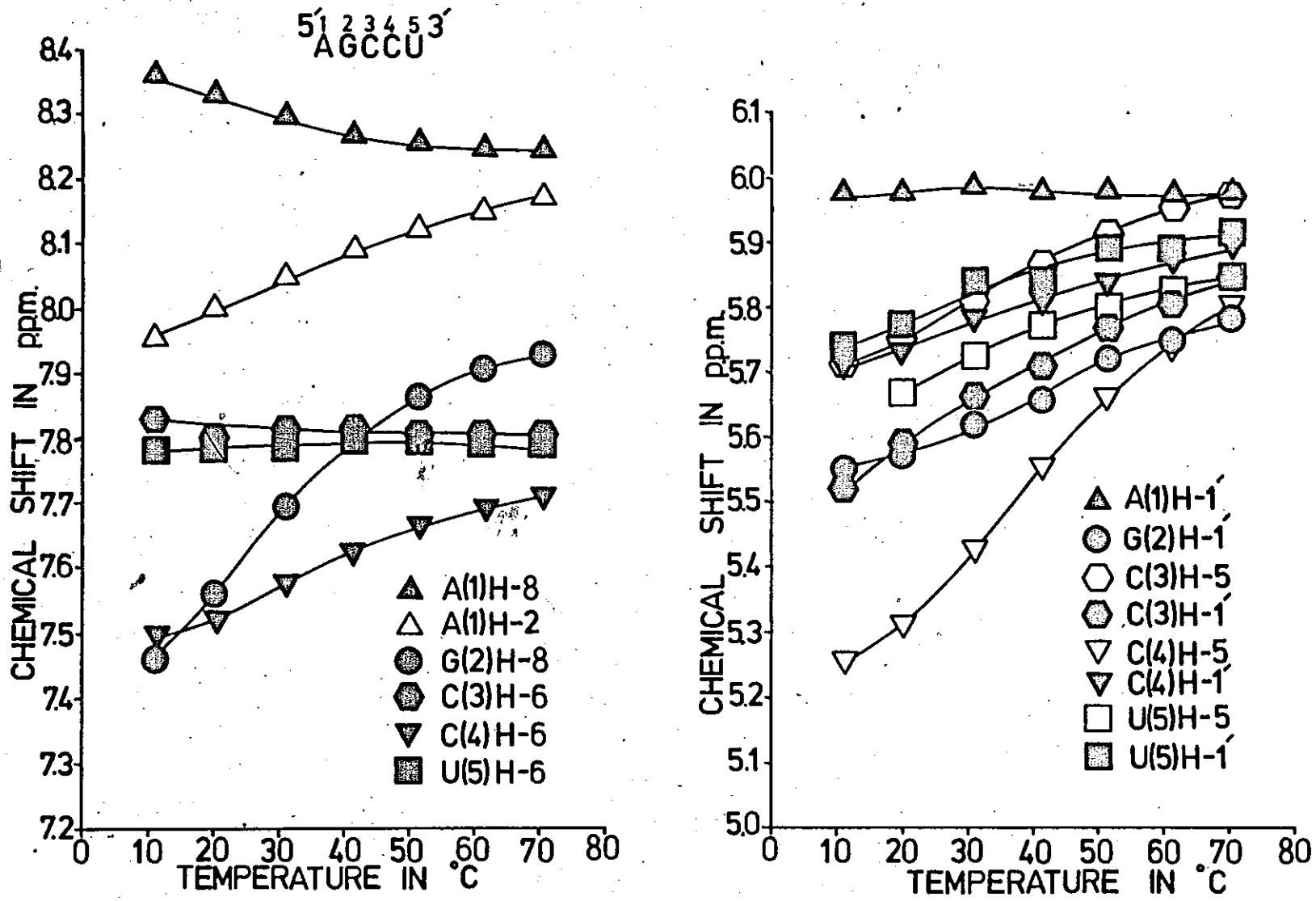


Figure 11. Variable temperature results for AGCCU.

AGACU may form a duplex with internal A·A oppositions that does not loop out but rather stacks within the duplex.

Figure 11 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 behaviour indicating duplex formation. These curves are more difficult to interpret because the C(3)H-6 curve shows a slight downfield shift, while the C(3)H-5 shows sigmoidal behaviour 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, duplex formation could occur with C·C opposition causing duplex instability. This duplex would also contain 50% G·C and 50% A·U base pair mixture with one non-complementary base pair opposition and four normal base pairs. Again, a T_m between 20°C and 29°C could be expected according to the guidelines of Lomant and Fresco (1975). In fact, an average of the five curves that show sigmoidal behaviour indicates a T_m of 25°C.

Figure 12 contains the chemical shift vs. temperature plots for AGGCU. Once again sigmoidal behaviour 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. Secondly, a duplex of the nature described by Lomant and Fresco (1975) could occur

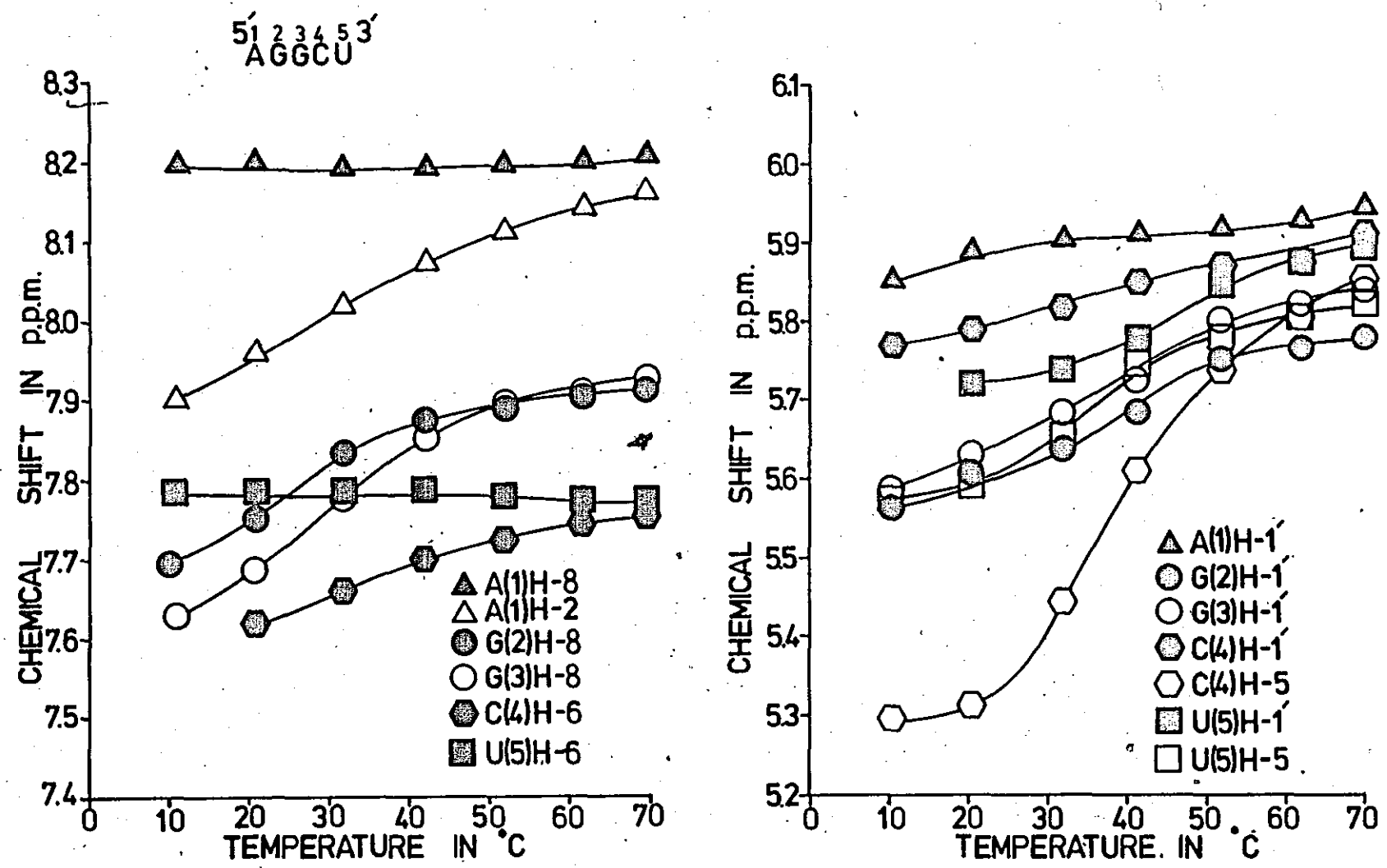


Figure 12. Variable temperature results for AGGCU.

resulting in a duplex of the type, $\overset{\rightarrow}{AGGCU}$ / UCGA / G, where one G residue is found in an extra-helical position. However, a third possible duplex, with two terminal 3'-non-bonded adenosines, similar to that formed with ACAUG (Neilson et al. 1980) could arise. This resultant duplex, $\overset{\rightarrow}{AGGCU}$ / UCGGA, 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 $\overset{\rightarrow}{GGCU}$:$\overset{\rightarrow}{GGCU}$, with two terminal G·U base pairs (Mizuno et al. 1981), forms with a T_m of ∿30°C (Alkema et al., unpublished observations). Therefore, it is unlikely that a duplex with G·G opposition forms from $\overset{\rightarrow}{AGGCU}$.

Figure 13 displays the chemical shift vs. temperature plot for $\overset{\rightarrow}{AGUCU}$. Unlike the other sequences this pentaribonucleotide shows absolutely no sigmoidal behaviour, only general upfield shifting of the resonances, despite a possible U·U opposition. Looping out of a 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. 1981b; Turner et al. 1981; Section 1), the presence of a non-bonded 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 $\overset{\rightarrow}{AGXCU}$ series of pentaribonucleotides, were entirely different from those obtained in an earlier study of the series of compounds, $\overset{\rightarrow}{CAXUG}$ (Romaniuk et al. 1979). A sequence effect was probably responsible, since the G·C base pairs are

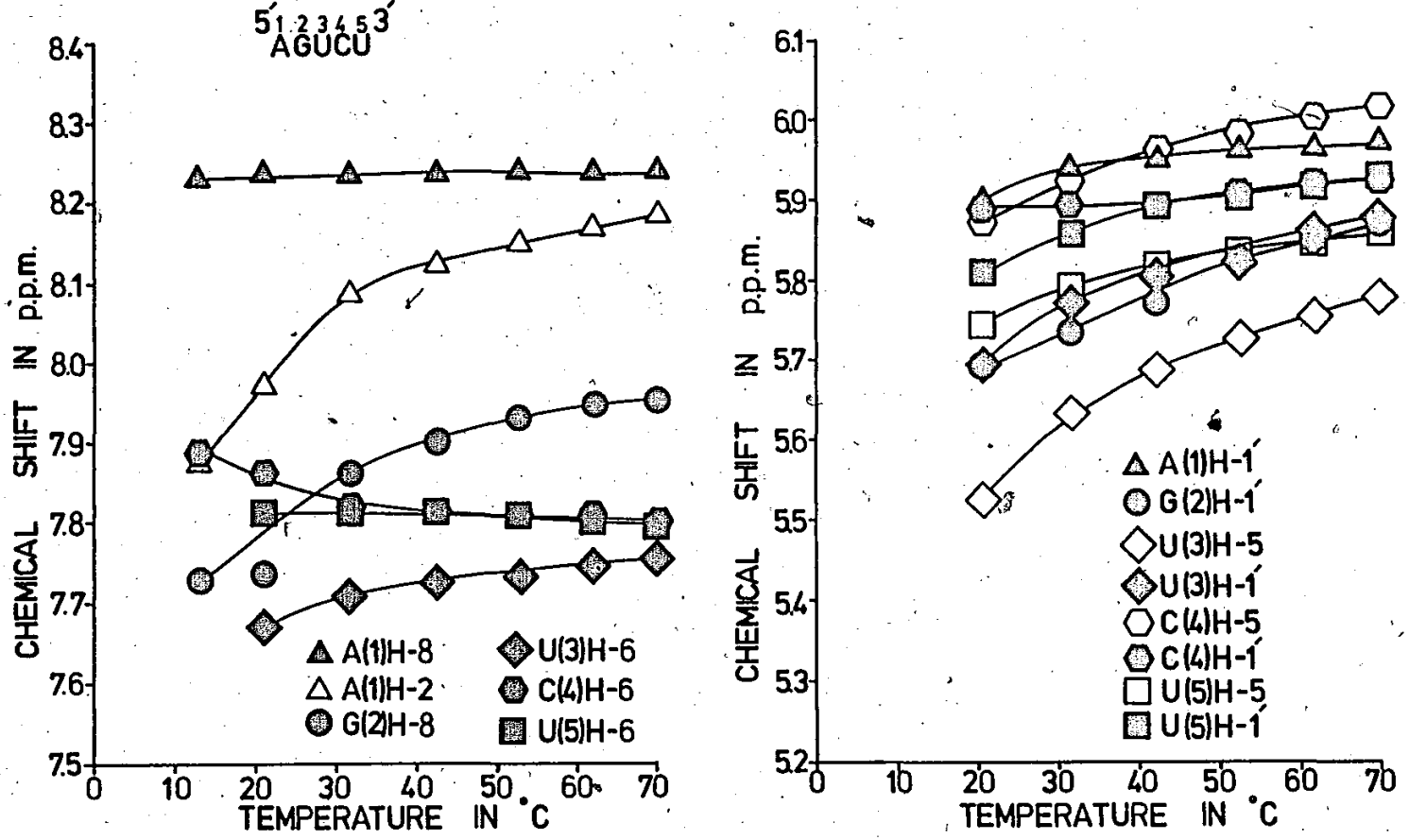


Figure 13. Variable temperature results for AGUCU.

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 Figures 14a and b. The T_m for AGACU:AGUCU represents an 11°C increase over that found for AGCU (Neilson et al. 1980), shown in Table 20. 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 fact that the A·U base pair is 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 $\overset{G-A}{\underset{C-U}{\cdot\cdot}}$ and $\overset{A-C}{\underset{U-G}{\cdot\cdot}}$ sequences are greater than for $\overset{A-U}{\underset{U-A}{\cdot\cdot}}$ and $\overset{U-U}{\underset{A-A}{\cdot\cdot}}$ 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 internal A·U base pair.

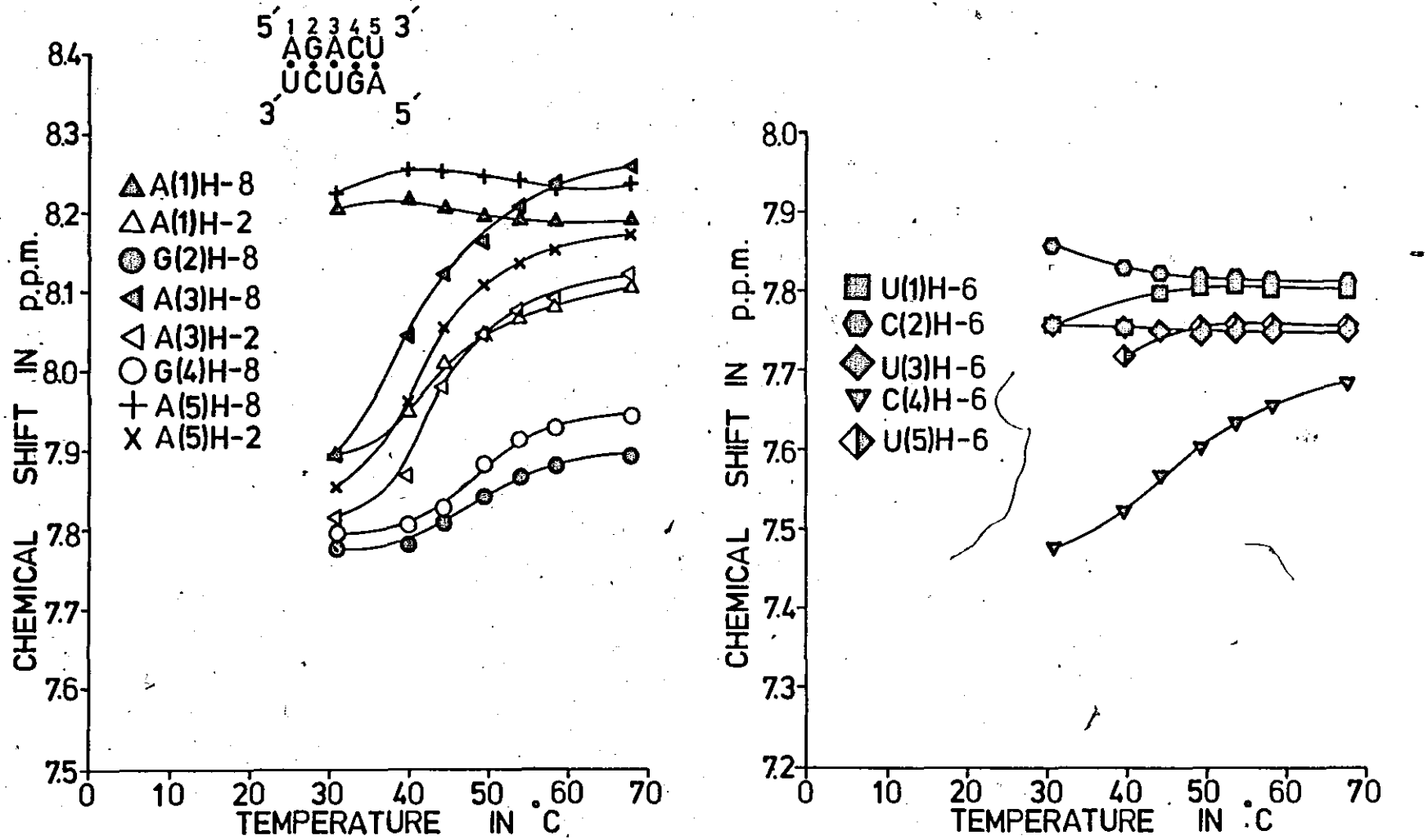


Figure 14a. Variable temperature plots for the aromatic base protons of duplex AGACU:AGUCU.

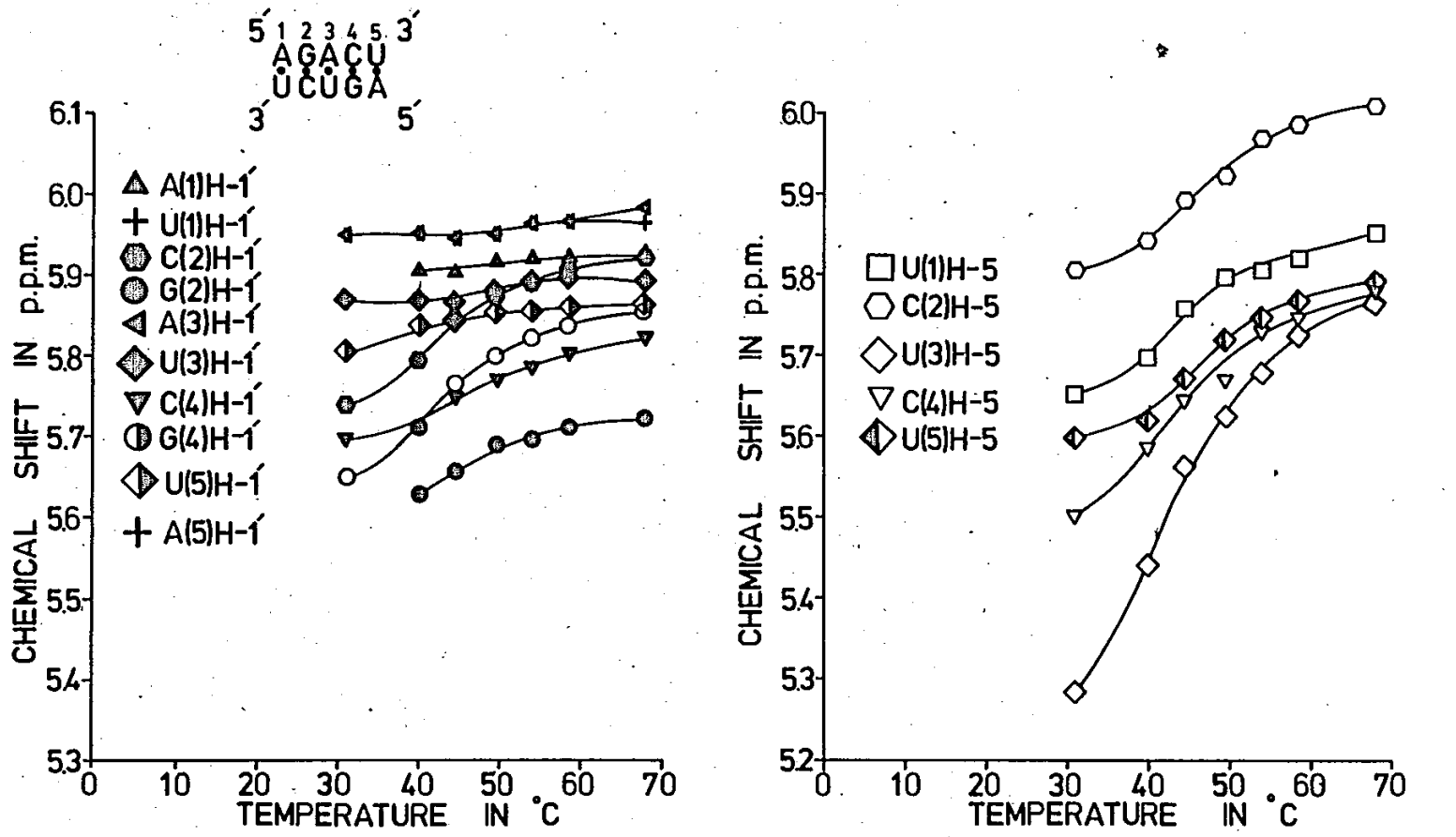


Figure 14b. Variable temperature plots for the anomeric H-1' and aromatic H-5 protons of duplex AGACU:AGUCU.

Figures 15a and b illustrate the chemical shift vs. temperature plots for the duplex, $A\vec{G}\vec{G}CU:A\vec{G}\vec{C}CU$. The average T_m for this duplex is 54°C, Table 20, and represents a 20°C increase over the T_m found for reference, $A\vec{G}CU$, 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 of duplex stability was demonstrated by inserting base pairs into different cores, for example, reference tetramer duplexes $A\vec{G}CU:A\vec{G}CU$ ($\overset{\rightarrow}{\underset{\leftarrow}{GC}}$ core) and $C\vec{A}UG:C\vec{A}UG$ ($\overset{\rightarrow}{\underset{\leftarrow}{AU}}$ core). Insertion of a G·C pair in a $\overset{\rightarrow}{\underset{\leftarrow}{GC}}$ core showed an increase of 6°C over a similar insertion into a corresponding $\overset{\rightarrow}{\underset{\leftarrow}{AU}}$ core (Table 20), 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, therefore, of a duplex containing a Watson-Crick base pair flanked by two G·C pairs will be 6°C greater than the corresponding duplex which contains the same base pairs flanked by two A·U pairs.

The two single strands, $A\vec{G}\vec{G}CU$ and $A\vec{G}\vec{U}CU$, gave a complementary duplex, $A\vec{G}\vec{G}CU:A\vec{G}\vec{U}CU$, containing an internal G·U base pair. Averaging the T_m's from the curves in Figure 16 gave a T_m of 44°C, which represents an increase in duplex stability of 10°C over the reference duplex, in contrast to $C\vec{A}\vec{G}UG:C\vec{A}\vec{U}UG$ where the GU base pair did not enhance

TABLE 20

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

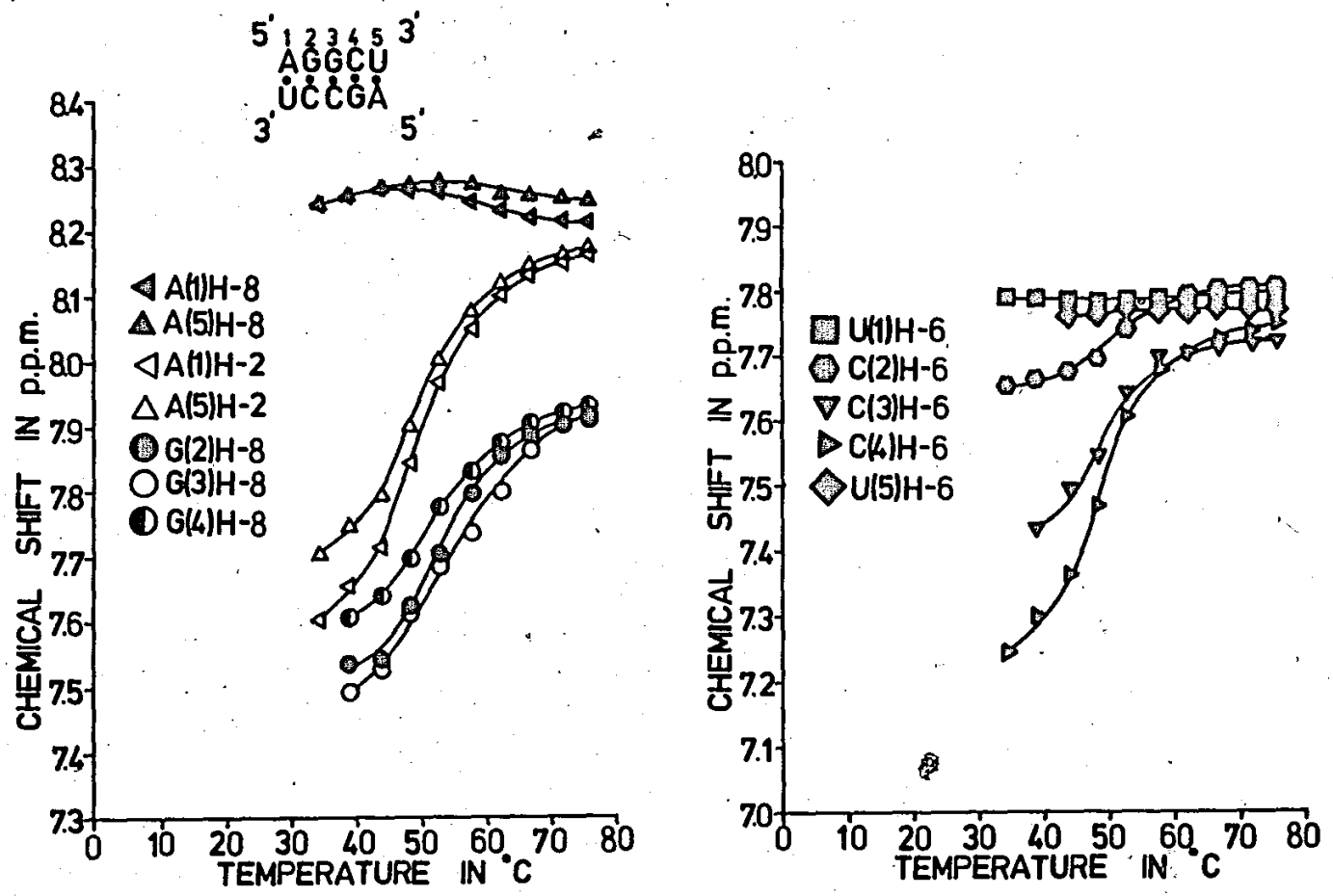
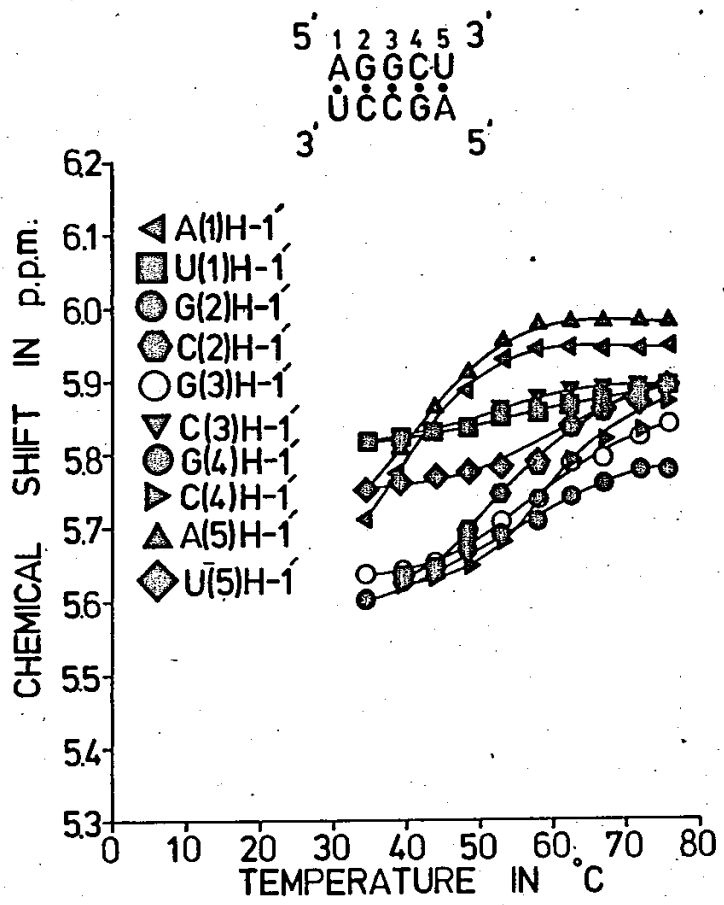


Figure 15a. Variable temperature plots for the aromatic base protons of duplex AGCCU:AGCCU.



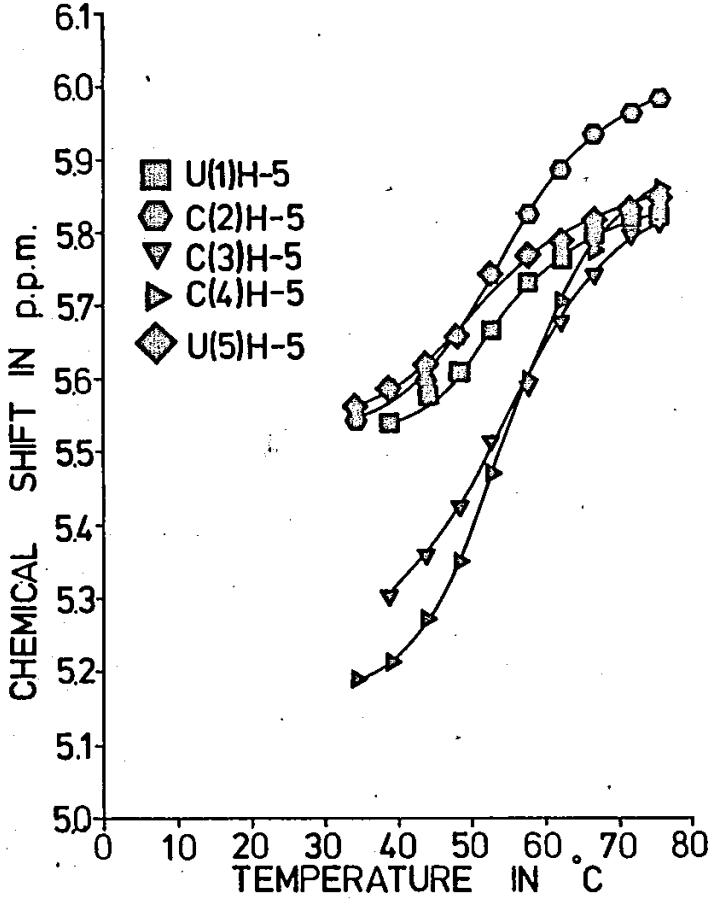


Figure 15b. Variable temperature plots for the anomeric H-1' and aromatic H-5 protons of duplex AGGCU:AGCCU.

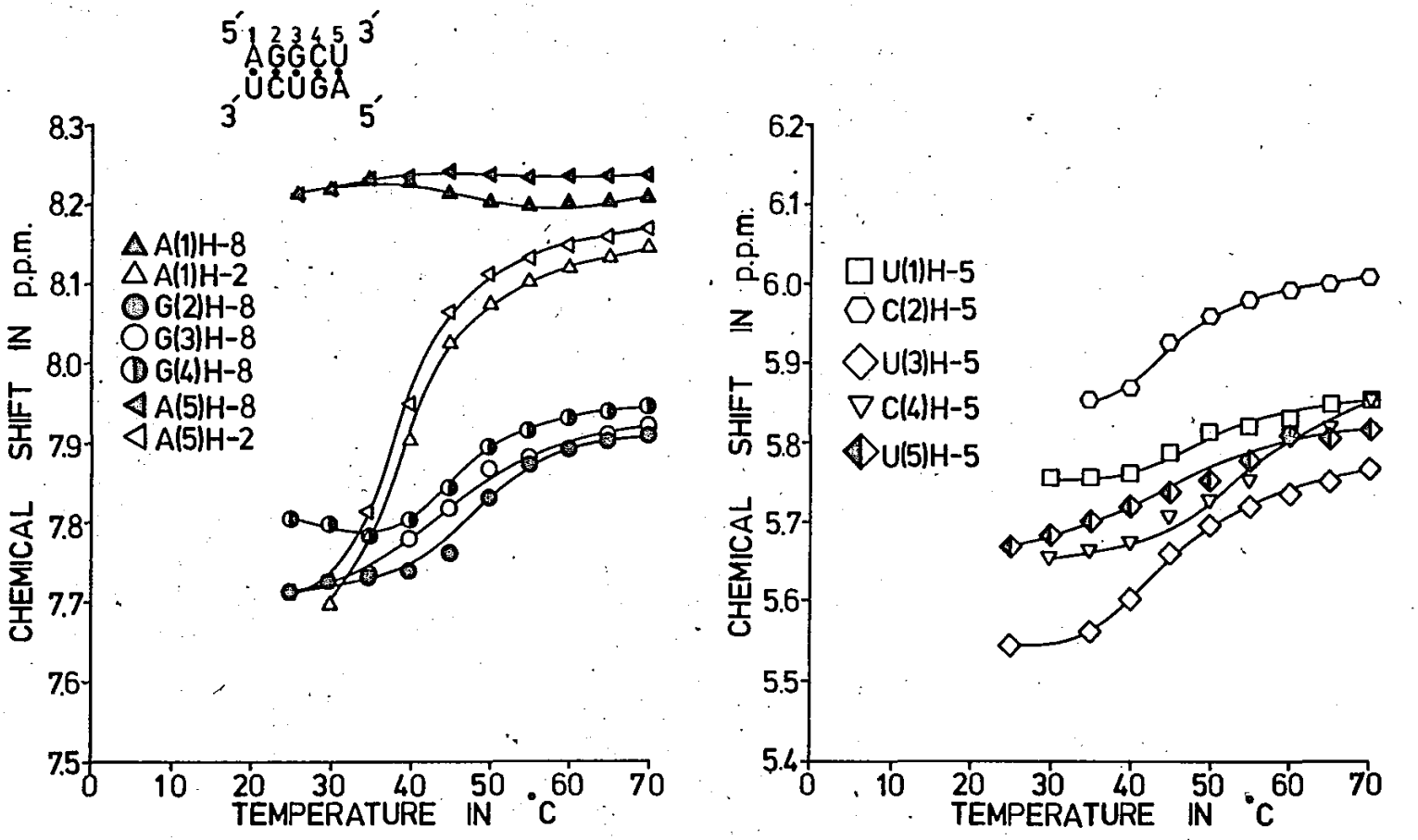


Figure 16. Variable temperature plots for the aromatic base protons of duplex AGGCU:AGUCU.

stability (Table 20) (Romaniuk et al. 1979). 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, especially, 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. Non-complementary base oppositions A·A and C·C exist within duplex AGXCU (where X = A or C). 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.

4.3 The Effects of Base Sequence and Dangling Bases on the Stability of Short Ribonucleic Acid Duplexes (Neilson et al. 1980).

The secondary structure of an RNA molecule consists of loops separated by Watson-Crick hydrogen bonded duplex regions (Fiers et al. 1975, 1976; Gross et al. 1978). Contribution to helical stability by stacked non-base paired regions adjacent to duplexes has been the subject of some research (Martin et al. 1971; Uhlenbeck et al. 1971; Tinoco et al. 1971; Tinoco et al. 1973). A dangling adenosine residue increases the melting temperature (T_m) of duplexes formed by oligoribonucleotides of sequence AmUn (Martin et al. 1971). Temperature jump

studies indicated longer non-base paired sequences enhance stability in short RNA helical regions (Grosjean et al. 1976; Yoon et al. 1976; detailed in Section 1).

Fraying of terminal base pairs in short duplex regions of RNA may also play a role in the stability of RNA-RNA interactions and the stability of large loops in secondary structures. Fraying effects were considered important in calculations of thermodynamic values derived from optical studies of the helix-coil transition of short $\vec{A}m\vec{U}n$ block oligonucleotides, although end effects would be exaggerated in short duplexes compared with those for longer sequences (Martin et al. 1971). Terminal A·U base pairs are known to be less stable than G·C base pairs (Uhlenbeck et al. 1971; Pohl, 1974; Grosjean et al. 1976). Direct evidence for fraying of terminal base pairs came from NMR studies on the helix-coil transition of the self-complementary hexanucleotides, $AA\vec{G}CUU$ (Borer et al. 1975; Kan et al. 1975) and d(ATGCAT) (Patel & Hilbers, 1975; Hilbers & Patel, 1975). On the other hand, NMR studies of short duplexes with terminal G·C base pairs have indicated negligible fraying (Arter et al. 1974; Hughes et al. 1978; Romaniuk et al. 1978).

The following work was a continuation of an earlier study of 3'-dangling uridines and adenosines (Romaniuk et al. 1978; Romaniuk, 1979) but introduces important new findings which must be considered in conjunction with these earlier studies. A set of three self-complementary tetramers, $C\vec{A}UG$ (T_m = 24°C), $A\vec{G}CU$ (T_m = 34°C) and $U\vec{G}CA$ (T_m = 33°C), each containing two G·C and two A·U pairs, but having considerably different properties, were used as reference duplexes. $C\vec{A}UG$:$C\vec{A}UG$ has terminal C·G base pairs which do not fray. $A\vec{G}CU$:$A\vec{G}CU$, however, contains terminal A·U

base pairs that exhibit indications of fraying. On the other hand, UĜCA:UĜCA which also has terminal A·U base pairs display no sign of fraying. Each tetramer was studied individually by ^{1}H-NMR spectroscopy. The corresponding duplexes containing terminal non-bonded nucleotides were then synthesized and monitored by NMR to observe the effects a dangling base has on duplex stability. A further look into the fraying phenomena was attempted by repeating the experiments with AAĜCUU (Borer et al. 1975; Kan et al. 1975) and then adding a 3'-dangling adenosine to determine whether or not dangling base residues would reduce fraying as previously suggested (Kallenbach & Berman, 1977).

4.3.1. Results

The following free oligoribonucleotides were used in this study: AĜCU, AGĈUA, AAĜCU, AAĜCUA, AAĜCUU, AAGĈUUA, CÂUG, CAÛGA, CAÛGU, ACÂUG, UCÂUG, UĜCA and UGĈAA. Sequences, AĜCU, AGĈUA, CÂUG, CAÛGA, CAÛGU (Romaniuk et al. 1978; Romaniuk, 1979) and AAĜCUU (Borer et al. 1975; Kan et al. 1975) have been described previously, but were an integral part of this study. The results reported here plus those in a following section firmly establish the role of terminal non-base paired residues in duplex stability.

The procedure of incremental analysis (Borer et al. 1975) was used to make the chemical shift assignments, which are summarized in Tables 21-23. The simultaneous irradiation of any pyrimidine H-6 resonances with the subsequent collapse of the corresponding pyrimidine H-5 resonances proved useful in the assignment of the anomeric (H-1') ribose resonances as well as the H-5 signals. Often UH-1' and CH-1' resonances were equivalent at 70°C, but separated at lower temperatures. At low

temperatures (<30°C), ribose H-1' proton signals were often observed as broad singlets (J1', 2' <1.0 Hz). Such a reduction in the coupling constant reflects a high degree of base stacking (Altona, 1975).

4.3.1.1. CÂUG Series (Table 21)

The resonances for CÂUG have been reported previously (Ganoza et al. 1979). Extension of the sequence to CÂUGA resulted in the appearance of two additional singlets and a doublet in the 70°C spectrum: A(5)H-8 (8.320 ppm), A(5)H-2 (8.189 ppm) and A(5)H-1' (6.060 ppm). Both AH-2 resonances are chemically equivalent at 70°C. Addition of 5'-adenosine residue resulted in an overall shielding throughout the sequence. Appearance of additional doublets, U(5)H-6 (7.828 ppm), U(5)H-5 (5.812 ppm) and U(5)H-1' (5,873 ppm) confirmed the spectrum of CÂUGU.

Extension of CÂUG to ACÂUG and UCÂUG resulted in the appearance of the expected additional resonances: A(1)H-8 (8.229 ppm), A(1)H-2 (8.129 ppm) and A(1)H-1' (5.985 ppm), and U(1)H-6 (7.762 ppm), U(1)H-5 (5.808 ppm) and U(1)H-1' (5.861 ppm) respectively. The shielding of the additional A residue extended part way along the sequence ACÂUG whereas the additional U residue caused a slight overall deshielding.

4.3.1.2. UĜCA Series (Table 22)

The resonances of ÛG were assigned by comparison with published data (Ezra et al. 1977). Extension to UĜC resulted in the appearance of three additional doublets in the 70°C spectrum: C(3)H-6 (7.809 ppm), C(3)H-1' (5.932 ppm) and C(3)H-5 (5.958 ppm).

The adenosine resonances of UĜCA spectrum were easily identified by the appearance of two additional singlets: A(4)H-8 (8.362 ppm) and

TABLE 21

CHEMICAL SHIFTS OF CAUG SERIES AT 70°C IN D_2O

PROTON	CAUG	CAUGA	CAUGU	ACAUG	UCAUG
			CHEMICAL SHIFTS[b]		
A(1)H-8				8.229	
A(2)H-8	8.354	8.343	8.352		
A(3)H-8				8.305	8.365
A(5)H-8		8.320			
A(1)H-2				8.129	
A(2)H-2	8.196	8.189	8.199		
A(3)H-2				8.134	8.177
A(5)H-2		8.189			
G(4)H-8	7.962	7.896	7.958		
G(5)H-8				7.950	7.963
C(1)H-6	7.665	7.644	7.662		
C(2)H-6				7.688	7.730
U(1)H-6					7.762
U(3)H-6	7.699	7.682	7.692		
U(4)H-6				7.688	7.705
U(5)H-6			7.828		
A(1)H-1'				5.985	
A(2)H-1'	6.043	6.047	6.037		
A(3)H-1'				5.997	6.047
A(5)H-1'		6.060			
G(4)H-1'	5.835	5.739	5.848		
G(5)H-1'				5.850	5.840
C(1)H-1'	5.773	5.783	5.772		
C(2)H-1'				5.842	5.814
C(1)H-5	5.905	5.923	5.912		
C(2)H-5				5.864	5.954
U(1)H-1'					5.861
U(3)H-1'	5.822	5.796	5.814		
U(4)H-1'				5.795	5.814
U(5)H-1'			5.873		
U(1)H-5					5.808
U(3)H-5	5.746	5.762	5.749		
U(4)H-5				5.731	5.733
U(5)H-5			5.813		

[a] pD = 7.0; concentrations: 9 - 12 mM

[b] Chemical shifts are in ppm relative to DSS using t-butyl alcohol-OD as an internal reference and are accurate to ±0.005 ppm.

TABLE 22

CHEMICAL SHIFTS OF UGCA SERIES AT 70°C IN D_2O[a]

PROTON	UG	UGC	UGCA	UGCAA
		CHEMICAL SHIFTS[b]		
A(4)H-8			8.362	8.270
A(5)H-8				8.278
A(4)H-2			8.202	8.069
A(5)H-2				8.164
G(2)H-8	8.004	8.015	7.975	7.966
C(3)H-6		7.809	7.711	7.692
U(1)H-6	7.723	7.731	7.716	7.710
A(4)H-1'			6.068	5.948
A(5)H-1'				6.020
G(2)H-1'	5.904	5.896	5.794	
C(3)H-1'		5.932	5.875	
C(3)H-5		5.958	5.853	5.838
U(1)H-1'	5.816	5.812	5.811	
U(1)H-5	5.828	5.832	5.811	5.807

Footnotes: see Table 21

A(4)H-2 (8.202 ppm) and a ribose doublet, A(4)H-1' (6.068 ppm). A general shielding trend extended along the sequence and agrees with the findings for $\vec{AGCU}$A.

Comparison of the 70°C spectrum of $\vec{UGCA}$A with its precursor $\vec{UGCA}$ shows the appearance of another set of adenine signals: A(5)H-8 (8.278 ppm), A(5)H-2 (8.164 ppm) and A(5)H-1' (6.020 ppm). Again a general shielding occurs.

4.3.1.2. $\vec{AGCU}$ Series (Tables 23a & b)

The resonances of $\vec{AG}$ were assigned by comparison with published data (Lee et al. 1976). Extension of the sequence to $\vec{AGC}$ resulted in the appearance of three additional doublets in the 70°C spectrum: C(3)H-6 (7.744 ppm), C(3)H-5 (5.876 ppm) and C(3)H-1' (5.886 ppm).

Aromatic base resonances of the $\vec{AGCU}$ spectrum were readily assigned by observing two additional doublets: U(4)H-6 (7.768 ppm), U(4)H-5 (5.821 ppm). Ribose U(4)H-1' coincided with the C(3)H-1' at 5.892 ppm.

Comparison of the 70°C spectrum of $\vec{AGCU}$A with that of $\vec{AGCU}$ shows the appearance of three new signals: A(5)H-8 (8.343 ppm), A(5)H-2 (8.186 ppm) and A(5)H-1' (6.063 ppm). Addition of the terminal A resulted in a general shielding of most of the resonances. The only difficulty this shielding caused was in the assignment of the AH-2 resonances. The more shielded resonance was assigned to A(1)H-2, which reflects the general shielding trend throughout the sequence.

The resonances attributable to the additional 5'-adenosine in A$\vec{AGCU}$ were routinely assigned by reference to the $\vec{AGCU}$ 70°C spectrum; A(1)H-8 (8.204 ppm), A(1)H-2 (8.084 ppm) and A(1)H-1' (5.891 ppm). This dangling adenosine results in some shielding of the neighbouring residues but is not as extensive as with the 3'-dangling adenosine in $\vec{AGCU}$A.

TABLE 23a

CHEMICAL SHIFTS OF THE AROMATIC PROTONS IN THE AGCU SERIES AT 70°C IN D_2O[a]

	CHEMICAL SHIFTS[b]							
PROTON	AG	AGC	AGCU	AGCUA	AAGCU	AAGCUA	@ 85°C AAGCUU	@ 85°C AAGCUUA
A(1)H-8	8.238	8.282	8.248	8.229	8.201	8.182	8.196	8.200
A(2)H-8					8.224	8.224	8.253	8.237
A(5)H-8				8.343				
A(6)H-8						8.323		
A(7)H-8								8.354
A(1)H-2	8.186	8.180	8.176	8.144	8.084	8.038	8.147	8.114
A(2)H-2					8.119	8.051	8.130	8.091
A(5)H-2				8.186				
A(6)H-2						8.114		
A(7)H-2								8.192
G(2)H-8	7.942	7.926	7.929	7.929				
G(3)H-8					7.844	7.810	7.880	7.866
C(3)H-6		7.744	7.737	7.704				
C(4)H-6					7.713	7.667	7.729	7.716
U(4)H-6			7.768	7.698				
U(5)H-6					7.767	7.686	7.767	7.738
U(6)H-6							7.783	7.716

[a] pD = 7.0; Concentrations: 9 - 12 mM

[b] Chemical shifts are in ppm relative to DSS using t-butyl alcohol-D as an internal reference and are accurate to ±0.005 ppm.

TABLE 23b

CHEMICAL SHIFTS OF THE ANOMERIC H-1' AND AROMATIC H-5 PROTONS OF THE AGCU SERIES AT 70°C IN D_2O[a]

	CHEMICAL SHIFTS[b]							
PROTON	AG	AGC	AGCU	AGCUA	AAGCU	AAGCUA	@ 85°C AAGCUU	@ 85°C AAGCUUA
A(1)H-1'	5.967	5.974	5.978	5.965	5.891	5.888	5.899	5.896
A(2)H-1'					5.930	5.922	5.960	5.942
A(5)H-1'				6.063				
A(6)H-1'						6.040		
A(7)H-1'								6.069
G(2)H-1'	5.842	5.812	5.791	5.798				
G(3)H-1'					5.710	5.695	5.768	5.742
C(3)H-1'		5.886	5.903	5.869				
C(4)H-1'					5.869	5.812	5.899	5.876
U(4)H-1'			5.892	5.817				
U(5)H-1'					5.869	5.790	5.899	5.830
U(6)H-1'							5.899	5.853
C(3)H-5		5.876	5.850	5.839				
C(4)H-5					5.788	5.730	5.866	5.825
U(4)H-5			5.821	5.793				
U(5)H-5					5.806	5.712	5.846	5.855
U(6)H-5							5.884	5.805

Footnotes - see Table 23a.

The spectrum for AAGCUA which contains both an additional 3' and 5' adenosines over AGCU was assigned by comparison to the spectra for AAGCU and AGCUA. The resonances corresponding to the A(1)H-8, A(1)H-2 and A(1)H-1' were assigned as 8.182 ppm, 8.053 ppm and 5.888 ppm respectively. The signals appearing at 8.232 ppm, 8.114 ppm and 6.040 ppm were assigned to A(6)H-8, A(6)H-2 and A(6)H-1', respectively.

The spectra for both AAGCUU and AAGCUUA were taken at 85°C to eliminate all secondary interactions which were in evidence, to a significant degree, even at 70°C. Use of the incremental procedure or comparison between similar sequences, to assign the spectral signals, depends on the assumption that specific inter- and intrastrand effects are minimized and that oligomers adopt a similar average conformation under this condition. Under these conditions the oligoribonucleotides are largely unstacked.

The spectrum for AAGCUU was assigned by comparison to AAGCU with the additional uridine signals appearing at: 7.783 ppm(U(6)H-6),5.884 ppm (U(6)H-5)and 5.899 ppm(U(6)H-1'). The U(6)H-1; resonance was coincidental with the U(5)H-1;, C(4)H-1 and A(1)H-1' signals but was assigned by integration of the peaks. Also the U(6)H-1' signal separated out at lower temperatures. The assignments were in agreement with those made by Borer (Borer et al. 1975).

Assignment of the additional adenosine residue in AAGCUUA was facile by comparison with AAGCUU. The adenosine signals were quite distinct and resonated at: 8.354 ppm (A(7)H-8), 8.192 ppm (A(7)H-2) and 6.069 ppm (A(7)H-1'). Once again the addition of an adenosine resulted in a shielding of many resonances, as previously noted with both AGCUA and UGCAA.

Chemical shift values for each aromatic and ribose signal were plotted against temperature (Figures 17-29). Cooperative behaviour of duplex formation was indicated by the sigmoidal nature of the curves. Table 24 contains the average T_m values for the duplexes from the $\vec{C}A\vec{U}G$, $U\vec{G}CA$ and $A\vec{G}CU$ series. Table 25 shows the T_ms for each curve in the $A\vec{G}CU$ and $AG\vec{C}UA$ duplexes and the average T_ms for the individual base pairs in both duplexes. Difference in G·C and A·U base pair values will be discussed later in the context of strand fraying.

4.3.2. Discussion

Secondary RNA structure results from more than mere Watson-Crick base pairing. Base stacking controlled by length and nearest neighbour interactions (i.e. sequence), non-complementary base-pairs, dangling bases, loop size and terminal fraying effects, also contribute to general stability. Optical studies can provide values for overall conformational change but fail to evaluate the algebraic contribution of each feature. Variable temperature ^{1}H-NMR spectroscopy has that capability. Already this methodology has been used to identify various non-bonded interactions and G-U base pairs as perturbations within short duplexes. The role of dangling bases in short RNA duplex stability will now be discussed.

The series of synthetic oligoribonucleotides $CA\vec{U}GA$, $CA\vec{U}GU$, $AC\vec{A}UG$, $UC\vec{A}UG$, each contain the common self-complementary sequence, $\vec{C}A\vec{U}G$, and an additional residue destined to become the 3'- or 5'-dangling base on duplex formation.

e.g. 2 $CA\vec{U}GA$ $\rightleftharpoons$ $\overset{CA\vec{U}GA}{\underset{AGUAC}{|||||}}$ (3'-dangling A duplex)

Stabilities of duplexes with 5'- and 3'-dangling bases, both adenine and uracil, were compared with the reference duplex, $\vec{C}A\vec{U}G:\vec{C}A\vec{U}G$ (Table 24). Clearly, the dangling bases increase stability, purines more than pyrimidines, with melting temperature differences of 12°C for adenine and 5°C for uracil (Figures 17-21).

The effect of a dangling adenosine on reference $\vec{C}A\vec{U}G$ is especially striking when compared with the reference duplex containing an internal A·A non-bonded pair, $CA\vec{A}UG:CA\vec{A}UG$, which has a T_m <0°C (Romaniuk et al. 1979). The dangling adenosine stabilizes the duplex by as much as 12°C whereas the internal A·A pair destabilizes the duplex by at least 24°C. Even with duplex, $AG\vec{A}CU:AG\vec{A}CU$ (T_m = 26°C) the A·A pair disrupts the reference, $A\vec{G}CU$ (T_m = 34°C), by 9°C (Section 4.2). This was noteworthy in light of the fact that all duplexes contain the same Watson-Crick pairs and each single strand contains the same bases. Apparently, in the dangling base duplex the adenosines extend and strengthen the base stacking of the duplex whereas the internal non-bonded adenosines disrupt duplex formation.

Another valuable comparison was possible with the reference duplex containing an additional A·U pair, $CA\vec{U}UG:CA\vec{A}UG$ (T_m = 28.5°C) (Romaniuk et al. 1979) indicating that the increase in T_m resulting from an additional internal A·U base pair (cf. $\vec{C}A\vec{U}G$) was similar to the increase that results from having a dangling uridine at each terminus, but less than the result from a dangling adenosine. In view of the fact that base stacking should be similar in all single strands and that the duplex containing the additional internal A·U pair has five Watson-Crick pairs compared to four in the dangling base duplex, a satisfactory

TABLE 24

MELTING TEMPERATURES OF BASE PAIRED DUPLEXES

DUPLEX	AVERAGE T_m (°C)	INCREASE IN T_m OVER REFERENCE
CAUG	24	ref
ACAUG	35	11
UCAUG	30	5
CAUGA	36	12
CAUGU	30	5
CAAUG	< 0	
UGCA	33	ref
UGCAA	46	13
AGCU	34	ref
AGGUA	45	11
AAGCU	30	- 4
AAGCUA	49	15
AAGCUU	54	ref.
AAGCUUA	65	11

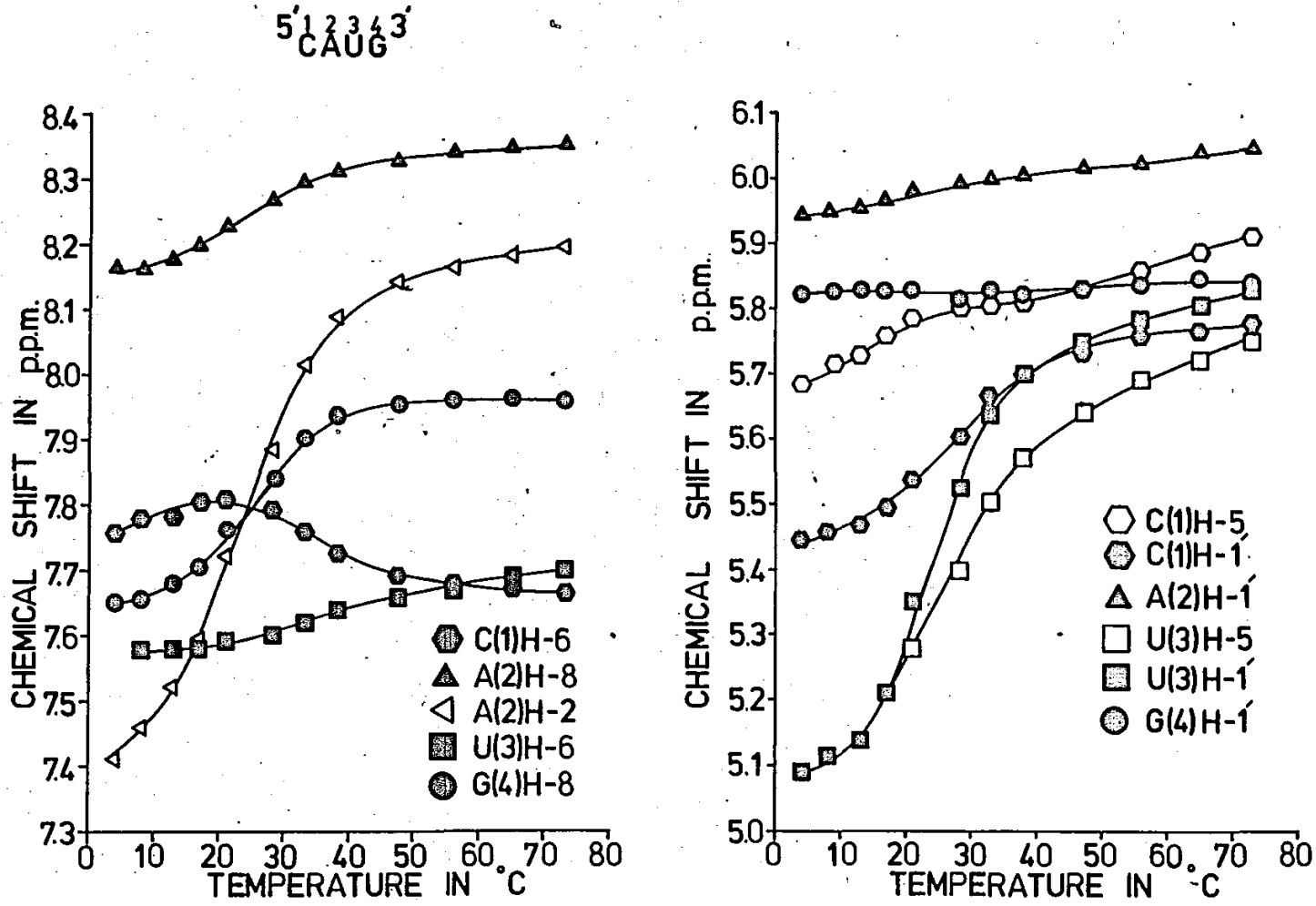


Figure 17. Results of the variable temperature study of the self-complementary duplex CAUG.

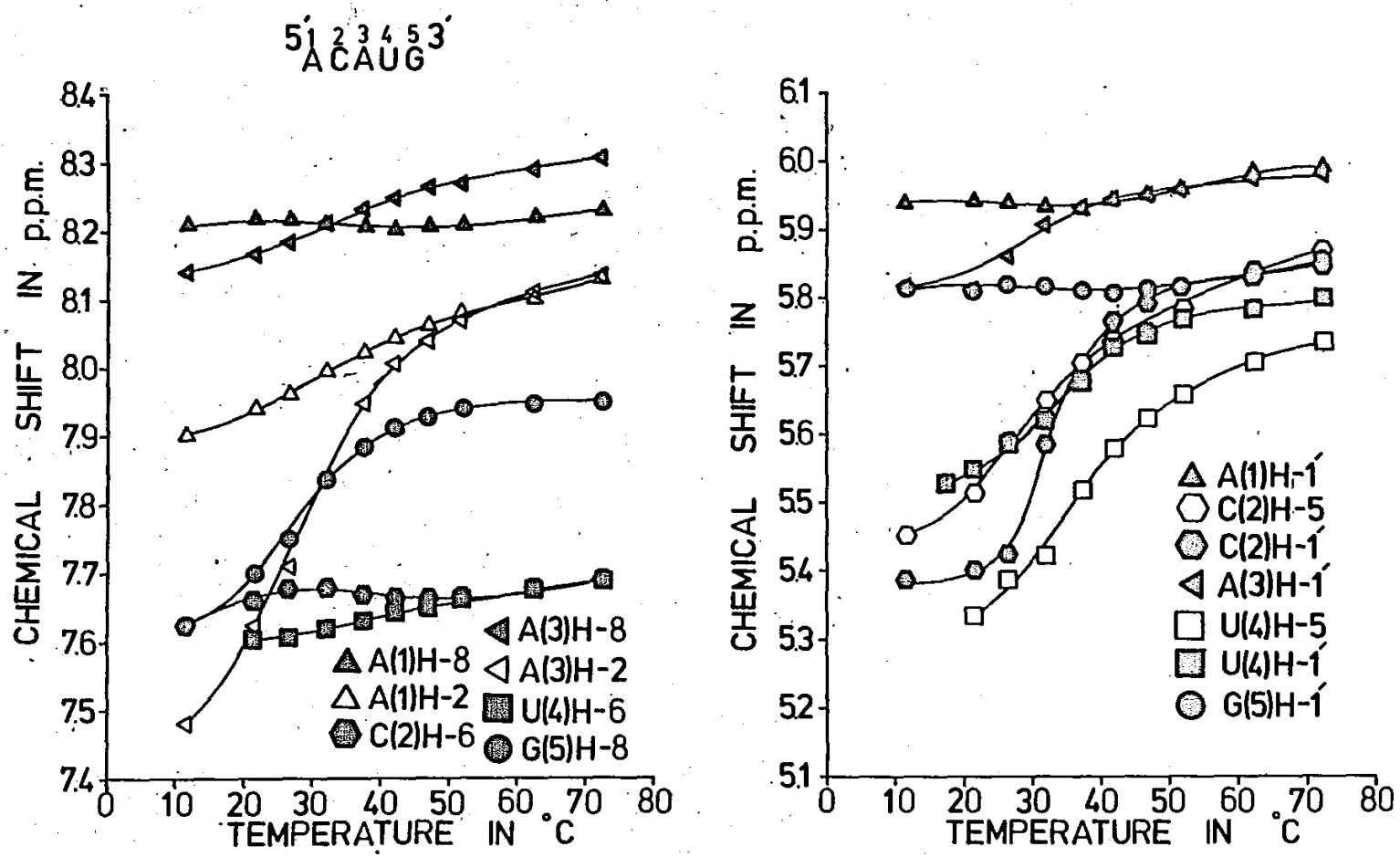


Figure 18. Results of the variable temperature study of ACAUG.

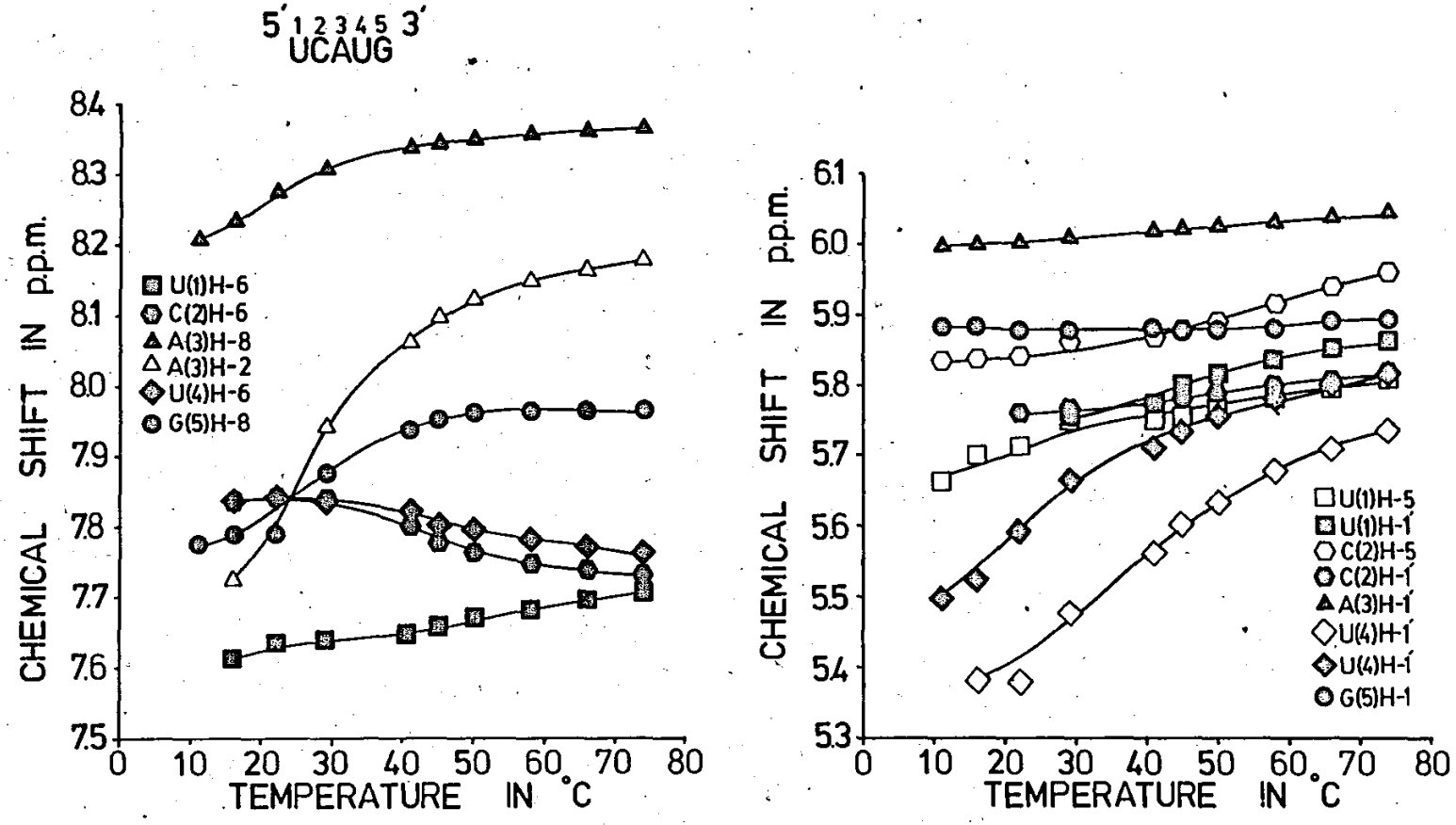


Figure 19. Results of the variable temperature study of UCAUG.

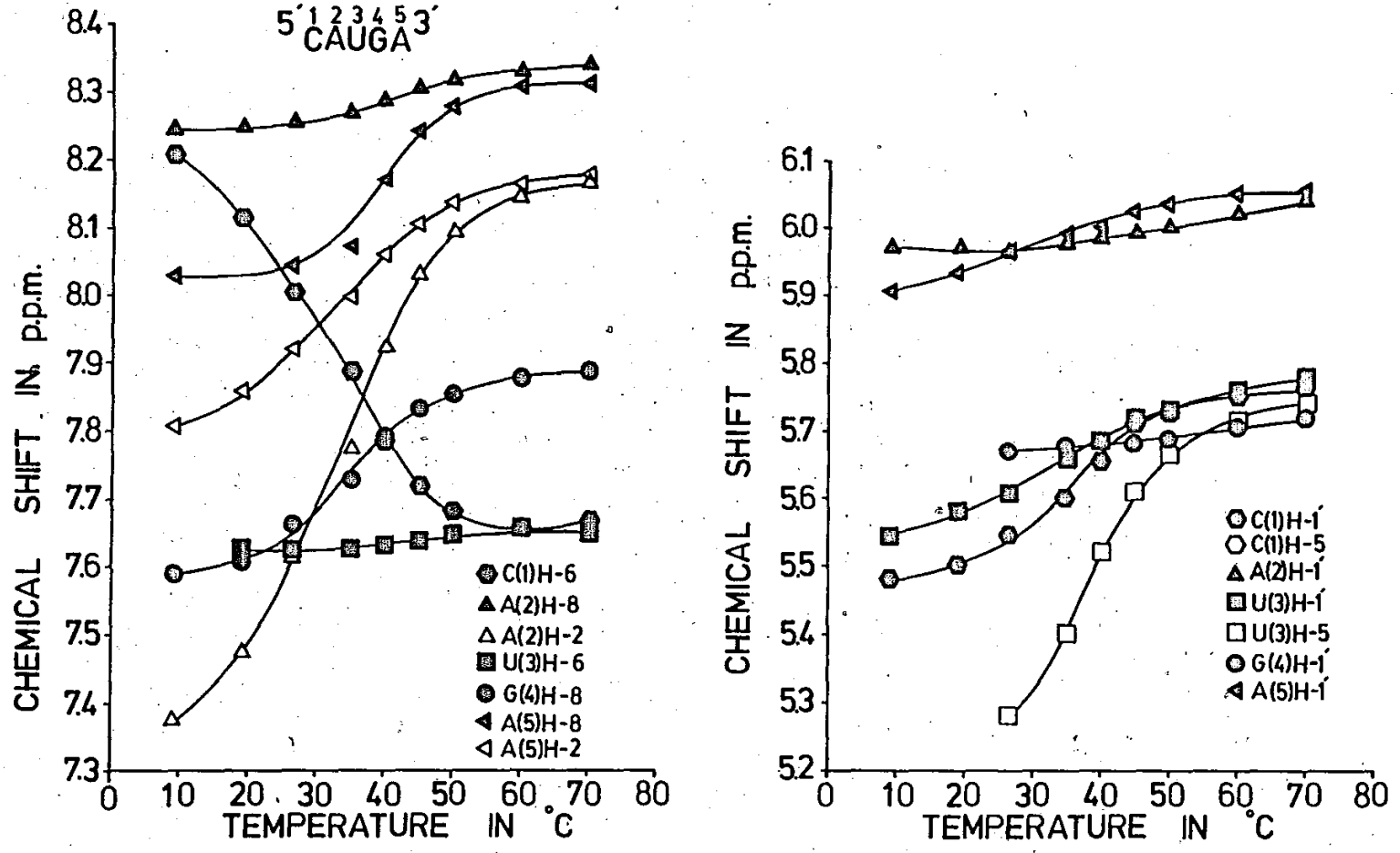
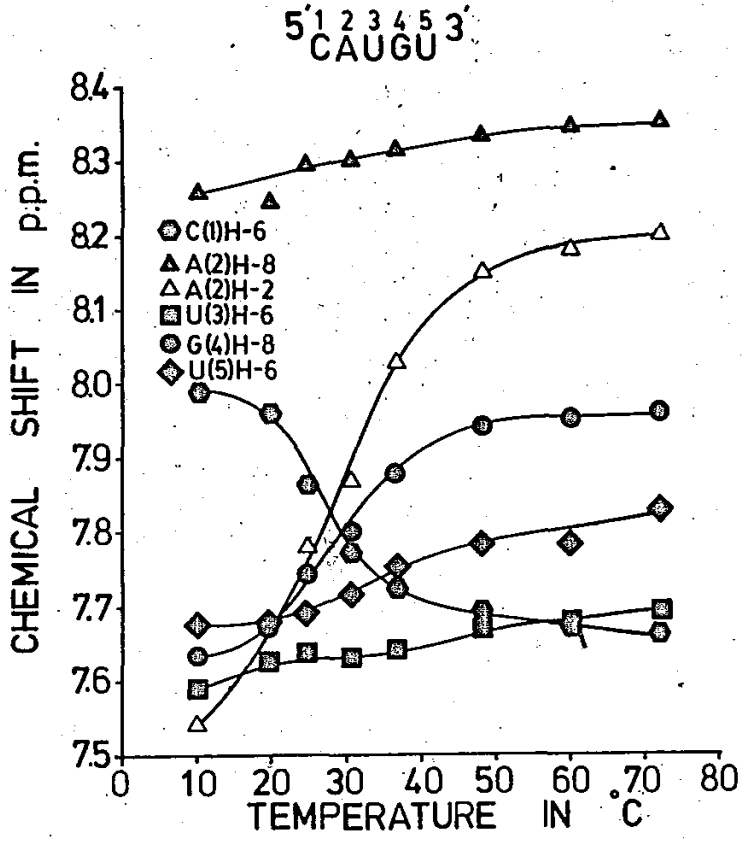


Figure 20. Variable temperature plots of CAUGA.



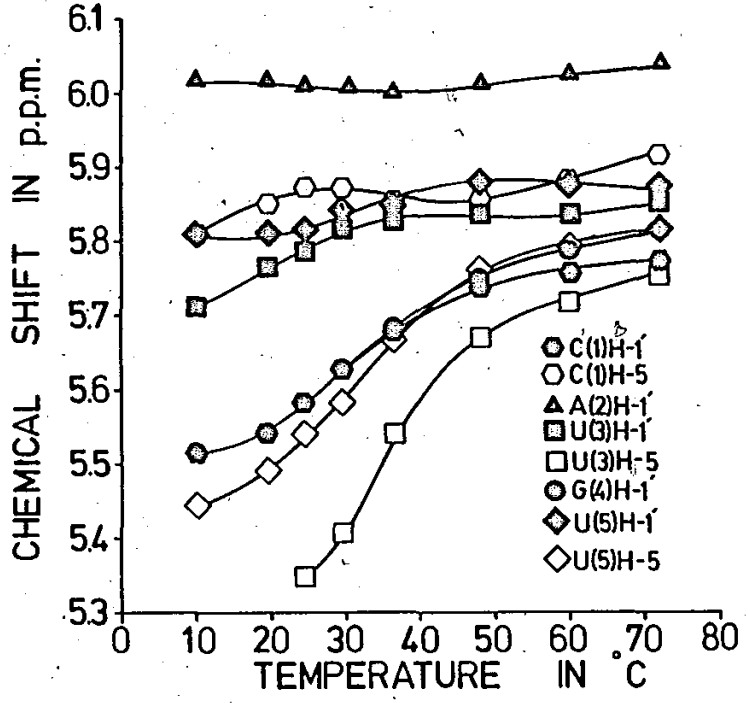


Figure 21. Variable temperature plots of CAUGU.

explanation for the magnitude of the dangling base contribution to duplex stability was not immediately apparent. However, if results for the insertion of an A·U base pair into tetramer duplex, $\vec{A}G\vec{C}U$ (T_m = 34°C), to give the duplex, $AG\vec{A}CU:AG\vec{U}CU$ (T_m = 46°C), were considered, it was noted that the additional A·U base pair increased the T_m by 12°C, similar to the effect of the dangling adenosine on $CA\vec{U}GA$. This implied that the three A·U base pair core of $CA\vec{A}UG:CA\vec{U}UG$ is actually less stable than expected. This may be due to the poor stacking of $\begin{matrix}\vec{A}\vec{A}\vec{U}\\ \dot{U}\dot{U}\dot{A}\end{matrix}$ sequences as suggested in Section 4.2.

This study demonstrated that dangling bases can add stability to sequences with non-fraying ends. Therefore, increases in the stability of duplexes with dangling residues was not merely a result of a reduction in terminal base pair fraying, but rather was due to an increase in the extent and strength of base stacking.

Stabilities of duplexes formed by $U\vec{G}CA$ and $UG\vec{C}AA$ were next considered (Table 24; Figures 22 and 23). The duplex $U\vec{G}CA$ (T_m = 33°C) was stabilized by 13°C when the 3'-dangling adenosine was added to give duplex $UG\vec{C}AA$ (T_m = 46°C) and was in agreement with the findings for $C\vec{A}UG$.

Although A·U base pairs occupy the terminal positions of this duplex, $U\vec{G}CA$, no fraying was observed (Figure 22). This result appeared to contradict a previous study with duplex $AA\vec{G}CUU$ (Kan et al. 1975; Borer et al. 1975), where terminal A·U base pairs appear to fray, i.e. the average T_m from the terminal A·U base pair plots were significantly lower than the T_ms for those of the internal G·C and A·U base pairs (Table 26). The fact that $U\vec{G}CA$ gave no indication of fraying suggests that fraying may be a sequence-dependent phenomenon, a result supported

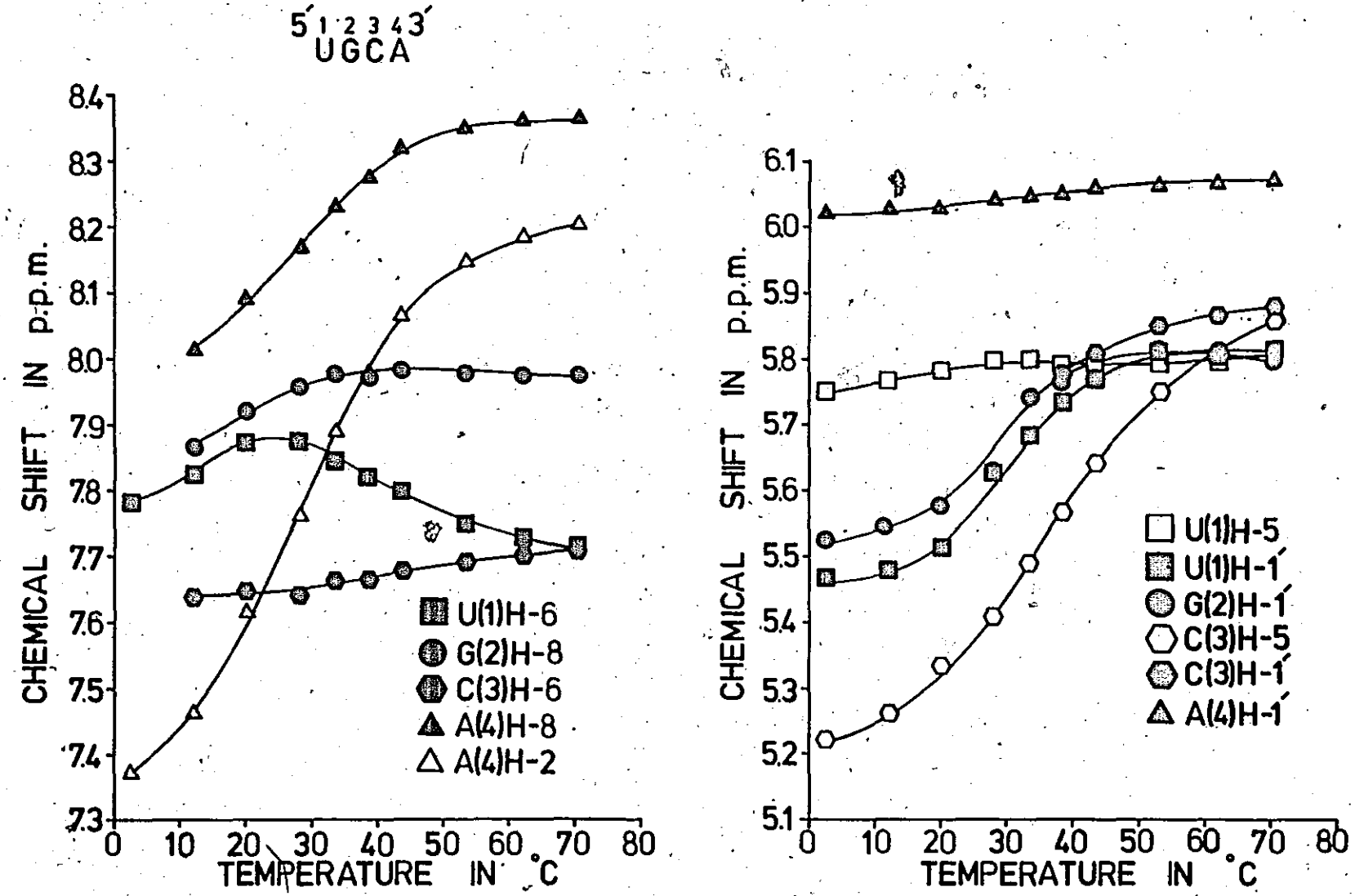


Figure 22. Results of the variable temperature study of the self-complementary duplex UGCA.

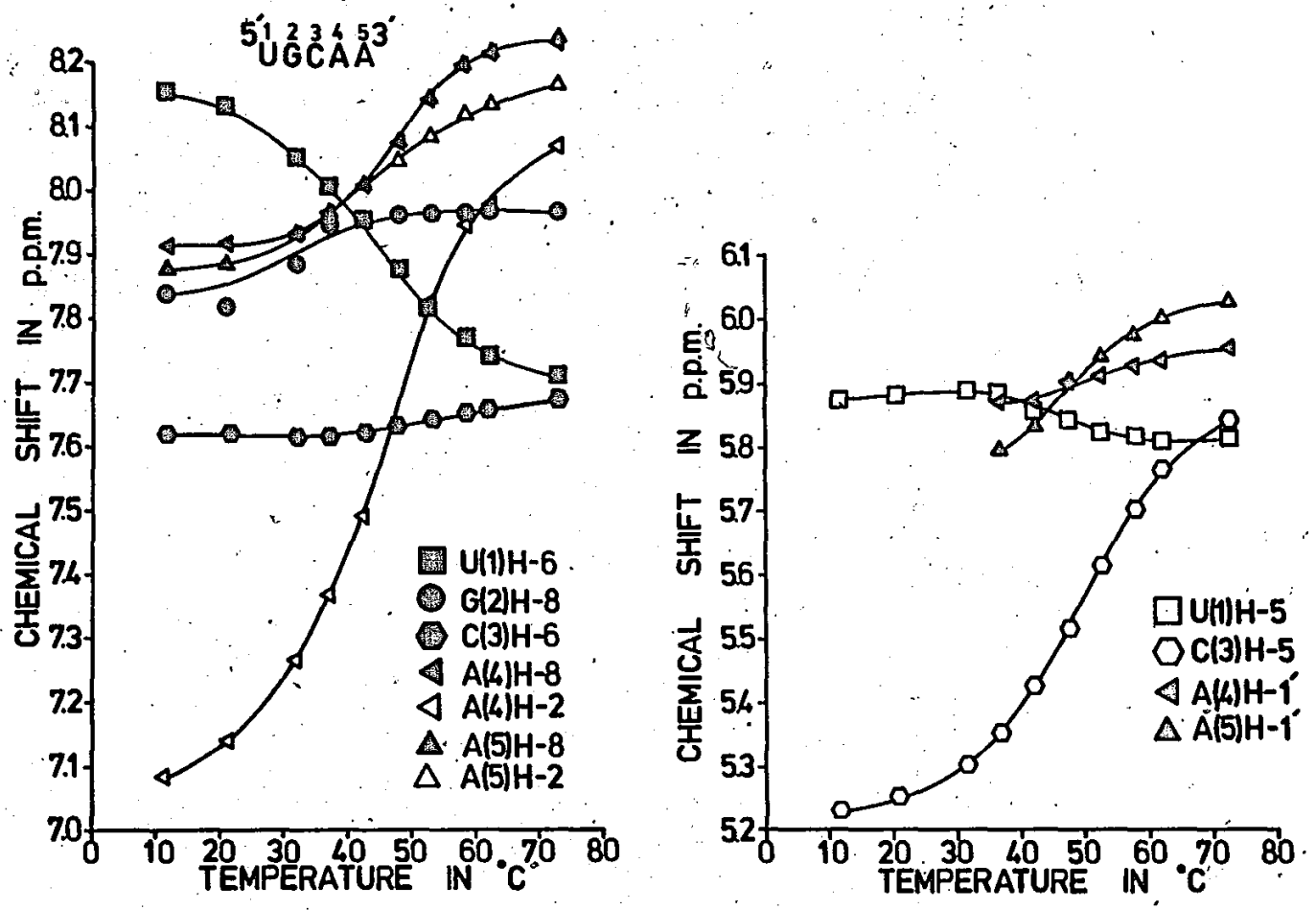


Figure 23. Results of the variable temperature study of UGCAA.

by work with the duplex formed by $\vec{A}GGA:U\vec{C}CU$. This duplex also has two terminal A·U and two internal G·C base pairs, but there was no significant difference in the T_ms of the various sigmoidal curves, indicating fraying was absent (Romaniuk, 1979).

A study of the stabilities for the $\vec{A}GCU$ series was significant in that the T_ms observed for the individual curves for $\vec{A}GCU$ (Figure 24) revealed differences between the terminal A·U and internal G·C base pairs (Table 25) (Romaniuk, 1979; Neilson et al. 1980), possibly indicating fraying. These results were in agreement with those for the earlier $AA\vec{G}CUU$ study (Borer et al. 1975; Kan et al. 1975).

Two tetramers, $\vec{A}GCU$ and $U\vec{G}CA$, each containing (1) an equal number of A·U and G·C base pairs and (2) terminal A·U base pairs, were now shown to possess different terminal base pair stabilities. Primary sequence would appear to play an important role in this phenomenon. Since all strands of both duplexes were completely stacked (all $J_{1',2'}$ <1.0 Hz) the less stable A·U pairs must possess a greater tendency to unwind in the $\vec{A}GCU$ duplex. Significantly, addition of adenosine to give duplex $AG\vec{C}UA$ (Figure 25) increased the T_m by 11°C to 45°C from 34°C for $\vec{A}GCU$. This increase was similar to that seen for $C\vec{A}UG$. Closer inspection of the T_ms for the individual base pairs (Table 25) revealed that the degree of instability was greater in the parent $\vec{A}GCU$ duplex ($\Delta T_m = T_m$ (G·C)$-T_m$(A·U)=5°C than in the $AG\vec{C}UA$ duplex ($T_m \simeq 2$°C), an observation that supports an earlier hypothesis that dangling bases could reduce fraying (Kallenbach & Berman, 1977). The extension of the base stacking in $AG\vec{C}UA$, resulting from the addition of the dangling adenosine, strengthens terminal A·U base pairs, so increasing their

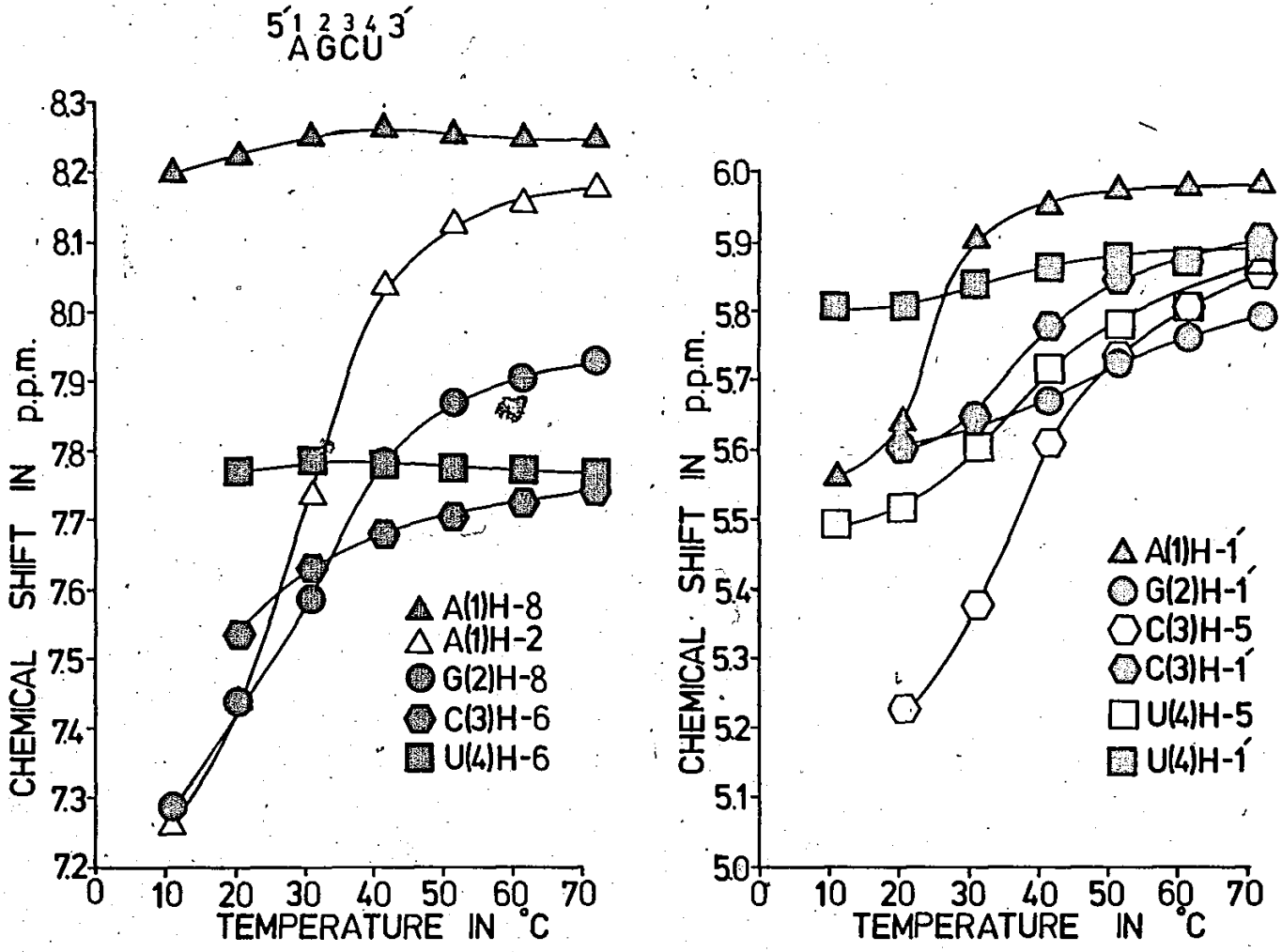


Figure 24. Variable temperature plots for the aromatic base and anomeric resonances for the self-complementary duplex AGCU.

TABLE 25

DETERMINATION OF MELTING TEMPERATURES

OF AGCU AND AGCUA

Resonance	T_m (°C)	Average T_m/ Base Pair	Average T_m for Duplex
(1) AGCU			
A(1)H-2	31		
A(1)H-1'	26		
U(4)H-1'	34	31	
U(4)H-5	34		34
G(2)H-8	33		
C(3)H-6	38		
C(3)H-5	34	36	
G(2)H-1'	43		
C(3)H-1'	33		
(2) AGCUA			
A(5)H-8	44		
A(1)H-8	44		
A(5)H-2	44		
A(1)H-2	43	44	
A(5)H-1'	44		
U(4)H-1'	43		
U(4)H-5	47		45
G(2)H-8	45		
C(3)H-6	45		
C(3)H-1'	45	46	
C(3)H-5	49		
G(2)H-1'	48		

individual T_m's relative to the internal G·C pairs. Examination of the variable temperature plots for $\vec{A}GCU$ and $AG\vec{C}UA$ (Figures 24 and 25) duplexes reveals T_m values cannot be explained by end-to-end duplex aggregation since all melting curves exhibit sigmoidal behaviour and a decrease in temperature dependence of shielding at lower temperatures (Patel, 1970).

In the $\vec{A}GCU$ duplex, all H-1' doublets were reduced to broad singlets ($J_{1',2'}$ <1.0 Hz) at temperatures (<30°C). This result was consistent with a high degree of base-stacking (Altona, 1975) necessary for duplex formation. The A(5)H-1' doublet of $AG\vec{C}UA$, however, remained well resolved ($J_{1',2'}$ = 2.4 Hz at 20°C) at temperatures below the duplex T_m, while the other anomeric signals were already singlets at 40°C. This observation implied that the dangling adenosine was less stacked relative to the residues involved in base pairing and so enjoyed more rotational freedom reminiscent of the orientation of the $C\vec{C}A$ arm in tRNAs (Rich, 1977). All bases in the $C\vec{A}UG$ series duplexes were highly stacked, including the dangling bases, consequently such rotational freedom of the dangling A in the $AG\vec{C}UA$ duplex is consistent with non-stable terminal A·U base pairs in the $\vec{A}GCU$ series.

The results in this study were considered open to question because of the relatively small difference seen in the T_ms for the terminal A·U and internal G·C pairs in $\vec{A}GCU$. Borer (1975), in his studies, found much larger differences in the T_ms for $AA\vec{G}CUU$. It was decided to repeat the experiments with $\vec{A}AGCUU$ to determine the extent of fraying, if any, then add an additional 3'- dangling adenosine and determine if any difference in fraying or duplex stability occurred (Figures 26 and 27).

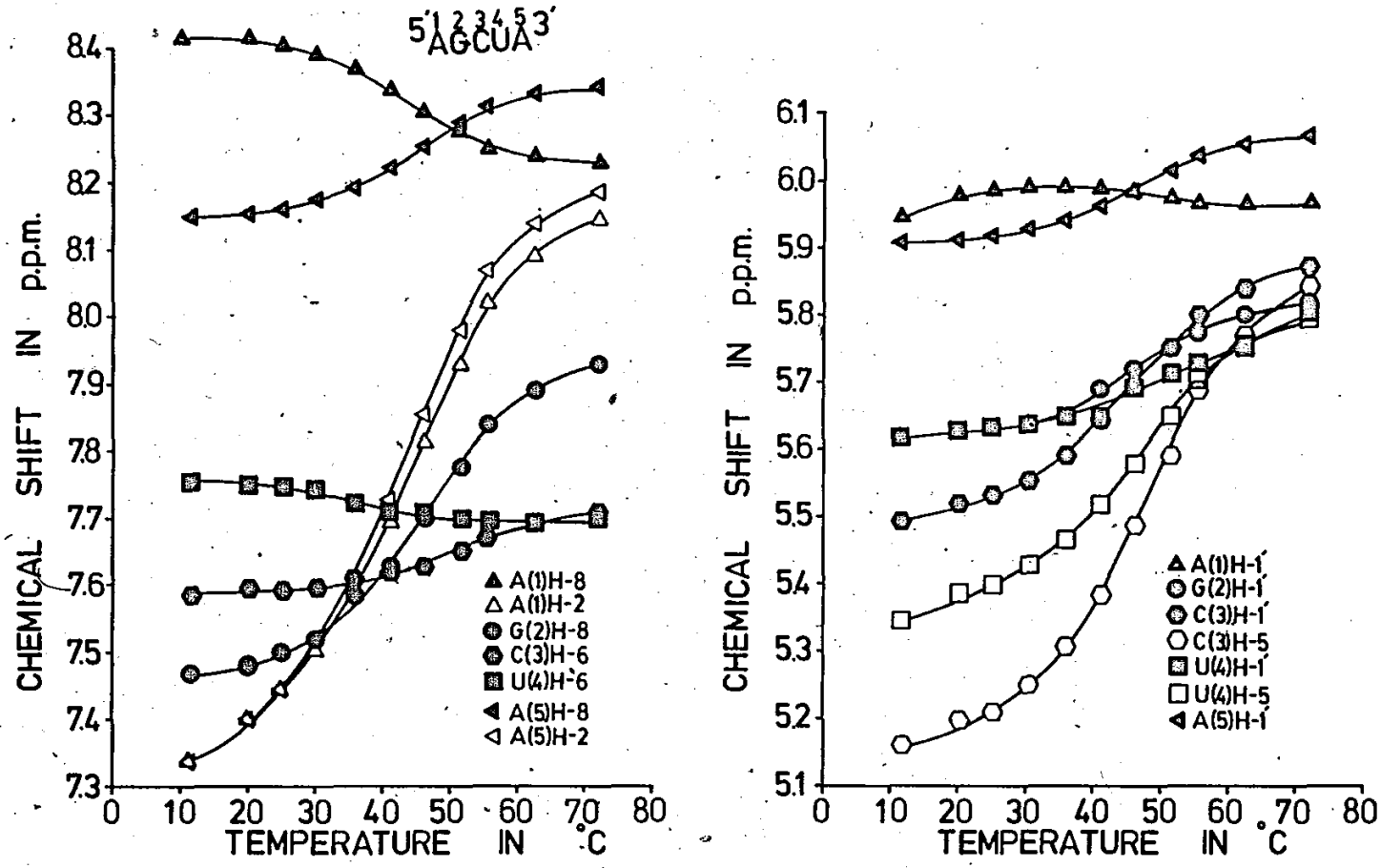


Figure 25. Variable temperature study of AGCUA.

Results for AAGCUU and AAGCUUA are presented in Table 26 and deviations from Borer's results are immediately obvious. Borer's results indicate that the terminal A·U and internal A·U base pairs melt well before the G·C core. On the other hand, studies with AAGCUU in this laboratory showed no significant differences in the T_ms of the individual base pairs. These results were difficult to explain since the assignments at high temperature agree with those of Borer. Also, the temperature versus chemical shift curves showed behaviour similar to those presented by Borer (Figures 26 and 27). However, Borer was able to follow the chemical shifts to far lower temperatures and, in addition, could back up his results with observations of NH resonances which are directly involved with H-bond formation in base pairing. Despite the differences in the T_ms for the individual base pairs, the average T_m for each duplex was very similar. The differences between the two studies cannot be satisfactorily explained without much additional work, including study of the NH resonances. However, the results obtained with the rest of the study are significant. When a dangling adenosine was added to AAGCUU to give the duplex AAGCUUA the T_m increased to 65°C. Once again the dangling adenosine stabilizes the duplex by 11°C in agreement with the previous results. Therefore, whether AAGCUU frays or not, the dangling adenosine increases duplex stability.

From Tables 24-26 a comparison of the increase in T_m values reveals that whether the sequence was suspected of fraying or not had no effect on the increase in duplex stability from a dangling adenosine. In each case a dangling adenosine stabilizes the duplex by 11°-13°C. A dangling base was thought to stabilize non-fraying duplexes simply by

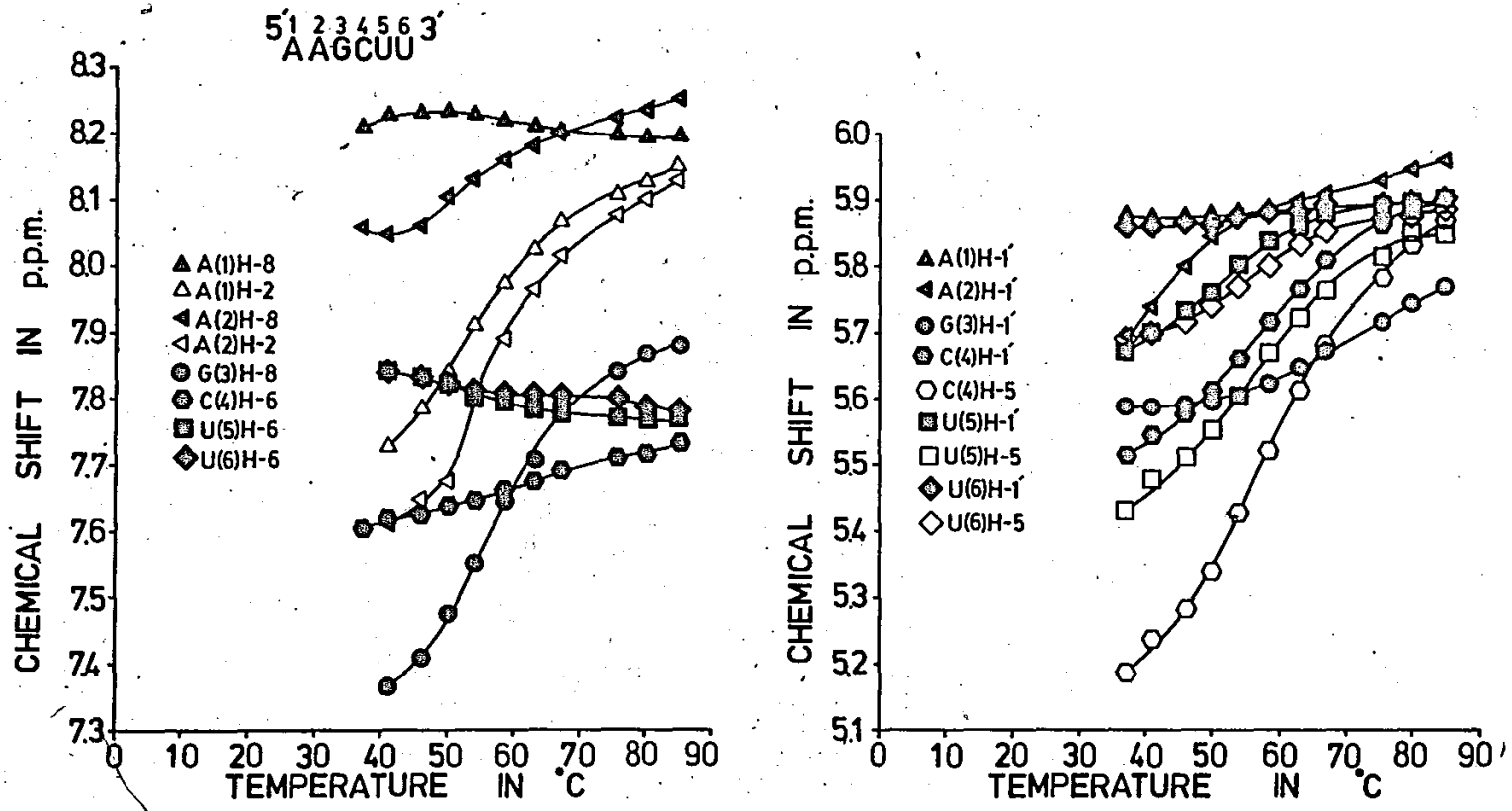


Figure 26. Variable temperature study of AAGCUU.

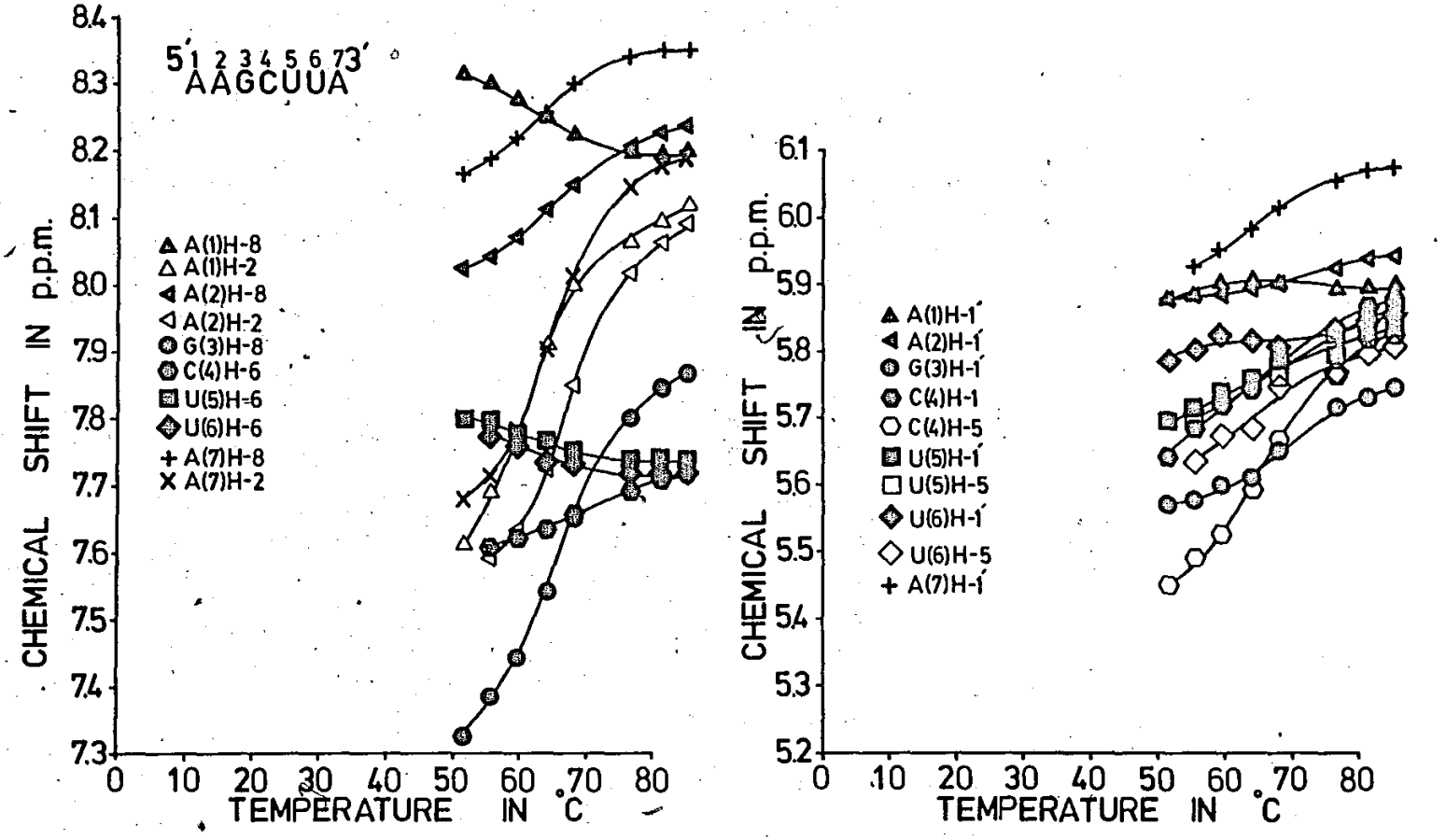


Figure 27. Variable temperature plots of the aromatic base and anomeric resonances of AAGCUUA.

TABLE 26

COMPARISON OF THE T_ms OBTAINED FOR THE A·U AND G·C BASE PAIRS IN AAGCUU SEQUENCE (T_m in °C)

Base Pair	AAGCUU (Alkema, 1981)	AAGCUU (Borer et al. 1975)	AAGCUUA
Terminal A·U	55	47	64
Internal A·U	52	48	66
Core G·C	55	57	67
Average T_m	54	53	65

extending and strengthening base stacking throughout the sequence and thus stabilizing the base pairs to a greater degree. On the other hand, with fraying duplexes a dangling base was thought to reduce the fraying of terminal base pairs (Kallenbach & Berman, 1977). However, it would still increase base stacking as well and, therefore, adenosine contributes in two ways to the stability of duplexes that fray. Accordingly, duplexes that have a tendency to fray should exhibit greater changes in T_m than duplexes that do not fray when a dangling adenosine is added. $\vec{A}GCU$ and $U\vec{G}CA$ have similar T_ms; but $\vec{A}GCU$ is thought to fray, while $U\vec{G}CA$ definitely does not. Also, when a dangling adenosine was added, both duplexes were stabilized to similar extents and in fact $U\vec{G}CA$ shows a greater increase in T_m. Similarly, $AA\vec{G}CUUA$ only increased by 11°C over $AA\vec{G}CUU$. $C\vec{A}UG$, which does not fray, was also stabilized by 11°-12°C when a 5'- or 3'-dangling adenosine was added. These results suggest that a dangling adenosine stabilizes the duplex only by extension and strenghtening of base stacking throughout the duplex and has little or no effect on fraying. Alternatively, it suggests that the fraying bases have little effect on duplex stability which is determined by the G·C core.

Tetramer, $\vec{A}GCU$, also served as reference duplex for the effects of a 5'-dangling adenosine in duplex $AA\vec{G}CU$ (Figure 28) and for the effects of a 5'- and 3'-dangling adenosine on the same sequence in the duplex $AA\vec{G}CUA$ (Figure 29). Table 24 presents the results obtained for these studies. The presence of a 5'-dangling adenosine on $AA\vec{G}CU$ actually destabilized the duplex (T_m = 30°C *vs.* 34°C). This dangling base did not appear to extend base stacking and in fact seemed to disrupt base

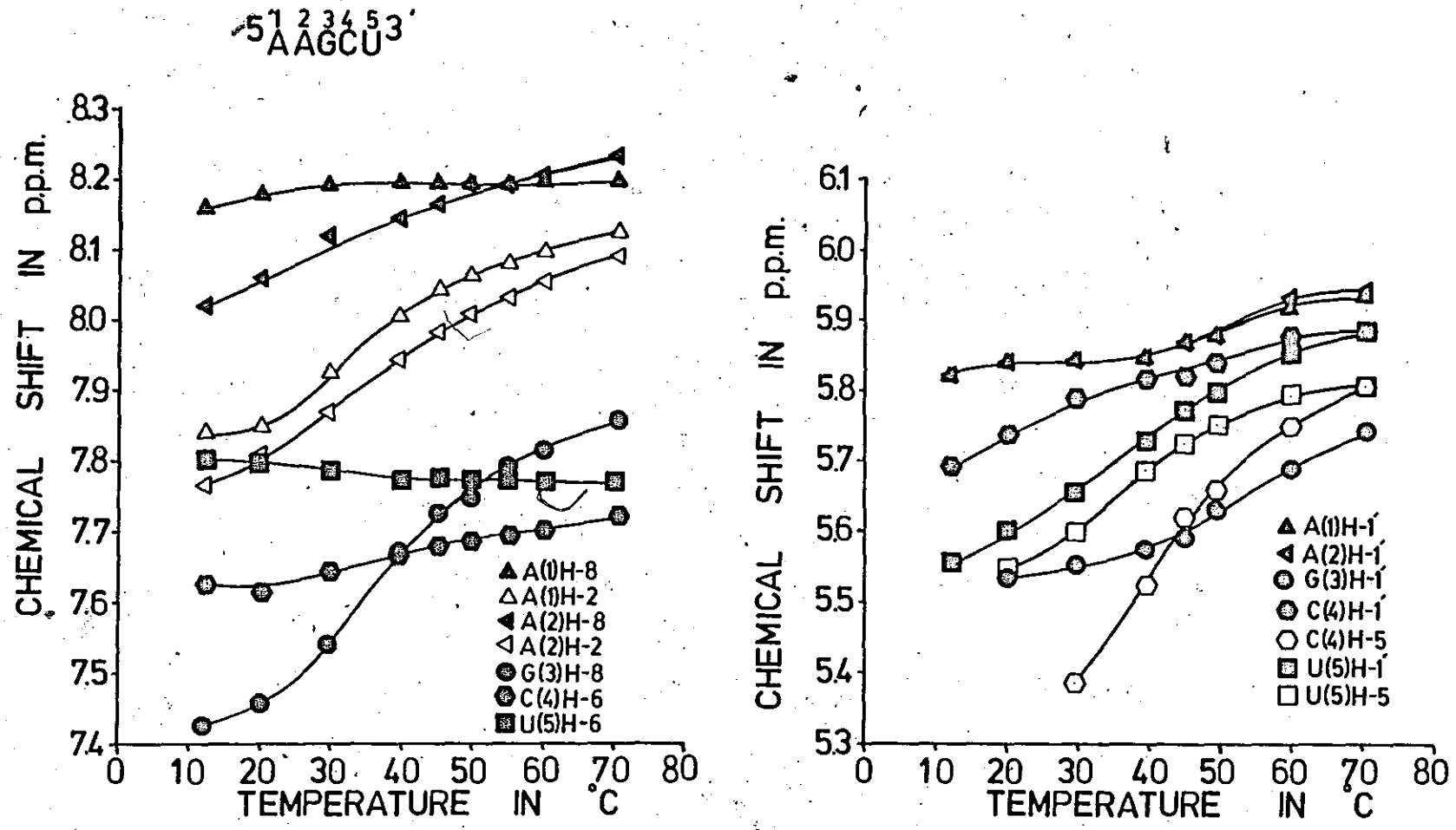


Figure 28. Variable temperature results for the duplex AAGCU.

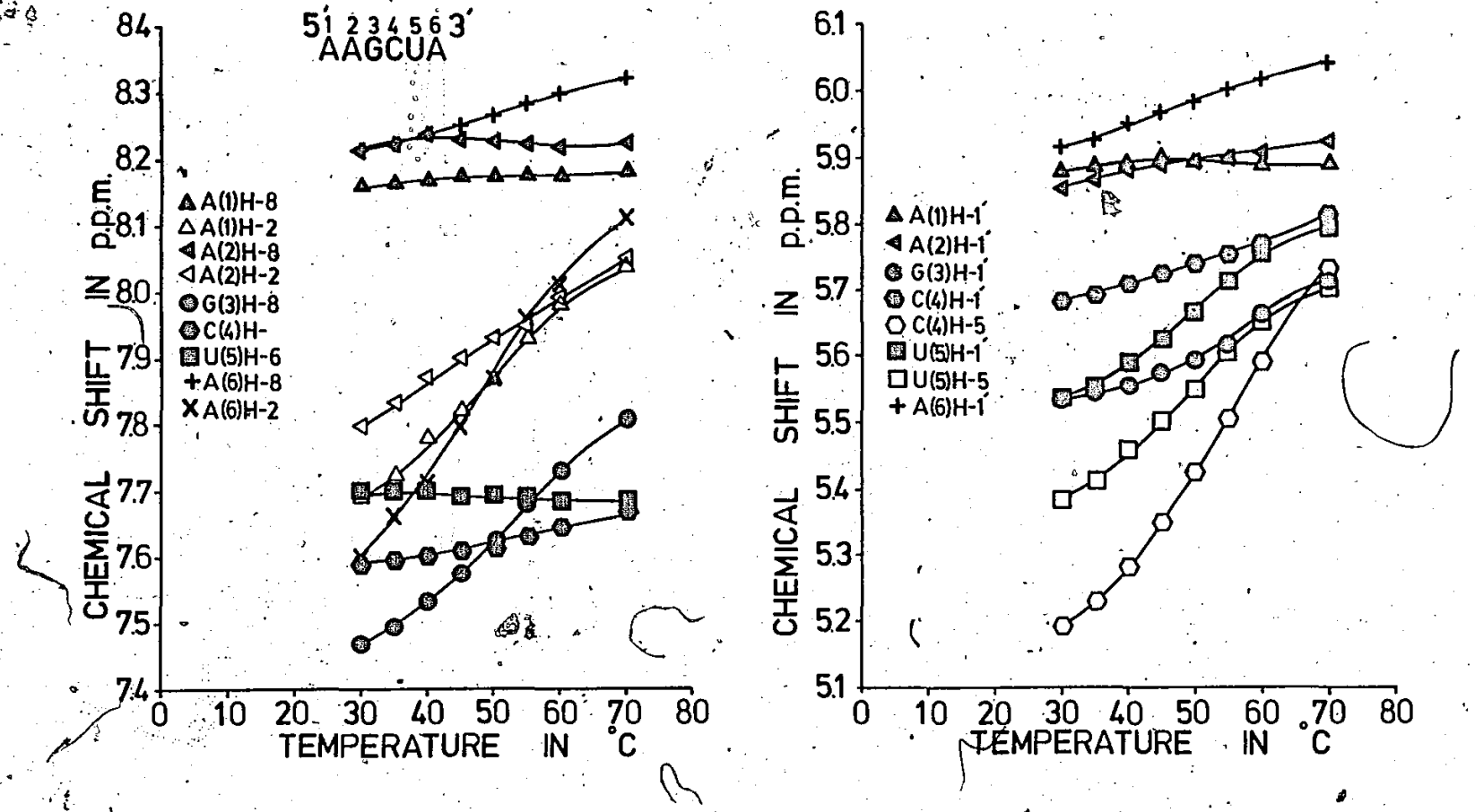


Figure 29. Variable temperature results for the duplex AAGCUA.

pairing. This result is in direct contrast to the studies of 5'-dangling bases on CÂUG. An explanation for this interesting finding was not available until after trimer GĈA was studied (Section 4.4).

Comparison of the duplex from AAĜCUA, which has two terminal A·A non-base-paired pairs, with its parent tetramer duplex AĜCU was significant (T_ms 49°C vs. 34°C). A·A pairs are centres of instability within duplexes; for example, duplex CAÂUG:CAÂUG cannot be detected (Romaniuk et al. 1979) and duplex AGÂCU:AGÂCU (T_m = 26°C) was significantly less stable than the parent AĜCU. However, in the duplex with terminal A·A pairs, sufficient geometric flexibility must exist to tolerate the adenosine residues in opposing positions and still maintain extension of base stacking to these same adenosine residues. A partially "unwound" hexamer duplex appeared to be more stable than a "tight" tetramer duplex with a similar Watson-Crick base-paired core. This finding has direct bearing on the junction of duplex and non-duplex regions within native RNA molecules. Purine residues in the immediate non-paired positions may contribute to overall stability despite their apparent steric interaction.

From the results obtained it is clear that dangling bases contribute significantly to the overall duplex stability. The studies presented here serve as models for duplex regions that have loops adjacent to them. The results support earlier theories that loop regions stabilize duplex regions (Grosjean et al. 1978) but that stability was dependent upon the nature (i.e. sequence) of base residues within the dangling sequence. The results also question the importance of base fraying and the effect that dangling bases have on fraying.

4.4. Triplet GpCpA Forms a Stable RNA Duplex (Alkema et al. 1981)

Perfect RNA duplexes containing three Watson-Crick base pairs are unstable under physiological conditions (Jaskunas et al. 1968; Brahms et al. 1969; Borer et al. 1974). Triribonucleotides, however, have been found to form stable duplexes if there are adjacent single stranded helical regions, such as in tRNA loops, where the bases already stacked in the helix account for the enhanced stability (Uhlenbeck et al. 1970; Grosjean et al. 1978). An alternate means of increasing base stacking, and so strengthening the overall duplex, is the presence of dangling bases (Martin et al. 1971; Neilson et al. 1980; Section 4.3).

In the constant search for factors that affect duplex stability and in the ongoing attempt to synthesize, analyze and characterize all 64 possible trimers, the first triribonucleotide to form a simple stable duplex, $\vec{GCA}$:$\vec{GCA}$ was discovered. Duplex $\vec{GCA}$:$\vec{GCA}$ contains two G·C Watson-Crick base pairs and two 3'-dangling adenosines. Remarkably, this duplex had a stability similar to that of the corresponding self-complementary tetramer duplex, formed from $\vec{UGCA}$, which contains four Watson-Crick pairs. The trimer, therefore, must derive its stability over the dinucleotide duplex $\vec{GC}$:$\vec{GC}$ by virtue of its 3'-dangling adenosine residues whose contributions to duplex stability approximate those of A·U pairs.

4.4.1. Results

Oligoribonucleotides, $\vec{GC}$, $\vec{GCA}$, $\vec{GCAA}$ and $\vec{AGC}$, were synthesized using the phosphotriester method. Variable temperature ^{1}H-NMR spectroscopy was used to monitor duplex stability. The chemical shifts were routinely assigned by comparison to UGCA and the chemical shift vs.

TABLE 27

NMR CHEMICAL SHIFT ASSIGNMENTS FOR GCA (7.3 mM)

OVER THE TEMPERATURE RANGE 70°C-0°C

	70.6°	59.6°	47.9°	38.5°	27.1°	20.0°	12.0°	0.6°	T_m °
A(3)H-8	8.348	8.340	8.310	8.264	8.191	8.156	8.120	8.085	30.5
A(3)H-2	8.191	8.153	8.069	7.955	7.790	7.703	7.624	7.564	32
G(1)H-8	7.920	7.925	7.941	7.969	8.009	8.031	8.045	8.053	32
C(2)H-6	7.709	7.699	7.685	7.669	7.647	7.642	7.641	7.634	NSB
A(3)H-1'	6.061	6.057	6.046	6.036	6.027	6.022	6.017	6.010	NSB
C(2)H-1'	5.883	5.850	5.808	5.744	5.652	5.600	5.554	5.516	32
G(1)H-1'	5.812	5.807	5.800	5.801	5.811	5.817	5.812	5.798	NSB
C(2)H-5	5.837	5.774	5.659	5.527	5.362	5.287	5.216	-	39

Average T_m = 33

NSB : No sigmoidal behaviour

temperature changes for the aromatic and ribose H-1' protons of GCA are shown in Figure 30 and listed in Table 27. The averaged T_m for the sigmoidal plots of these protons is 33°C at 7.3 mM. The plots displayed in Figure 30 are only consistent with a GCA duplex containing two Watson-Crick base pairs. The chemical shift of CH-5 changed by 0.615 ppm to higher field over the temperature range 70°-10°C, and this upfield movement is characteristic of a CH-5 on a cytidine involved in normal G·C Watson-Crick base pairing as shown by the 0.559 ppm upfield shift for the CH-5 in the $\vec{U}GCA$ duplex (see Table 28 and Figure 22 in Section 4.3.2.). Protons, AH-8 and AH-2, on the dangling adenosines also exhibit pronounced upfield chemical shift changes during $G\vec{C}A$ duplex formation, indicating their involvement in duplex formation. In addition, the $J_{1',2'}$ coupling constants for the ribose H-1' protons of the guanosine and cytidine residues collapse to <0.5 Hz below 30°C while the $J_{1',2'}$ values for the adenosine decrease but do not become <0.5 Hz until temperature approaches 0°C. This was indicative of strong GC stacking, while the 3'- adenosine unit still retained some flexibility in the duplex (Cheng et al. 1980).

4.4.2. Discussion

Although the trimer, $G\vec{C}A$, contains a purine-pyrimidine-purine sequence, these results provided an interesting contrast to those obtained from studies with similar type base sequences which preferred internal bulge base conformations (Lee & Tinoco, 1980) at lower temperatures.

The spectacular stability of the $G\vec{C}A$ duplex , with its 3'-dangling adenosine, was even more dramatic when compared to the

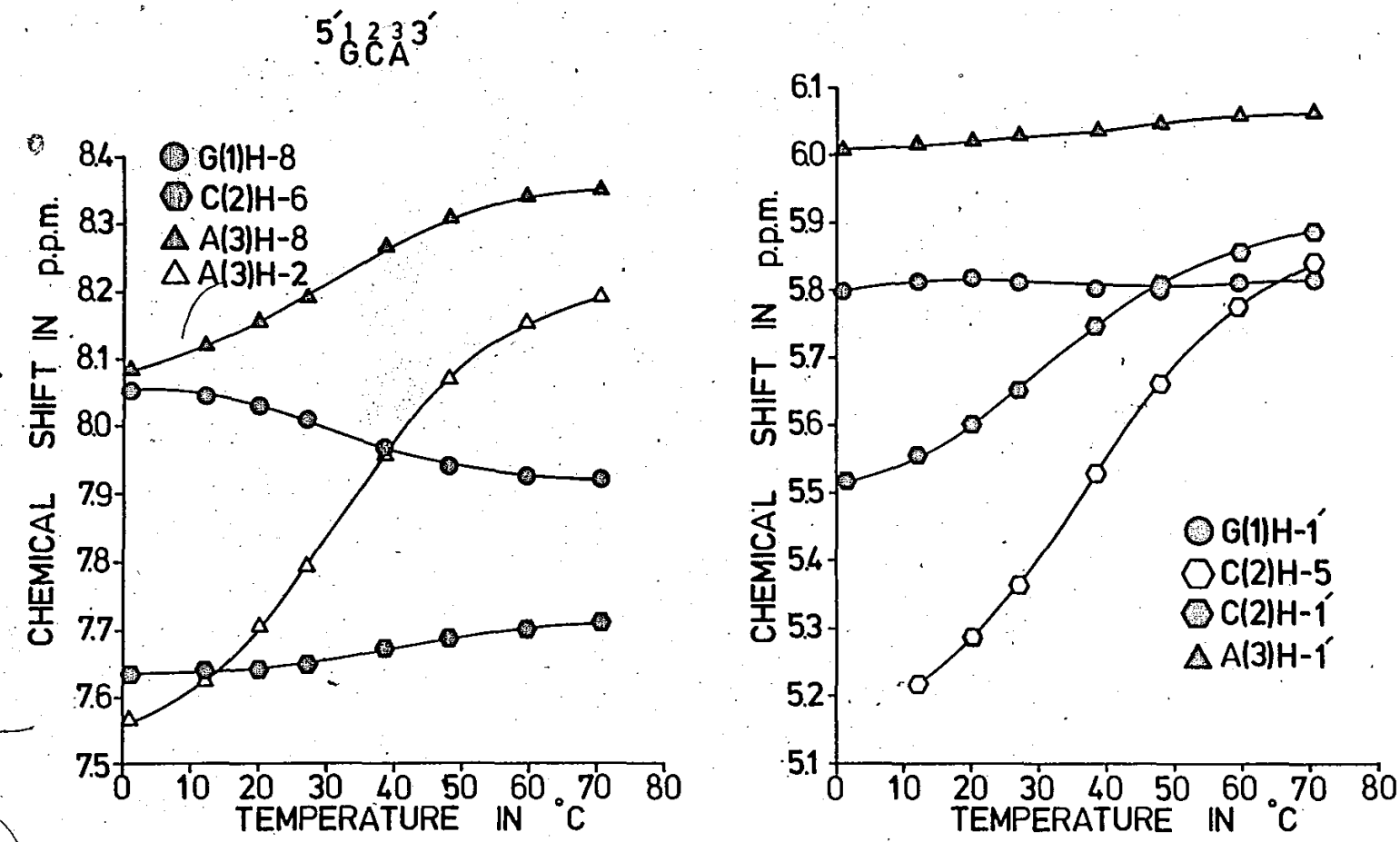


Figure 30. Variable temperature plots of the aromatic base and anomeric resonances of duplex GCA.

TABLE 28

NMR CHEMICAL SHIFT ASSIGNMENTS FOR UGCA (8.2 mM)

OVER THE TEMPERATURE RANGE 70°C-0°C

Resonance	70.6°	62.0°	53.3°	43.4°	38.5°	33.4°	27.9°	19.8°	12.1°	2.5°	T_m °
AH-8	8.362	8.359	8.349	8.320	8.274	8.232	8.167	8.092	8.014	7.977	30.5
AH-2	8.202	8.183	8.147	8.066	7.978	7.887	7.760	7.616	7.463	7.369	30.5
GH-8	7.975	7.975	7.978	7.984	7.978	7.978	7.958	7.919	7.864		40.5
UH-6	7.716	7.727	7.747	7.793	7.819	7.845	7.874	7.874	7.825	7.781	NSB
CH-6	7.711	7.700	7.692	7.675	7.664	7.661	7.642	7.646	7.638		NSB
AH-1'	6.068	6.063	6.062	6.057	6.049	6.045	6.039	6.027	6.027	6.021	32.0
CH-1'	5.875	5.865	5.845	5.808	5.777	5.812	5.812	5.819	5.806		NSB
CH-5	5.853	5.806	5.749	5.640	5.565	5.488	5.407	5.335	5.259	5.221	38.0
UH-5	5.811	5.798	5.790	5.793	5.791	5.798	5.798	5.782	5.764	5.749	NSB
UH-1'	5.811	5.803	5.811	5.769	5.733	5.682	5.627	5.511	5.477	5.467	30.5
GH-1'	5.794	5.812	5.811	5.769	5.770	5.739	5.627	5.526	5.646	5.524	32.0

Average T_m = 33

NSB: No sigmoidal behaviour

TABLE 29

NMR CHEMICAL SHIFT ASSIGNMENTS FOR AGC (16 mM)

OVER THE TEMPERATURE RANGE 70°C-0°C

	72.1°	60.6°	50.0°	37.7°	31.7°	20.6°	9.6°	0.4°
AH-8	8.238	8.235	8.235	8.227	8.224	8.213	8.191	8.191
AH-2	8.170	8.143	8.116	8.072	8.050	8.004	7.945	7.909
GH-8	7.934	7.909	7.882	7.842	7.820	7.768	7.687	7.627
CH-6	7.748	7.737	7.725	7.707	7.698	7.674	7.643	7.627
AH-1'	5.969	5.960	5.953	5.935	5.926	5.899	5.846	5.821
CH-5	5.872	5.834	5.788	5.724	5.698	5.639	5.533	5.493
CH-1'	5.889	5.866	5.854	5.834	5.826	5.802	5.767	5.740
GH-1'	5.809	5.778	5.754	5.702	5.678	5.630	5.566	5.526

trinucleotide, AGC, which contains a 5'-dangling adenosine. Chemical shift *vs.* temperature changes for the aromatic and ribose H-1' protons of $\vec{A}GC$ at 16.0 mM are shown in Table 29 and Figure 31. The small chemical shift changes observed are typical for all reported trinucleotides (Lee & Tinoco, 1980; Everett *et al.* 1980) and implies that for the normal molar concentration range (1-10 mM) the averaged T_m value would be <0°C. The behaviour of the CH-5 proton of $\vec{A}GC$ is also indicative of there being no significant interaction of cytidine and guanosine in the normal Watson-Crick manner, at temperatures down to 0°C. Observations that 3'-dangling residues contribute more to duplex stability than corresponding 5'-dangling residues have been previously reported (Gennis & Cantor, 1970; Section 4.3.2.). The greater helical overlap of a 3'-dangling base residue generates increased aromatic ring-current interaction within a strand, enhancing base stacking, which in turn strengthens duplex formation. Additional studies have indicated that within single strands, chemical shift parameters for 3'-adjacent bases are greater than those for corresponding 5'-neighbours. Larger shift parameters (Kearns & Shulman, 1974) are used for flanking 3'-neighbours when assigning ring protons in A·U and G·C hydrogen-bonded systems.

Comparison of $G\vec{C}A$ with $U\vec{G}CA$ is significant. The chemical shift *vs.* temperature data for the aromatic and ribose H-1' protons of UGCA at 8.2 mM are contained in Table 28 and its average T_m was 33°C. Remarkably, the $G\vec{C}A$ duplex which contains only two Watson-Crick base pairs and two dangling adenosine residues, is equal in stability to the $U\vec{G}CA$ duplex which contains four Watson-Crick base pairs. It is considered that a combination of factors: base-stacking, hydrophobic

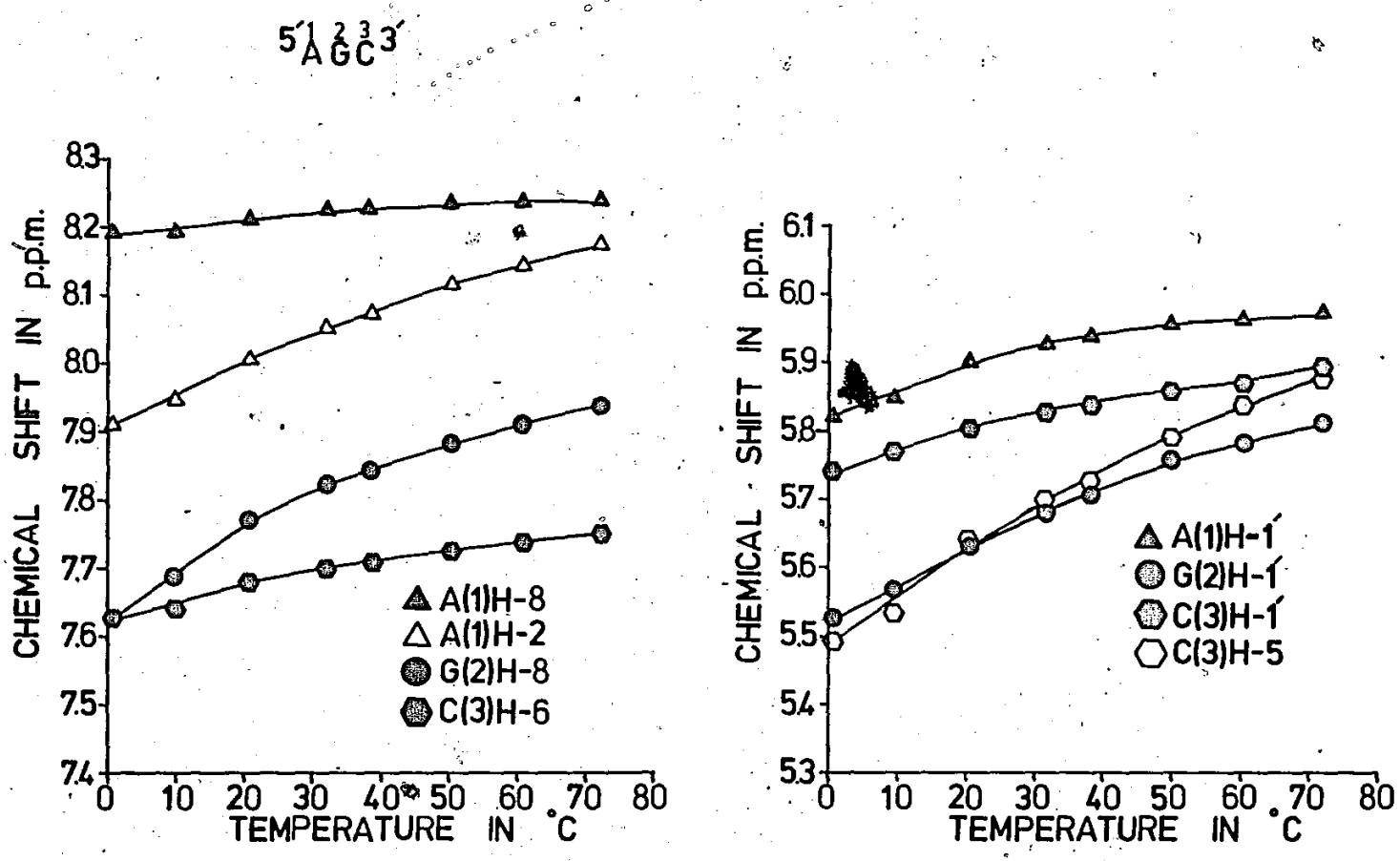


Figure 31. Variable temperature results for the triribonucleotide AGC.

TABLE 30

NMR CHEMICAL SHIFT ASSIGNMENTS FOR GCAA (7.3 mM)

OVER THE TEMPERATURE RANGE 70°C-0°C

	70.8°	60.8°	51.0°	45.9°	40.5°	35.7°	30.2°	25.1°	19.6°	9.4°	T_m °
A(3)H-8	8.313	8.306	8.278	8.257	8.224	8.166	8.098	8.080	8.074	8.047	36
A(4)H-8	8.313	8.306	8.278	8.257	8.224	8.190	8.138	8.120	8.117	8.047	38.5
A(4)H-2	8.181	8.166	8.150	8.138	8.126	8.114	8.098	8.080	8.074	8.047	35
A(3)H-2	8.101	8.065	8.001	7.943	7.857	7.760	7.632	7.516	7.418	7.256	31
GH-8	7.909	7.912	7.921	7.928	7.949	7.967	7.989	8.013	8.045	8.047	31
CH-6	7.717	7.713	7.707	7.699	7.688	7.676	7.646	7.640	7.635	7.625	35
A(4)H-1'	6.038	6.024	6.001	5.979	5.957	5.932	5.884	5.874	5.880	5.874	33.5
A(3)H-1'	5.977	5.957	5.940	5.930	5.917	5.905	5.884	5.832	5.829	5.789	NSB
CH-5	5.849	5.808	5.739	5.685	5.612	5.530	5.437	5.357	5.283		36
CH-1'	5.867	5.849	5.817	5.794	5.757	5.713	5.657	5.609	5.572	5.517	33
GH-1'	5.809	5.793	5.806	5.794	5.795	5.803	5.803	5.808			

Average T_m = 34

NSB: no sigmoidal behaviour

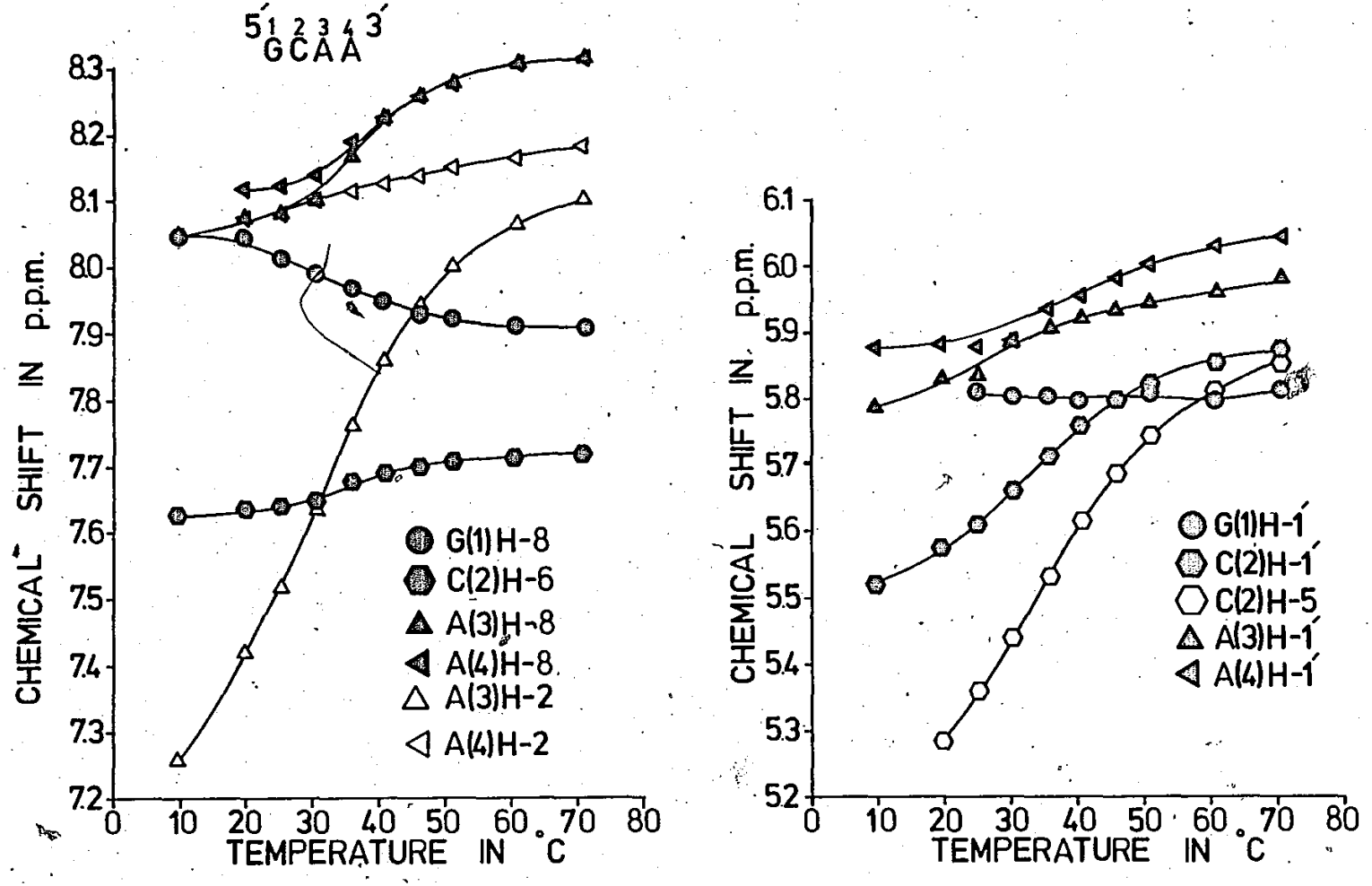


Figure 32. Results of the variable temperature study of GCAA.

interactions, solvation and entropic effects, as well as Watson-Crick hydrogen-bonding, contribute to duplex stability.

Stability of the $\overrightarrow{GCAA}$ duplex was also studied and its T_m found to be 34°C at 7.3 mM (Table 30, Figure 32). Behaviour was similar to that for $\overrightarrow{GCA}$ and it was noteworthy that the effects of 3'-terminal dangling adenosines were cooperative. However, the residue immediately adjacent to the base-paired region appears to make the major contribution to duplex stability, considering the additional 3'-adenosine made little contribution to overall duplex stability.

4.5. Contribution from Non-Paired Residues to the Direction of Stacking and its Importance in Duplex Stability (Alkema et al. 1981).

The preceding sections strongly indicated that the secondary structure of RNA results from more than simple Watson-Crick base pairing. Non-complementary base pairs, dangling bases, loop size and terminal fraying effects, also contribute to general stability. These factors all affect base stacking and it is becoming apparent that base stacking is as important to duplex stability as base pairing. Base stacking is also controlled by length and nearest-neighbour interactions (i.e. sequence). However, the results obtained with $\overrightarrow{GCA}$, when compared to $\overrightarrow{AGC}$, have brought forward an influence on base stacking that strongly affects duplex stability. The role of the direction of base stacking in short RNA duplex stability will now be considered.

The results obtained in the $\overrightarrow{GCA}$ study were significant not only because it was the first trimer to form a duplex but also, because of the comparison with $\overrightarrow{AGC}$ which did not duplex. The 5'-dangling adenosine in $\overrightarrow{AGC}$ did not stabilize the duplex. This finding was supported by the

results for duplex AAGCU (T_m = 29°C) which also has a 5'-dangling adenosine and was less stable than the reference duplex, AGCU (T_m = 34°C). However, the duplex AGCUA with a 3'-dangling adenosine, was considerably more stable than the reference, with a T_m = 45°C (Table 24). Although the 5'-dangling adenosines did result in some shielding of the other resonances, as with a 3'-dangling adenosine, this shielding was not an indication of its effect on duplex stability. The problem was further complicated when the results for ACAUG and CAUGA were considered. In these duplexes adenosine stabilized to the same extent whether 3' or 5' (Table 24, Section 4.3.2).

In addition to providing the changes in chemical shift with temperature to indicate duplex formation, ^{1}H-NMR spectra also provide information on the solution conformation of the individual nucleosides within a sequence and thus of the entire sequence. This information can be obtained by measuring the distance between peaks in the H-1' doublet (in Hertz). The anomeric proton signal is split by coupling with the C-2' proton and is called the $J_{1',2'}$ coupling constant. These coupling constants are temperature dependent and show a gradual, if sometime erratic, decrease in magnitude as the temperature is lowered. Low values of the coupling constant are associated with high degrees of base stacking and a 3'-endo ribose ring conformation (Evans & Sarma, 1974). High values are associated with low degrees of base stacking and a 2'-endo ribose ring conformation. Therefore, as the temperature is lowered, the coupling constants decrease indicating a shift in the equilibrium from the 2'-endo state to the 3'-endo state, Figure 2. A simple equation:

$$10(10-J_{1',2'}) = \% \text{ C(3')-endo}$$

provides an estimate as to what percent each base exists in the C-3'- or C-2'-endo conformation. Graphing the percent C-3'-endo conformation or the $J_{1',2'}$ coupling constant of each base enables the researcher to determine which nucleotide stacks first and last.

The studies with the trimer series, $\vec{AG}X$ and $\vec{CA}X$, (Everett et al. 1980; Section 4.1), discovered that two different trends of stacking were possible. In general, the magnitude of the $J_{1',2'}$ values, in the $\vec{CA}X$ series, decreased in the order X>A>C at all temperatures, whereas in the $\vec{AG}X$ series the opposite trend, A>G>X, existed. Figures 33 and 34 show the plots of $J_{1',2'}$ against temperature for two representative cases, $\vec{CA}G$ and $\vec{AG}U$. In two cases, $\vec{CA}C$ and $\vec{CA}U$, there was no clear trend and in some cases, such as $\vec{AG}U$, the trend was A, G>U. However, not one case was observed where the trend was broken by following an opposite order for a given series (e.g. A<G<X). At face value the data implied that the $\vec{AG}X$ series of compounds stacked from the 3'-end to the 5'-end while the $\vec{CA}X$ series stacked from the 5'-end to the 3'-end. The reasons for this interesting divergence in behaviour of the two series were unclear, especially since the trend for stacking ability was accepted as being A>G>C>U. These results seem to imply that the direction of stacking is sequence dependent.

A look at the coupling constants for $\vec{GC}A$ and $\vec{AG}C$ revealed that each of these sequences had different and distinct stacking patterns (Figures 35 and 36). This suggested that the direction of base stacking might play a role in duplex stability. Consideration of this new information led to a reinvestigation of other duplexes with dangling bases. An intriguing sequence related effect soon emerged.

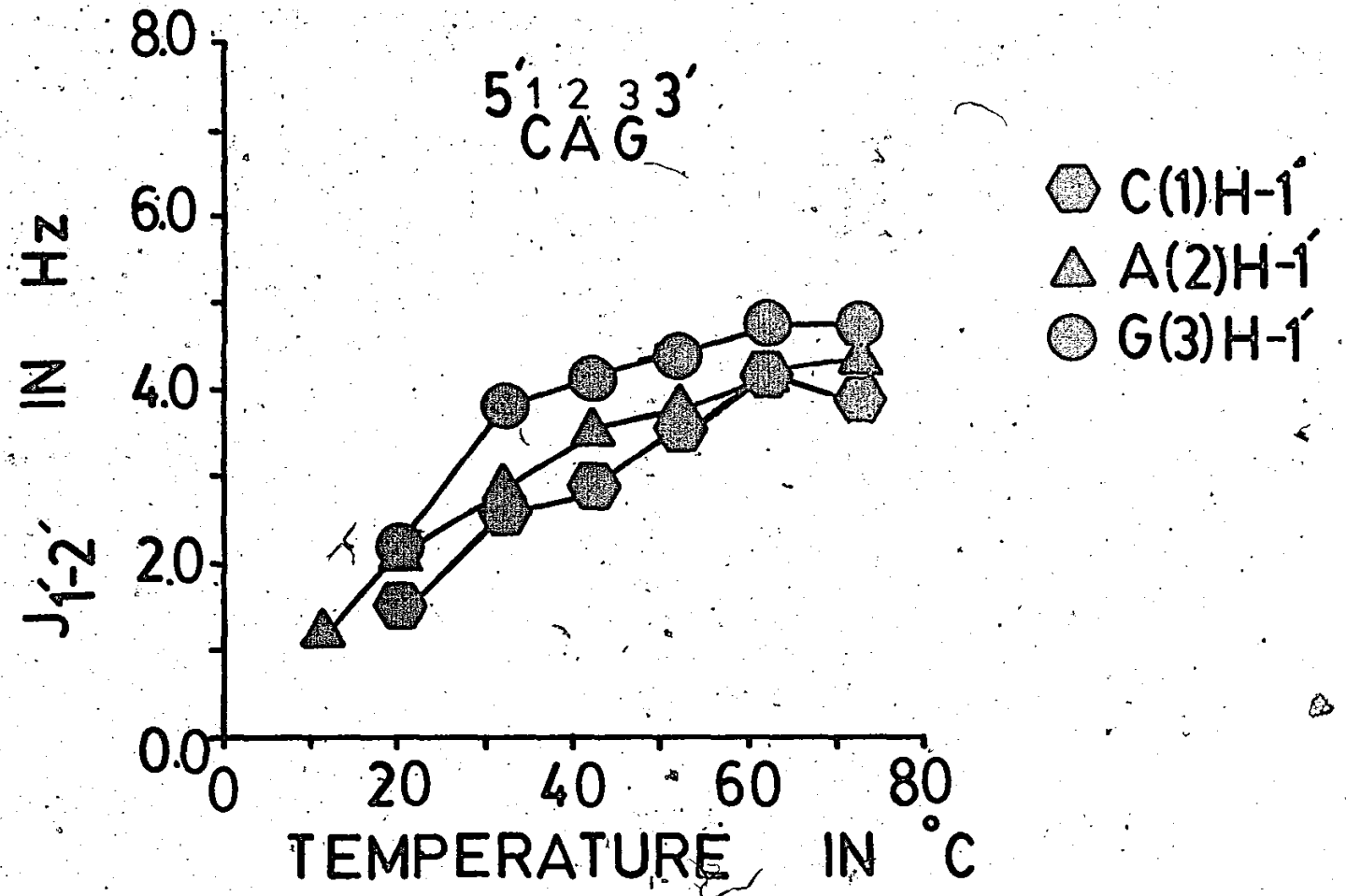


Figure 33. Results of the variable temperature study of the $J_{1'-2'}$ coupling constants for CAG.

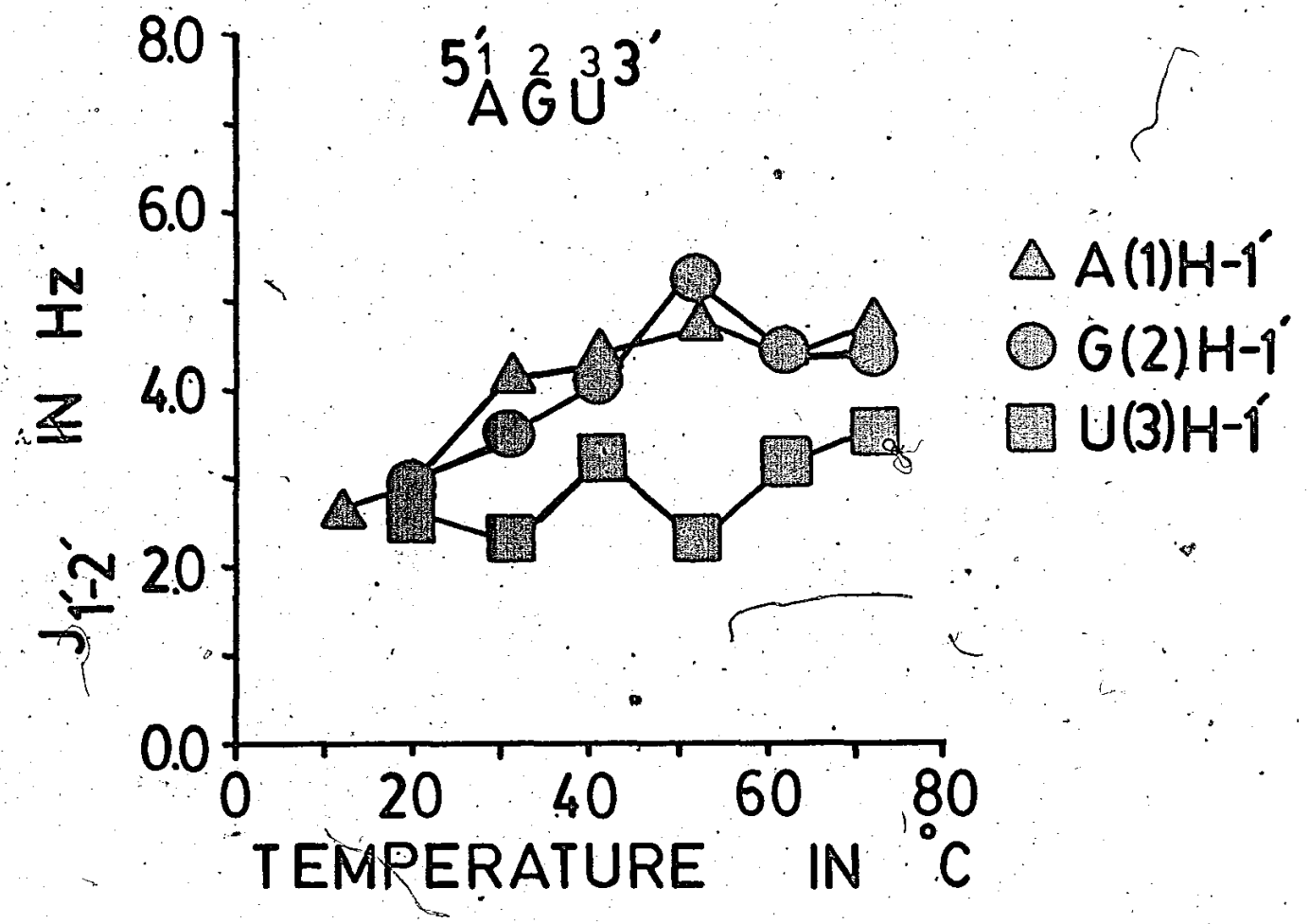


Figure 34. Results of the variable temperature study of the $J_{1'-2'}$ coupling constants of AGU.

Sequences, GĈA, AĜC, AĜCU, AGĈUA, AAĜCU, AAĜCUU, AAĜCUUA, AAĜCUA, UĜCA and UĜCAA, all fit into a general type with a GC core, terminal AU base pairs and a 3'- and/or 5'-dangling adenosine. Duplexes formed from GĈA, UGĈAA, AGĈUA and AAĜCUUA all clearly demonstrate the increased helix stability derived from the contribution of a non-base paired adenosine residue. Comparison of GĈA with UĜCA (Section 4.4), both of which exhibit a T_m of 33°C, dramatically illustrated that the stability of GĈA was primarily due to the presence of 3', well stacked, non-base paired adenosine residues which were equivalent to A·U pairs in their contribution to duplex stability. On the other hand, the corresponding duplex from AĜC, also contains a GC core but had 5'-dangling adenosines and exhibited an estimated T_m ≈0°C. These adenosine residues, also non-base-paired, must contribute less to overall duplex stability. Presumably, the greater helical overlap of a 3'-base residue generated a greater aromatic ring-current interaction within a strand, enhancing overall base stacking which, in turn, strengthened duplex formation. This observation clearly contrasted with those seen earlier in the case of oligomer CÂUG (T_m = 24°C) where the dangling adenosine residues increased stability by +11°C, irrespective of whether the base was 3' or 5' (Neilson et al. 1980; Section 4.3). It should be noted, however, that the 3'-dangling adenosine flanked a guanosine residue in CAÛGA and the 5'-dangling adenosine flanked a cytidine residue in ACÂUG.

The $J_{1',2'}$ coupling constants observed at all the H-1' signals in GĈA and AĜC showed a gradual decrease in magnitude as the temperature was lowered. Figures 35 and 36 show the plots of $J_{1',2'}$ against

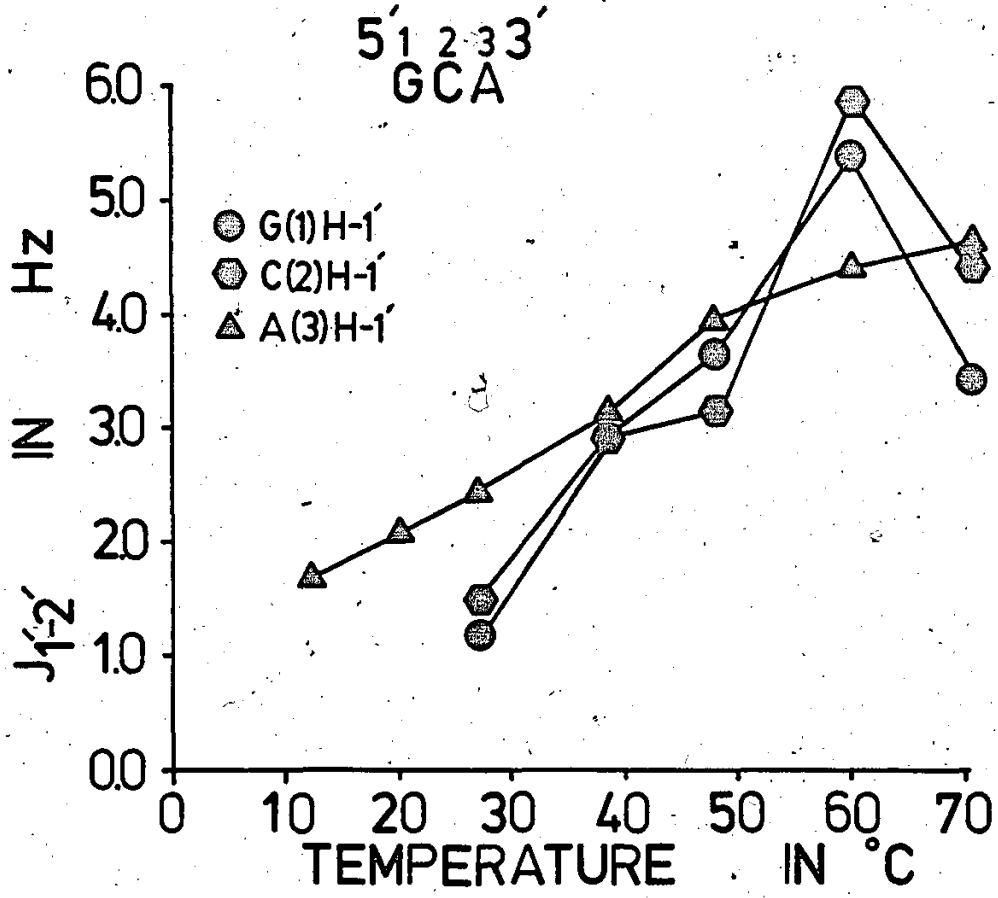


Figure 35. The coupling constant ($J_{1'-2'}$) versus temperature plots for the H-1' protons of GCA.

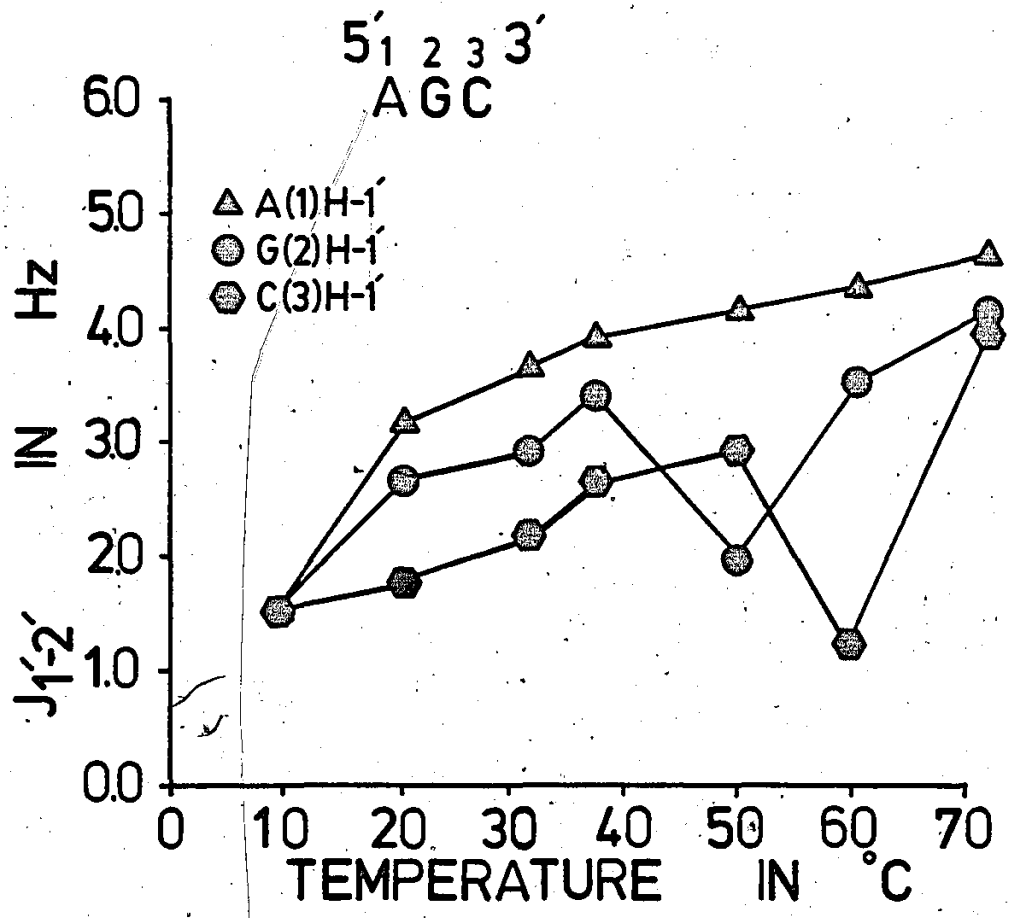


Figure 36. The coupling constants ($J_{1'-2'}$) versus temperature plots for the H-1' protons of AGC.

temperature for these two trimers. Two opposite trends in the behaviour of these coupling constants were noted. In the trimer $\vec{A}GC$ the magnitude of the $J_{1',2'}$ values decreased in the order A>G>C, whereas in the oligomer $\vec{G}CA$ the opposite direction of stacking (A>C>G) existed.

In the dinucleoside monophosphate series, low values of the coupling constants have been associated with a C(3')-endo ribose ring conformation and a high degree of base stacking (Lee et al. 1976; Ezra et al. 1977). If a similar relationship exists between the magnitude of $J_{1',2'}$ and base stacking for trimers as for dimers, then the present data imply that the trimer $\vec{A}GC$ stacks from the 3'- end to the 5'- end, while the $\vec{G}CA$ oligomer stacks from the 5'- end to the 3'- end.

The direction of stacking in the two trimers appeared to have a significant influence on the process of duplex formation. The 5' to 3' direction of stacking enhanced duplex formation while the 3' to 5' direction seemed to have an unfavourable effect on duplex stability.

To investigate this idea further, some additional trimers which, theoretically could form duplexes similar to $\vec{G}CA$, were studied. Table 31 lists a number of trimers investigated and indicates the direction of stacking found in each. Surprisingly, only GCA formed a stable RNA duplex. Even $\vec{C}GA$, which was expected to duplex did not (Figure 37). Although $\vec{C}GA$ displayed the favourable 5'→3' stacking direction (Figure 38), the CG core, known to be less stable than a GC core, appears to prevent duplex formation of the trimer. Still the chemical shift vs. temperature plots showed strong upfield movement indicative of strong base stacking and the effect of the 3'-dangling adenosine. On the other hand, $\vec{A}CG$, with its 5'-dangling adenosine, showed no direction of stacking and

TABLE 31

Direction of stacking in a variety of triribonucleotides

SEQUENCE	DIRECTION OF STACKING	T_m IN °C
GCA	5' → 3'	33
AGC	3' → 5'	< 0
CGA	5' → 3	< 0
ACG	None*	< 0
GCU	3' → 5'	< 0
UGC	None*	< 0
GGC	3' → 5'	< 0
CGG	5' → 3'	< 0
CGC	3' → 5'	< 0
CCG	5' → 3'	< 0

* $J_{1',2'}$ vs. temperature plots did not indicate a clear direction of stacking

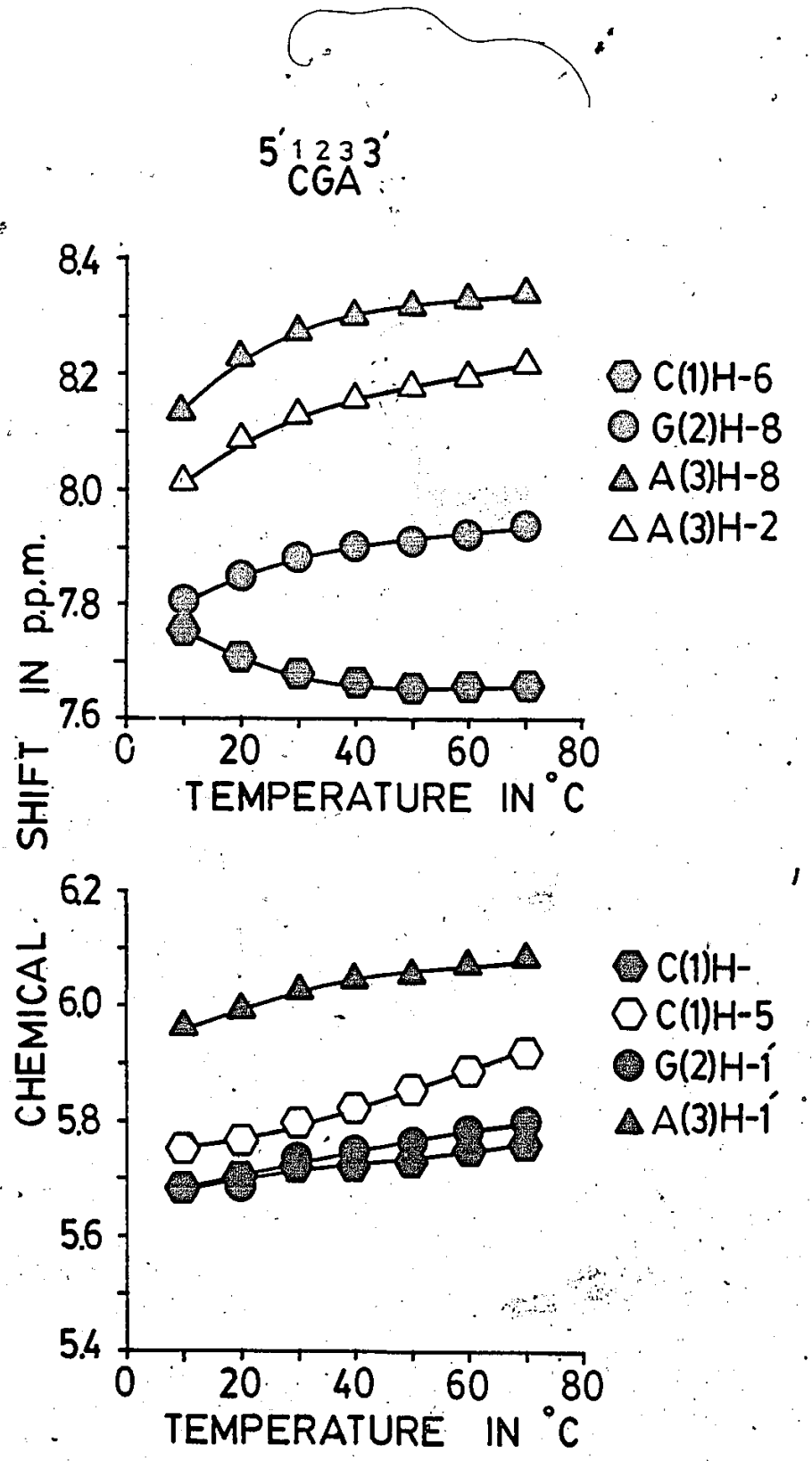


Figure 37. Variable temperature plots of the aromatic base and anomeric resonances of CGA.

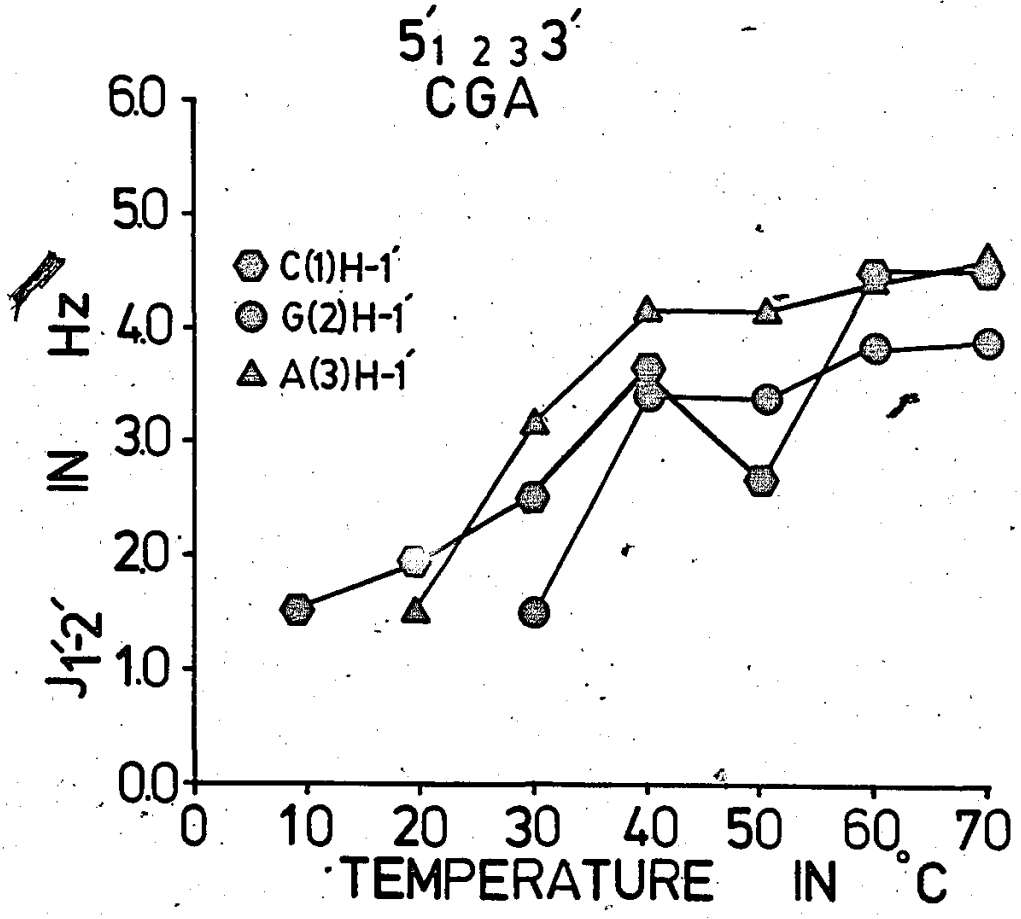


Figure 38. The coupling constant ($J_{1'-2'}$) versus temperature plots for the H-1'-protons for CGA.

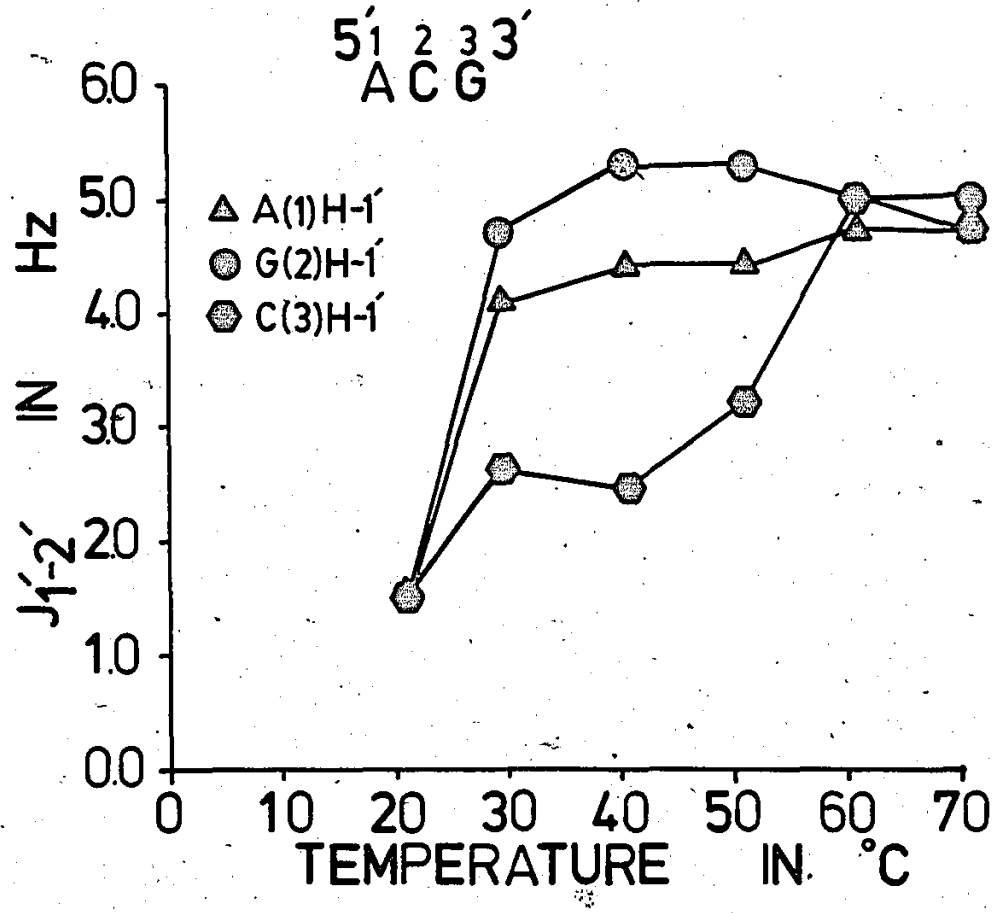


Figure 39a. The coupling constants ($J_{1'-2'}$) versus temperature plots for the H-1' protons of ACG.

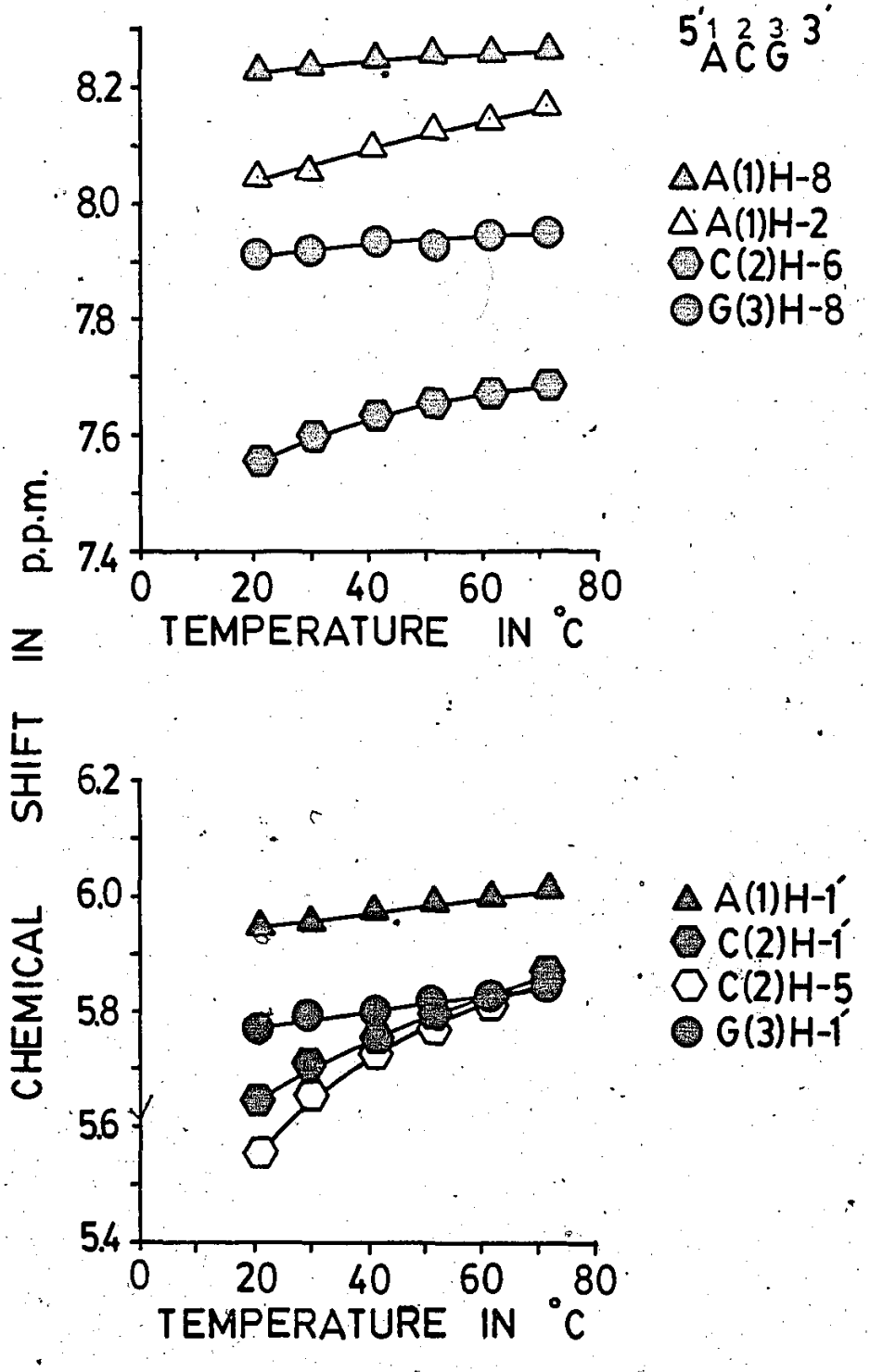


Figure 39b. Variable temperature plots of the aromatic base and anomeric resonances of ACG.

almost straight lines in the temperature vs. chemical shift plots (Figure 39a & b).

The difference in the stacking patterns of the trimers with terminal adenosines suggest that adenosine may influence the overall stacking pattern of the sequence. In addition, a CG core contributes less to overall duplex stability than a GC core, as suggested in earlier studies (Borer et al. 1974). This result is supported by work with $\vec{A}CGU$ which has a T_m of 29°C as opposed to $A\vec{G}CU$ with a T_m of 34°C. These two duplexes differ only in the orientation of their G·C base pairs, indicating that a CG sequence destabilizes the duplex.

With the results obtained with $G\vec{C}A$ and $A\vec{G}C$ in mind, attention was turned to the direction of stacking in sequences $AA\vec{G}CU$ and $AG\vec{C}UA$. Duplex formed by oligomer $AA\vec{G}CU$, which contains a 5'-dangling adenosine residue, also exhibits an overall 3'- end to 5'- end base stacking (Figure 40), while the duplex from $\overrightarrow{AGCUA}$, which contains a 3'-dangling adenosine, exhibits the more stabilizing 5'- end to 3'- end base stacking (Figure 41), and has a T_m = 46°C (Table 24). Duplex $AA\vec{G}CU$ has an even lower T_m (30°C) (Table 24) than that of the parent duplex $A\vec{G}CU$ (T_m = 34°C). 5'-dangling adenosine residues apparently destabilize the duplex formed by $AA\vec{G}CU$ by altering the base stacking direction, subsequently lowering the helical overlap of the base residues and possibly increasing instability of the terminal base pairs in the $AA\vec{G}CU$ duplex. In contrast, in duplexes formed by oligomers $AG\vec{C}UA$ and also $UG\vec{C}AA$ the 3'-dangling adenosine residue extends the 5'- end to 3'- end base stacking, strengthening the terminal A·U base pairing, and increasing the overall stability.

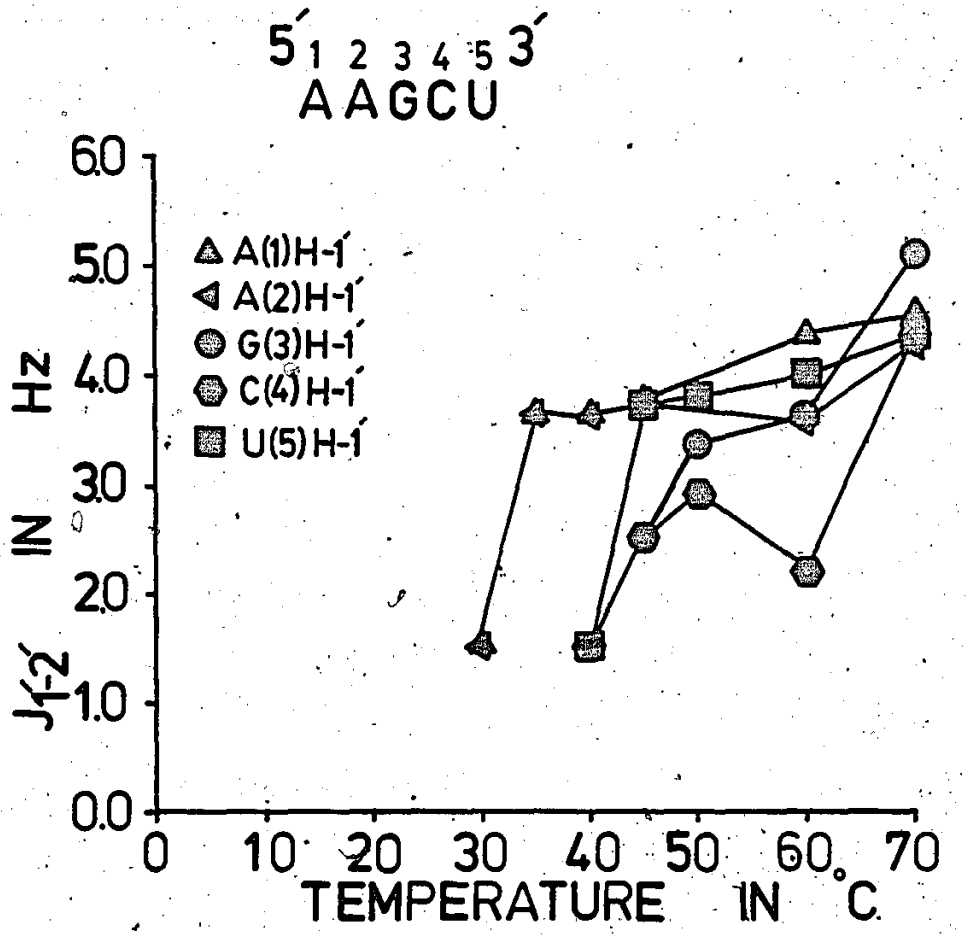


Figure 40. The coupling constant ($J_{1'-2'}$) versus temperature plots for the H-1' protons of AAGCU.

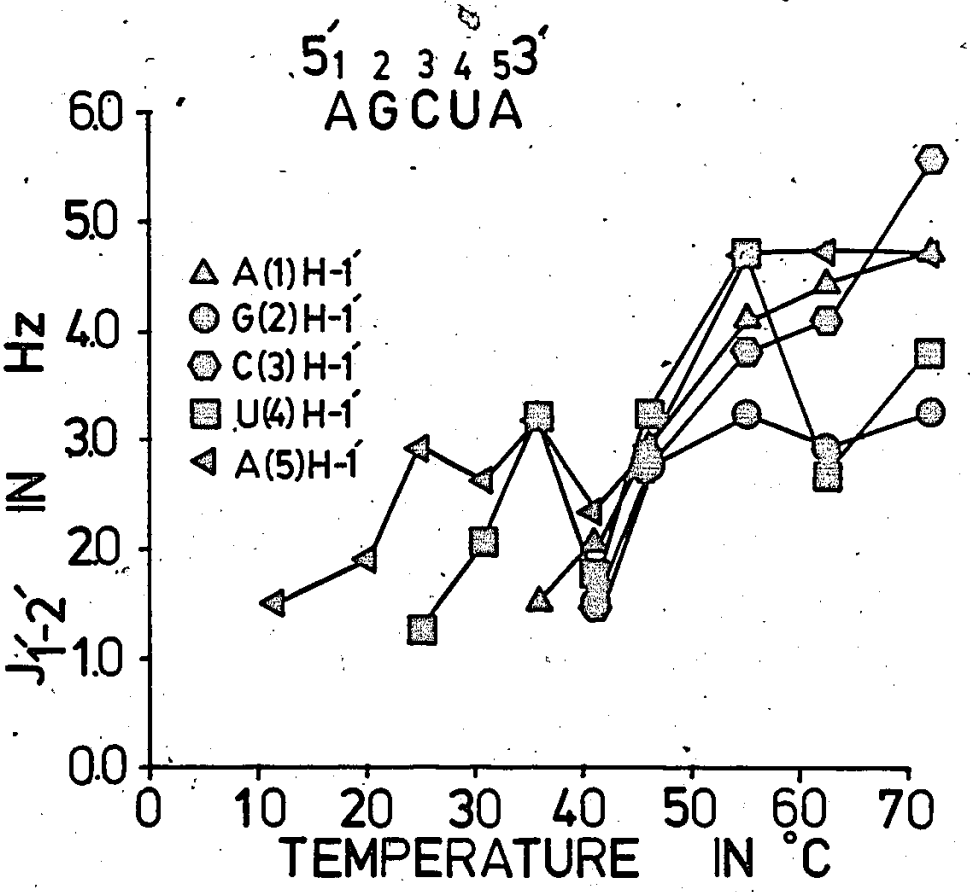


Figure 41. The coupling constant ($J_{1',-2'}$) versus temperature plots for the H-1' protons of AGCUA.

With both the $\vec{A}AGCUU$ and $\vec{A}AGCUUA$ duplexes the bases stack in the favourable 5' to 3' direction overall. However, in both cases the 5'-adenosine was the last base to stack (Figures 42 and 43). This result indicates that the 5'-adenosine behaves in a manner similar to the 5'-dangling adenosine of $\vec{A}AGCU$ where strand stacking occurred in the opposite direction. Because the 5'-adenosine is H-bonded in A·U base pairs in both $\vec{A}AGCUU$ and $\vec{A}AGCUUA$ duplexes, it does not appear to destabilize duplex formation, the lack of stacking interaction is compensated for by the H-bonding (Gupta & Sasisekharan, 1978a).

Duplex formed from $\vec{A}AGCUA$ contains <u>two</u> 3'- and <u>two</u> 5'-dangling adenosines and exhibited a slightly higher T_m (48°C) than duplex $\vec{A}GCUA$. Apparently the effect of the 3'-dangling adenosines predominated and directed base stacking to adopt the 5' to 3' direction (Figure 44). The additional 5'-dangling adenosines became an extension to this established stacking and so enhanced duplex stability.

Considering the results presented above, why did both 5'- and 3'-dangling residues on tetramer $\vec{C}AUG$ stabilize to the same extent (Section 4.4)? When the $J_{1',2'}$ coupling constants for these sequences were studied, no pattern or direction of stacking could be detected. Instead the coupling constants appeared to collapse at higher temperatures, than in $\vec{C}AUG$, indicating that complete stacking of the bases occurred earlier and thus would result in base pairing at higher temperatures and therefore a higher T_m.

The results presented in this section demonstrate the importance of the direction of base stacking on duplex stability. This phenomenon appears to be sequence dependent and is, therefore, of importance in

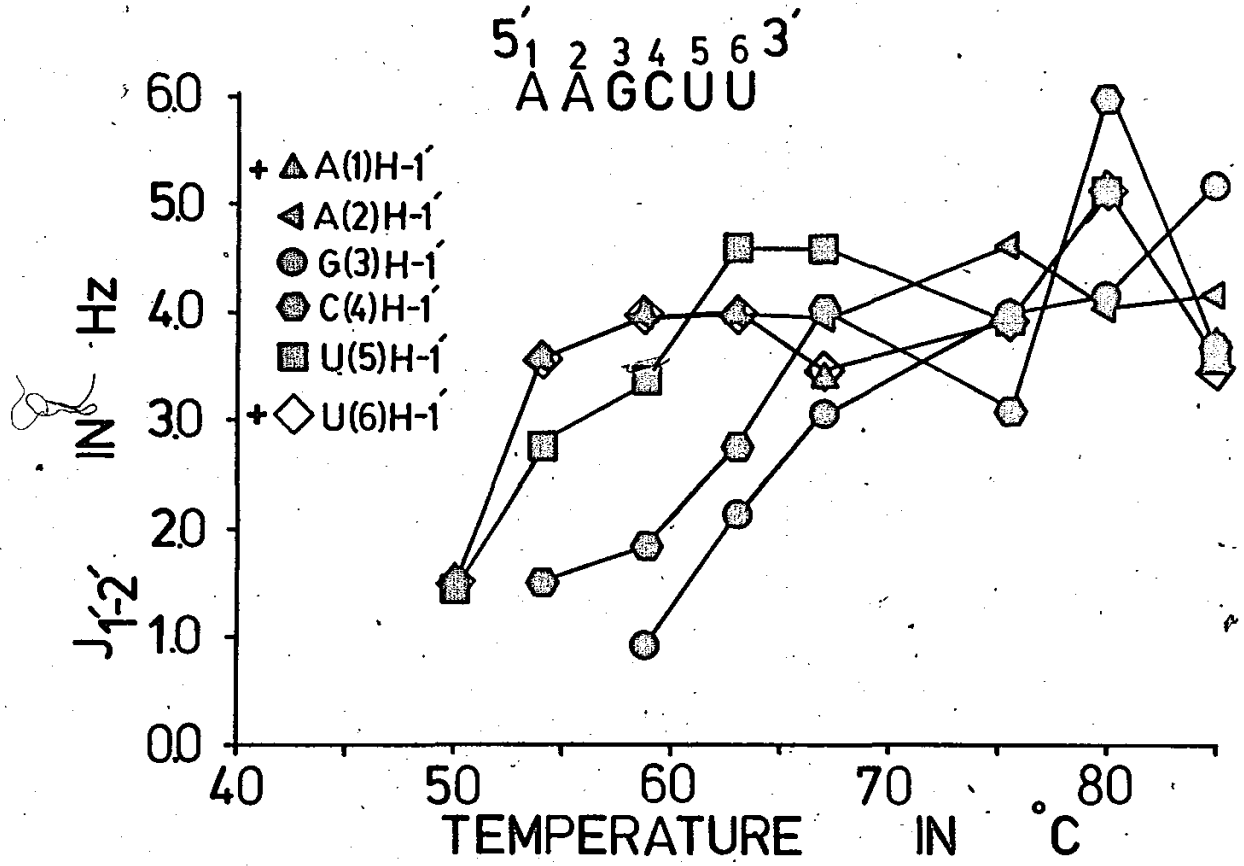


Figure 42. The coupling constant ($J_{1'-2'}$) versus temperature plots for the H-1' protons of AAGCUA.

+ = coincidental

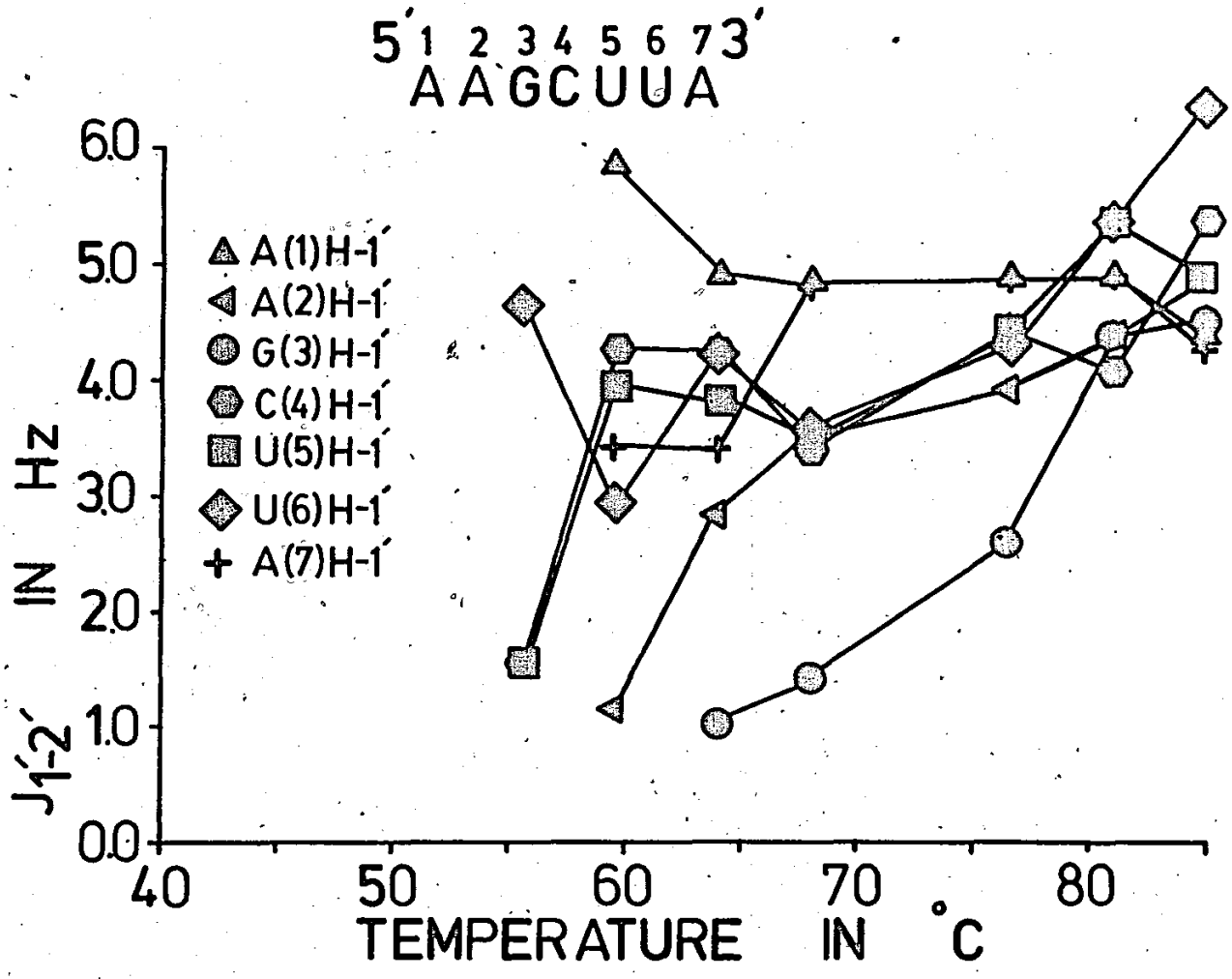


Figure 43. Variable temperature study of the $J_{1'-2'}$ coupling constants of AAGCUUA.

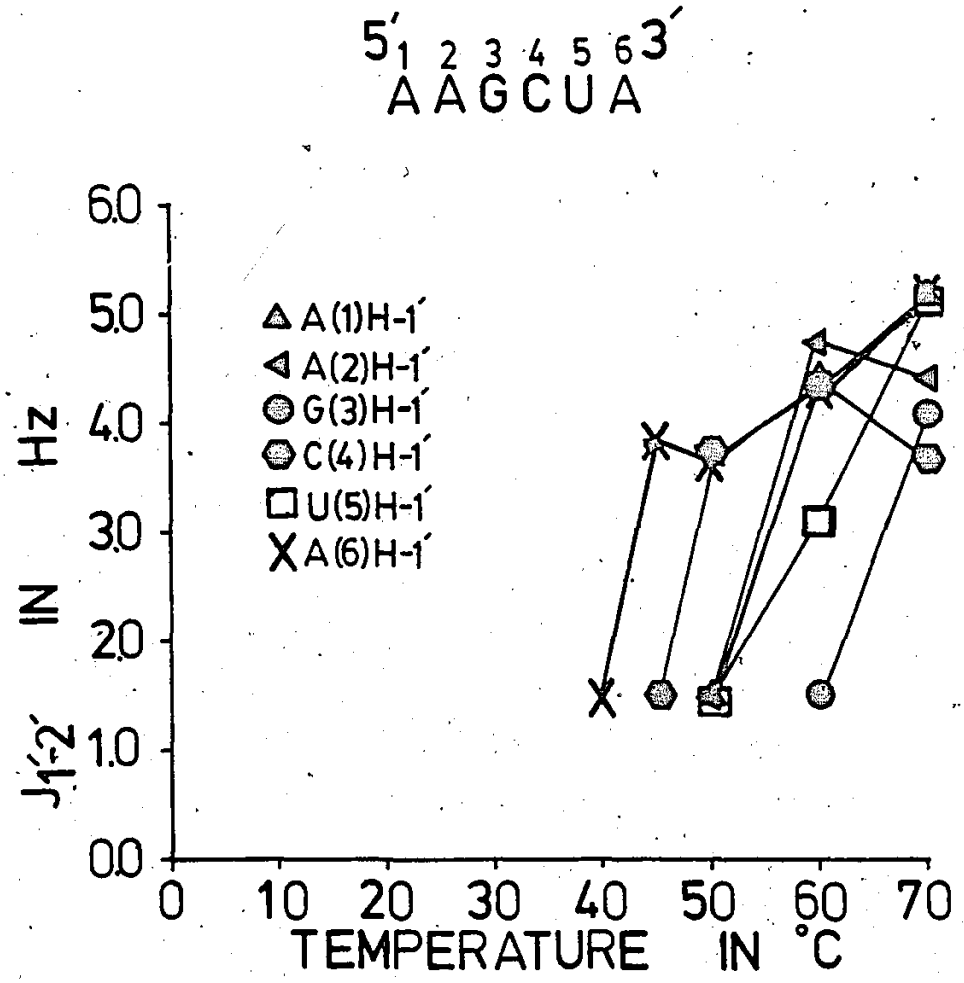


Figure 44. The coupling constant ($J_{1',2'}$) versus temperature plots for the H-1' protons of AAGCUA.

the consideration of sequences to be used in stability studies. It is tempting to speculate on the biological significance of possible changes in duplex stability promoted by reverse base stacking as a result of interaction with proteins or other nucleic acid molecules, for example, in polymerase recognition sites.

4.6. Invariant Adenosine Residues Stabilize tRNA D-Stems (Alkema et al. 1981; accepted by FEBS Letters)

Examination of the duplex formed by AAGCUA proved that terminal A·A non-base paired oppositions could exist and in fact stabilized duplex formation significantly. If the terminal adenosines are considered to be the beginning of loops adjacent to duplex RNA stems:

e.g. 

then this is direct evidence that loops stabilize stem regions as previously suggested (Pohl, 1974; Grosjean et al. 1976). These results also suggest a possible role for the frequent occurrence of adenosines at positions 14 and 21 of tRNAs.

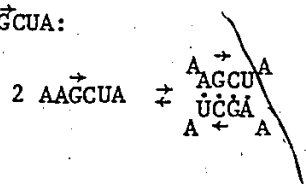
Sequence inspection of reported tRNAs reveals that two adenosine residues nearly always occupy positions 14 and 21 at the D-loop stem junction. Apparently, the nature of these residues has been conserved and perhaps they are involved in some tRNA function, for example, amino-acyl-tRNA synthetase recognition (Dubock et al. 1971; Roe et al. 1973). Their presence may also satisfy a structural aspect controlling native conformation, namely an invariant A·U base pair (Kim et al. 1974; Rich, 1977). The controversy may be resolved by proposing that the invariant adenosines stabilize D-stem duplexes; features of secondary cloverleaf

structure (Holley et al. 1965). D-stems contain only three or four Watson-Crick base pairs while other stem duplexes have five or more. Melting studies have shown D-stems to be the least stable regions (Hilbers & Shulman, 1974).

Three-dimensional tertiary structure of tRNA showed adenosine residues 14 and 21 (see Figure 45), to be coplanar, and each base stacked to the adjacent D-stem duplex (Rich, 1977). Steric tolerance exists, however, since the adenosine at position 14 was displaced within the helix as a result of its participation in a tertiary Sobell-type A·U base pair (Haschmeyer & Sobell, 1963) with an invariant uridine at position 8. This displacement from the normal RNA-A helical geometry at the ends of a duplex can be considered as partial strand unwinding. Extension of base stacking to the invariant adenosine residues, is still possible, and therefore enhances overall duplex stability.

In the continuing quest to evaluate the various factors affecting RNA duplex stability, it has been established that base stacking plays a significant role. The contribution from non-base paired 3'-terminal (dangling) adenosines was a major factor influencing overall helical stability (Section 4.4). Internal A·A non-bonded pairs have been confirmed as centres of instability, for example. CAÂUG:CAÂUG (T_m <0°C) and AGÂCU:AGÂCU (T_m ≈26°C).

Model studies on D-stem melting can be carried out using the duplex formation of AAĜCUA:



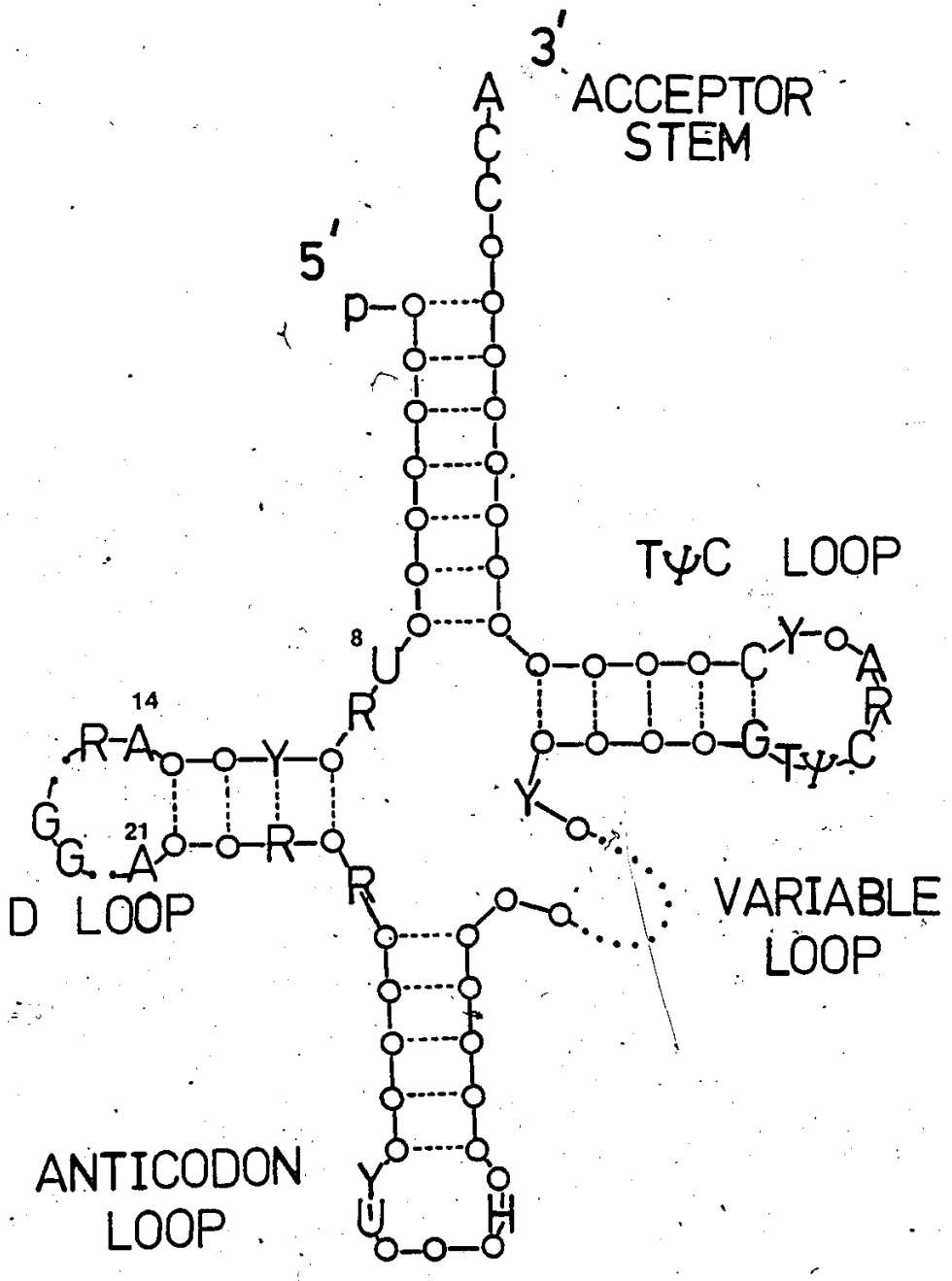


Figure 45. General tRNA cloverleaf structure indicating the invariant and semi-invariant bases for tRNA's other than initiator tRNA's, Y stands for pyrimidine, R for purine, H for hypermodified purine. Numbering system corresponds to that of yeast $tRNA^{Phe}$ (Rich, 1977).

Comparison of the effects for opposing terminal adenosines with a single terminal adenosine is presented in Table 32.

Table 32. Melting temperatures of synthetic duplexes.

Oligomer	T_m (°C)
AG̃CU	33
AGC̃UA	45
AAG̃CUA	48

Clearly, terminal non-bonded adenosine residues contribute to duplex stability. The ability of the short duplex to partially unwind at the ends allows the terminal adenosines to exist in opposition to each other; however, the displacement was not of sufficient magnitude to interrupt the extended base stacking which was enhanced by these adenosines.

When these model studies are applied to tRNA secondary structure, D-stems flanked by two adenosines at the neck-loop junction will be more stable than stems lacking adjacent non-bonded adenosines. "Evolution" of tRNA structure has resulted in a delicately balanced steric arrangement where a tertiary Sobel-type A·U base pair displaces an adenosine residue sufficiently to remain opposite another adenosine , but not to interfere with extended base-stacking interactions. The invariance of adenosines at positions 14 and 21 ensures a more stable D-stem as well as distinct loop formation due to strand separation.

5. CONCLUSIONS

The phosphotriester method, developed by Neilson and coworkers, for the preparation of short oligoribonucleotide sequences, was extensively employed in the synthesis of a large variety of RNA sequences. These sequences were designed with several features, found in naturally occurring RNA, that would allow the study of perturbations affecting RNA helix structure and stability. Features investigated include: non-complementary base oppositions, non-base paired regions found adjacent to duplex regions, changes in base sequence, G·U base pairs, invariant adenosines and base stacking.

The study of short RNA sequences requires a technique capable of measuring conformational changes in the individual nucleoside residues. Proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy has that capability. By monitoring changes in the local microenvironments of the various hydrogens found on the nucleoside residues, it was possible to discern how each nucleotide in the sequence was affected by the perturbations introduced.

Monitoring the changes in chemical shift over a specified temperature range (70°C-0°C) enables the determination of the melting point (Tm) of any duplex. Changes in the Tm due to disruptions in the duplexes provide a means for measuring the effect of the alteration on helix stability. Changes in the anomeric coupling constant, measured from the spectra, make it possible to observe changes in ribofuranoside conformation and in the base stacking of the sequence.

In order to use NMR to its maximum capacity in the study of nucleic acids, it was essential to be able to make unambiguous assignment of the spectral resonances of most protons found in any given sequence. To this end, series of trimers, $C\vec{A}X$ and $A\vec{G}X$ (X ≡ A, G, C or U), were studied to determine the shielding effects that adding a nucleotide to the end of a sequence would have on the chemical shifts of the dimers $C\vec{A}$ and $A\vec{G}$. The magnitude of these shieldings decreased in the order A>G>C>U. Similarly, the sequences $C\vec{X}G$, $A\vec{X}G$, $A\vec{G}XC$, $CA\vec{X}UG$ and $AG\vec{X}CU$ were studied to determine the effects of inserting a nucleotide residue into a given sequence, had on the chemical shifts of neighbouring residues. In all cases, the order was A>G>C>U. These results, coupled with the incremental assignment procedure of Borer et al. (1975) provides an accurate assignment procedure.

Synthesis of the pentaribonucleotide series, $AG\vec{X}CU$, provided a system for the study of the effects of an internal non-complementary base opposition on duplex stability. All four pentamers could theoretically form a duplex with an internal non-bonded base pair. Earlier results with the $CA\vec{X}UG$ series indicated duplex formation did not occur (Romaniuk et al. 1979). However, the $AG\vec{X}CU$ study revealed that the neighbouring G·C base pairs provided sufficient stability for duplex formation to occur when X was adenosine or cytidine. Duplex $AG\vec{A}CU$:$AG\vec{A}CU$ had a Tm of 26°C and duplex $AG\vec{C}CU$:$AG\vec{C}CU$ had a Tm of 25°C. It was suggested that the stacking interactions of the mismatched bases along their individual strands allowed duplex formation to occur, although it was significantly destabilized. $AG\vec{G}CU$ also formed a stable duplex with a Tm of 30°C, however, this duplex was probably of a different nature,

with two internal G·C base pairs, two terminal G·U base pairs and two 5'-terminal dangling adenosines. The poor stacking interactions of uridine weaken the overall stacking of the single strand, AGÛCU, and disrupted duplex formation by this pentaribonucleotide.

Mixing of two appropriate pentamers in this study made it possible to determine the effects of an additional A·U, G·C or G·U base pair on duplex stability when compared with the parent duplex AĜCU (Tm = 34°C). In addition, these experiments provided data that could be used to compare the effects on flanking G·C or A·U base pairs on duplex stability by using the results from the CAX̂UG study (Romaniuk et al. 1979). The duplex formed by AGÂCU:AGÛCU had a Tm of 45°C while the duplex AGĜCU:AGĈCU had a Tm of 54°C. It was determined that the G·C base pairs flanking the addition internal base pair provided an additional 6°C of thermal stability as compared with flanking A·U base pairs.

Investigation of the duplex formed by AGĜCU:AGÛCU (Tm = 44°C) which contains an internal G·U base pair, found that this G·U base pair was equivalent to an internal A·U base pair in its contribution to duplex stability. This provided an interesting contrast to earlier results which indicated that an internal G·U base pair was destabilizing (Romaniuk et al. 1979). The difference was attributed to the stronger, neighbouring G·C base pairs being able to accommodate any helical distortion due to the G·U base pair.

The effects of non-base paired regions adjacent to duplex regions was considered important in codon-anticodon interactions (Grosjean et al. 1976). To study this phenomenon closely, a number of sequences with 5' and/or 3' terminal non-base paired (dangling) adenosines were assembled

and examined. These studies showed that a terminal 3'-dangling adenosine contributed an average of 11°C to the Tm of the duplex or is approximately equal to an additional A·U base pair. This study was also the first to reveal that a 3'-dangling adenosine could contribute significantly more to duplex stability than a 5'-dangling adenosine, although this effect was also sequence dependent. It was found that the duplex formed by AGCUA had 3'-dangling adenosines and a Tm of 46°C compared to that of AGCU (Tm = 34°C). However, the duplex formed by AAGCU had 5'-dangling adenosines and had a Tm of only 30°C. In this case, the 5'-dangling adenosines actually destabilized the duplex.

The study of GCA was the first to demonstrate that a triribonucleotide could form a stable RNA duplex. This duplex had two G·C base pairs and two 3'-dangling adenosines with a Tm of 33°C. The stability of this trimer was equal to that of the tetramer duplex formed by UGCA with four Watson-Crick base pairs. Interestingly, the trimer AGC, which could form a duplex similar to GCA but with 5'-dangling adenosines, showed no indication of duplex formation.

The results from dangling adenosines and GCA studies were interesting in that they indicated that duplex stability was dependent on more than simple Watson-Crick base pair formation. Dangling adenosines through extended base stacking enhance duplex stability. Also of significance, it appeared that the direction of base stacking was involved in duplex formation and stability. Moreover, the location of the dangling adenosine appeared to influence the direction of base stacking.

A closer look at the coupling constants of the sequences with

dangling bases concluded that with a 3'-dangling adenosine, such as in $\overrightarrow{AGC}UA$ or $\overrightarrow{GC}A$, the base stacked from the 5' end to the 3' end and duplex stability was enhanced. However, sequences with 5'-dangling adenosines, such as $\overrightarrow{AG}C$ and $\overrightarrow{AAG}CU$, the bases stacked from the 3' to 5' end and the duplex stability decreased. Therefore, the direction of base stacking definitely plays a role in helix stability.

The sequence, $\overrightarrow{AAG}CUA$, was studied and found to form a duplex with both 3'- and 5'-dangling adenosines and a Tm of 48°C. This sequence provided a model for the D-stem region of tRNA which contains invariant adenosines at positions 14 and 21 and purines at positions 9 and 26. It was proposed that these invariant adenosines stabilized the D-stem which is the weakest duplex region in tRNA (Hilbers & Shulman, 1974).

These results establish the importance of base sequence in duplex stability and also show that single non-base paired oppositions within a duplex can occur. In addition, the importance of non-base paired regions in determining duplex stability was confirmed. The contribution of base stacking and the direction of base stacking was emphatically demonstrated. These studies also firmly establish the utility and success of NMR in the study of nucleic acid molecules.

BIBLIOGRAPHY

Adamiak, R.W., Arentzen, R.W. and Reese, C.B. (1977) Tetrahedron Letters, 1431.

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

Alkema, D., Bell, R.A., Hader, P.A. and Neilson, T. (1981) in "Biomolecular Stereodynamics". Sarma, R.H., ed., Elmsford, NY, Adenine Press, p. 417.

Alkema, D., Bell, R.A., Hader, P.A. and Neilson, T. (1981) Accepted for publication in "Biochemistry".

Alkema, D., Bell, R.A., Hader, P.A. and Neilson, T. (1981) Accepted for publication in "FEBS Letters".

Alma, N.C.M., Harmsen, B.J.M., Hull, W.E., van der Marel, G., van Boom, J.H. and Hilbers, C.W. (1981) Biochemistry 20, 4419.

Altona, C. (1975) in "Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions". Sundaralingam, M. and Rao, S.T., eds., Baltimore, University Park Press, p.613.

Arnott, S., Hukins, D.W.L., Dover, S.D., Fuller, W. and Hodgson, A.R. (1973) J. Mol. Biol. 81, 107.

Arnott, S., Dover, S.D. and Wonacott, A.J. (1969) Acta. Crystallography, Section B, 25, 2192.

Arter, D.B., Walkers, G.C., Uhlenbeck, O.C. and Schmidt, P.G. (1974) Biochem. Biophys. Res. Comm. 61, 1089.

Behe, M. and Fesenfeld, G. (1981) Proc. Natl. Acad. Sci. 78, 1619.

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

Blackburn, B.J., Greg, A.A., Smith, I.C.P. and Hruska, F.E. (1970) Can. J. Chem. 48, 2866.

Bolton, P.H. and James, T.L. (1080) Biochemistry 19, 1388.

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

Borer, P.N., Dengler, B., Tinoco, Jr. I. and Uhlenbeck, P.C. (1974) J. Mol. Biol. 86, 843.

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

Bossi, L. and Roth, J.R. (1980) Nature 286, 123.

Brahms, J., Aubertin, A.M. and Dirheimer, G. (1969) Biochemistry 8, 3269.

Brown, D.M., Mazrath, D.I., Neilson, A.H. and Todd, A.P. (1956) Nature 177, 1124.

Broyde, S. and Hingerty, B. (1979) Biopolymers 18, 2905.

Broyde, S., Wartell, R., Stellman, S., Hingerty, B. and Langridge, R. (1975a) Biopolymers 14, 1597.

Broyde, S., Stellman, S. and Wartell, R. (1975b) Biopolymers 14, 2625.

Broyde, S. and Hingerty, B. (1978) Nucl. Acids Res. 5, 2729.

Bubienko, E., Uniak, M.A. and Borer, P.N. (1981) Accepted for publication in "Biochemistry".

Camerman, N., Fawcett, J.K. and Camerman, A. (1976) J. Mol. Biol. 107, 601.

Cheng, D.M., Dhingra, M.M. and Sarma, R.H. (1978) Nucl. Acids Res. 5, 4399.

Cheng, D.M., Danyluk, S.S., Dhingra, M.M., Ezra, E.S., Macloss, M., Mitra, C.K. and Sarma, R.H. (1980) Biochemistry 19, 2491.

Clarke, B.F.C. (1977) Prog. Nucl. Acids Res. and Mol. Biol. 20, 1.

Coleman, J.E., Armitage, I.M., Chlebowski, J.F., Otvos, J.D. and Uiterkamp, A.J.M.S. (1979) in "Biological Applications of

Cornelius, A.G., Haasnoot, J.H.J., den Hartog, J.F., de Rooij,M., van Boom, J.H. and Cornetis, A. (1979) Nature 281, 235.

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

Davies, D.B. (1978) Prog. Nucl. Magn. Reson. Spectro. 12, 135.

De Santis, P., Morosetti, S., Palleschi, A. and Savino, M. (1981a) Biopolymers 20, 1727.

De Santis, P., Morosetti, S., Palleschi, A. and Savino, ,. (1981b) Biopolymers 20, 1707.

Dhingra, M.M. and Sarma, R.H. (1978) Nature 272, 798.

Dickerson, R.E. and Drew, H.R. (1981) J. Mol. Biol. 149, 761.

Dower, S.K. and Davek, R.A. (1979) in "Biological Applications of Magnetic Resonance". Shulman, R.G. ed., New York, NY. Academic Press, p. 271.

Dudock, N., Di Peru, C., Seileppi, K. and Reszelback, R. (1971) Proc. Natl. Acad. Sci. USA 68, 681.

Early, T.A. and Kearns, D.R. (1979) Proc. Natl. Acad. Sci. USA 76, 4165.

Early, T.A. and Kearns, D.R. (1980) Nucl. Acids Res. 8, 5795.

Early, T.A., Kearns, D.R., Hillen, W. and Wells, R.D. (1981a) Biochemistry 20, 3756.

Early, T.A., Kearns, D.R., Hillen, W. and Wells, R.D. (1981b) Biochemistry 20, 3764.

Early, T.A., Olmstead, J., Kearns, D.R. and Lezuis, A.G. (1978) Nucl. Acids Res. 5, 1955.

Eckstein, F. and Rizk, I. (1967a) Angew. Chem. Inter. Edit. 6, 695.

Eckstein, F. and Rizk, I. (1967b) Angew. Chem. Inter. Edit. 6, 949.

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

England, T.E. and Neilson, T. (1977) Can. J. Chem. 55, 365.

Evans, F.E., Lee, C. and Sarma, R.H. (1975) Biochem. Biophys. Res. Comm. 63, 106.

Evans, F.E. and Sarma, R.H. (1976) Nature 263, 567.

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

Ezra, E.S., Kondo, M.S., Ainsworth, C.F. and Danyluk, S.S. (1976) Nucl. Acids Res. 3, 2549.

Ezra, F.S., Lee, C.-H., Kondo, N.S., Danyluk, S.S. and Sarma, R.H. (1977) Biochemistry 16, 1977.

Fiers, W., Contreras, R., Dueriuck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molesman, F., Raeymaekers, A., van den Berghe, A., Volckaert, G. and Ysebaert, M. (1976) Nature 260, 500.

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Merregaert, J., Min Jou, W., Raeymaekers, A., Volckaert, G., Ysebaert, M., van de Kerckhove, J., Molf, F. and van Montagu, M. (1975) Nature 256, 273.

Fink, T.R. and Crothers, D.M. (1972) J. Mol. Biol. 66, 1.

Frier, S.M., Hill, K.O., Dewey, T.G., Marky, L.A., Breslauer, K.J. and Turner, D.H. (1981) Biochemistry 20, 1419.

Fuller, W. and Hodgson, A. (1967) Nature 215, 817.

Ganoza, M.C., Fraser, A.R. and Neilson, T. (1978) Biochemistry 17, 2769.

Gadian, D.G., Radda, G.K., Richards, R.E. and Seely, P.L. (1979) in "Biological Applications of Mag. Reson". Shulman, R.G., ed., New York, NY., Academic Press, p. 463.

Geissner-Prettre, C., Pullman, B. Borer, P.N., Kan, S.-L. and Ts'O, P.O.P. (1976) Biopolymers 15, 2277.

Geissner-Prettre, C. and Pullman, B. (1976) Biochem. Biophys. Res. Comm. 70, 578.

Gennis, R.B. and Cantor, C.R.(1970) Biochemistry 9, 4714.

Ghosh, K. and Ghosh, H.P. (1970) Biochem. Biophys. Res. Comm. 40, 135.

Gralla, J. and Crothers, D.M. (1973a) J. Mol. Biol. 73, 497.

Gralla, J. and Crothers, D.M. (1973b) J. Mol. Biol. 78, 301.

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

Gray, D.M. and Ratcliff, R.L. (1977) Biopolymers 16, 1331.

Gray, D.M., Lui, J.-J., Ratcliff, R.L. and Allen, F.S. (1981) Biopolymers 20, 1337.

Gregoire, R.J. and Neilson, T. (1978) Can. J. Chem. 56, 487.

Griffin, B.E., Jarman, M., Reese, C.B., Sulston, J.E. and Trentham, D.R. (1966) Biochemistry 5, 3638.

Grosjean, H., Soll, D.G. and Crothers, D.M. (1976) J. Mol. Biol. 103, 499.

Grosjean, H.J., Dettenan, S. and Crothers, D.M. (1978) Proc. Natl. Acad. Sci. USA 75, 610.

Gross, H.J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H. and Sanger, H.L. (1978) Nature 273, 203.

Gupta, G. and Sasisekharan, Y. (1978a) Nucl. Acids Res. 5, 1639.

Gupta, G. and Sasisekharan, Y. (1978b) Nucl. Acids Res. 5, 1655.

Haasnoot, C,A,G., den Hartog, J.H.J., de Rooij, J.F.M., van Boon, J.H. and Altina, C. (1980) Nucl. Acids Res. 8, 169.

Haschmeyer, A.E.V. and Sobell, H.M. (1963) Proc. Natl. Acad. Sci. USA 50, 872.

Hilbers, C.W. and Shulman, R.G. (1974) Proc. Natl. Acad. Sci. USA 71, 3239.

Hilbers, C.W. and Patel, D.J. (1975) Biochemistry 14, 2656.

Hingerty, B. and Broyde, S. (1978) Nucl. Acids Res. 5, 127.

Holley, R.W., Apgar, J., Everett, G.A., Madison, J.T., Maquisse, M., Merrill, S.H., Penwich, J.R. and Zamii, A. (1965) Science 147, 1462.

Hughes, D.W., Bell, R.A., England, T.E. and Neilson, T. (1978) Can. J. Chem. 56, 2243.

Hurd, R.E. and Reid, B.R. (1979a) Biochemistry 18, 4005.

Hurd, R.E. and Reid, B.R. (1979b) Biochemistry 18, 4017.

Hurd, R.E. and Reid, B.R. (1980) J. Mol. Biol. 142, 181.

Imoto, T., Shibata, S., Akasaka, K. and Hatano, H. (1977) Biopolymers 16, 2705.

Jaskunas, S.R., Cantor, C.R. and Tinoco, Jr., I. (1968) Biochemistry 7, 3164.

Johnston, P. D. and Redfield, A.G. (1977) Nucl. Acids Res. 4, 3599.

Johnston, P.D. and Redfield, A.G. (1978) Nucl. Acids Res. 5, 3913.

Johnston, P.D., Figueroa, N. and Redfield, A.G. (1979) Proc. Natl. Acad. Sci. USA 76, 3130.

Johnston, P.D. and Redfield, A.G. (1981) Biochemistry 20, 1147.

Jones, C.R. and Kearns, D.R. (1975) Biochemistry 14, 2660.

Kallenback, N.R. and Berman, H.M. (1977) Quart. Rev. Biophys. 10, 138.

Kan, L.S., Barrett, J.C. and Ts'O, P.O.P. (1973) Biopolymers 12, 2409.

Kan, L.S., Borer, P.N. and Ts'O, P.O.P. (1975) Biochemistry 14, 4864.

Kastrup, R.V., Young, M.A. and Krugh, T.R. (1978) Biochemistry 17, 4855.

Kataguri, M., Ikatura, K. and Narang, S.A. (1974) J. Am. Chem. Soc. Comm. 355.

Katz, L. and Penman, S. (1966) J. Mol. Biol. 15, 220.

Kearns, D.R., Patel, D.J. and Shulman, R.G. (1971) Nature 229, 338.

Kearns, D.R. and Shulman, R.G. (1974) Acc. Chem. Res. 7, 33.

Kearns, D.R. (1976) Prog. Nucl. Acids Res. and Mol. Biol. 18, 91.

Kim, S.-H., Suddath, F.L., Quigley, C., McPherson, A., Sussman, S.L., Wang, A.H.J., Seeman, N.C. and Rich, A. (1974) Science 185, 435.

Kondo, N.S., Holmes, H.M., Stempel, L.M. and Ts'O, P.O.P. (1970) Biochemistry 9, 3479.

Kroon, P.A., Kreishman, G.P., Nelson, J.H. and Chan, S.I. (1974) Biopolymers 13, 2571.

Krol, A., Machett, M.A., Braulant, C. and Ebel, J.P. (1978) Nucl. Acids Res. 5, 4933.

Krugh, T.R., Laing, J.W. and Young, M.A. (1976) Biochemistry 15, 1224.

Krugh, T.R. and Young, M.R. (1975) Biochem. Biophys. Res. Comm. 62, 1025.

Ladner, J.E., Jack, A., Robertus, J.D., Brown, R.S., Rhodes, D., Clarke, B.F.C. and Klug, A. (1975) Proc. Natl. Acad. Sci. USA 72, 4414.

Lake, J.A. (1977) Proc. Natl. Acad. Sci. USA 74, 1903.

La Mar, G.N. (1979) in "Biological Applications of Magnetic Resonance." Shulman, R.G., ed., New York, NY, Academic Press, p. 305.

Lee, C.-H. and Tinoco, Jr., I. (1978) Nature 274, 609.

Lee, C.-H. and Tinoco, Jr., I. (1980) Biophysical Chem. 11, 283.

Lee, C.-H., Ezra, F.S., Kondo, N.S. and Sarma, R.H. (1976) Biochemistry 15, 3627.

Lerner, D.N. and Kearns, D.R. (1981) Biopolymers 20, 803.

Letsinger, R.L. and Mahadevan, V. (1965) J. Am. Chem. Soc. 87, 3526.

Letsinger, R.L. and Mahadevan, V. (1966) J. Am. Chem. Soc. 88, 5319.

Letsinger, R.L. and Ogilvie, K.K. (1967) J. Am. Chem. Soc. 89, 4801.

Letsinger, R.L. and Ogilvie, K.K. (1969) J. Am. Chem. Soc. 91, 3350.

Lezius, A.G. and Domin, E. (1973) Nat. New Biol. 244, 169.

Lohrmann, R., Soell, D., Hayatsu, H., Ohtsuka, E. and Khorana, H.G. (1966) J. Am. Chem. Soc. 88, 819.

Lomant, A.J. and Fresco, J.R. (1972) J. Mol. Biol. 66, 49.

Lomant, A.J. and Fresco, J.R. (1975) Prog. Nucl. Acids Res. and Mol. Biol. 15, 185.

Martin, F.H., Uhlenbeck, O.C. and Doty, P. (1971) J. Mol. Biol. 57, 201.

Michelson, A.M. and Todd, A.R. (1955) J. Chem. Soc. 2632.

Milman, G., Langbridge, R. and Chamberlin, M.J. (1967) Proc. Natl. Acad. Sci. USA 57, 1804.

Mizuno, H. and Sundaralingan, M. (1975) Nucl. Acids Res. 5, 4451.

Mizuno, H., Tomita, K. and Nakagawa, E. (1981) J. Mol. Biol. 148, 103.

Neilson, T. (1969) Chem. Comm. 1139.

Neilson, T. and Werstiuk, E.S. (1971a) Can. J. Chem. 49, 493.

Neilson, T. and Werstiuk, E.S. (1971b) Can. J. Chem. 49, 3004.

Neilson, T., Wastrodowski, F.V. and Werstiuk, E.S. (1973) Can. J. Chem. 51, 1068.

Neilson, T. and Werstiuk, E.S. (1974) J. Am. Chem. Soc. 96, 2295.

Neilson, T., Deugau, K.V., England, T.E. and Werstiuk, E.S. (1975) Can. J. Chem. 53, 1093.

Neilson, T., Romaniuk, P.J., Alkema, D., Hughes, D.W., Everett, J.R. and

Bell, R.E. (1980) Nucl. Acids Res. Sym. Series No. 7., p. 283.

Newmann, J.M., Guschlbauer, W. and Tran-Dinh, S. (1979) Eur. J. Biochem. 100, 141.

Ninio, J. (1979) Biochemie 16, 1133.

Noller, H.F. and Woese, C.R. (1981) Science 212, 403.

Ogilvie, K.K., Beaucage, S.L. and Entwistle, D.W. (1976) Tetrahedron Letters, 1255.

Olson, W.K. (1978) Biopolymers 17, 1015.

Olson, W.K. (1980) in "Nucleic Acid Geometry and Dynamics". Sarma, R.H., ed., Elmsford, NY., Pergamon Press, p. 383.

Patel, D.J. and Hilbers, C.W. (1975) Biochemistry 14, 2651.

Patel, D.J. (1976) Biophys. Biochem. Acta 442, 98.

Patel, D.J. and Canuel, L.L. (1976) Proc. Natl. Acad. Sci. USA 73, 674.

Patel, D.J. and Canuel, L.L. (1977) Proc. Natl. Acad. Sci. USA 74, 2624.

Patel, D.J. (1979) Biopolymers 18, 553.

Patel, D.J. and Canuel, L.L. (1979) Proc. Natl. Acad. Sci. USA 76, 24.

Patel, D.J. (1980) in "Nucleic Acid Geometry and Dynamics". Sarma, R.H., ed., Elmsford, NY., Pergamon Press, p. 185.

Pattabiraman, N., Rao, S.N. and Sasisekharan, V. (1980) Nature 284, 187.

Pohl, F.M. (1974) FEBS Letters 38, 202.

Porschke, D., Uhlenbeck, O.C. and Martin, F.H. (1973) Biopolymers 12, 1313.

Porschke, D. (1976) Biochemistry 15, 1495.

Quigley, G.J., Seeman, N.C. and Wang, A.H.-J. (1975) Nucl. Acids Res. 2, 2329.

Ravetch, J., Gralla, J. and Crothers, D.M. (1974) Nucl. Acids Res. 1, 109.

Reese, C.B. and Trentham, D.R. (1965) Tetrahedron Letters, 2467.

Reese. C.B. (1978) Tetrahedron 34, 3143.

Reid, B.R. (1981) Ann. Rev. Biochem. 50, 969.

Reid, B.R., McCollum, L., Ribeiro, M.S., Abbak, J. and Hurd, R.E. (1979) Biochemistry 18, 3996.

Rich, A. (1977) Acc. Chem. Res. 10, 385.

Rich, A., Quigley, G.J. and Wang, A.H.-J. (1980) in "Nucleic Acid Geometry and Dynamics". Sarma, R.H., ed., Elmsford, NY., Pergamon Press, p. 273.

Robertus, J.D., Ladner, J.E., Finch, J.T., Rhodes, D., Brown, R.S., Clarke, B.F.C. and Klug, A. (1974) Nature 250, 546.

Robillard, G.T., Hilbers, C.W., Reid, B.R., Gangloff, J., Dirheimer, G. and Shulman, R.G. (1976) Biochemistry 15, 1883.

Roe, B., Michael, M. and Dudock, D. (1973) Nat. New Biol. 246, 135.

Romaniuk, P.J., Neilson, T., Hughes, D.W. and Bell, R.A. (1978a) Can. J. Chem. 56, 2249.

Romaniuk, P.J., Hughes, D.W., Gregoire, R.J., Neilson, T. and Bell, R.A. (1978b) J. Am. Chem. Soc. 100, 3971.

Romaniuk, P.J., Hughes, D.W., Gregoire, R.J., Bell, R.A. and Neilson, T. (1979) Biochemistry 18, 5109.

Romaniuk, P.J. (1979) Ph.D. Thesis, McMaster University, Hamilton, Ont.

Rosenberg, J.M., Seeman, N.C., Kim, J.J.P., Suddath, F.L., Nicohlos, H.B. and Rich, A. (1973) Nature 243, 150.

Rubin, J., Brennan, T. and Sundaralingam, M. (1972) Biochemistry 11, 3112.

Sakore, T.D., Jain, S.C., Tsai, C. and Sobell, H.M.(1977) Proc.Natl. Acad. Sci. USA 74, 188.

Sarma, R.H. (1980) in "Nucleic Acid Geometry and Dynamics". Sarma, R.H., ed., Elmsford, NY., Pergamon Press, p. 1.

Schaller, H., Weimann, G., Lerch, B. and Khorana, H.G. (1963) J. Am. Chem. Soc. 85, 3821.

Seeman, N.C., Sussman, J.L., Berman, H.M. and Kim, S.H. (1971) Nat. New Biol. 233, 90.

Seeman, N,C., Rosenberg, J.M., Suddath, F.L., Kim, J.J.P. and Rich, A. (1976) J. Mol. Biol. 104, 109.

Seeman, N.C. (1980a) in Nucleic Acid Geometry and Dynamics". Sarma, R.H., ed., Elmsford, NY., Pergamon Press, p. 47.

Seeman, N.C. (1980b) in 'Nucleic Acid Geometry and Dynamics". Sarma, R.H., ed., Elmsford, NY., Pergamon Press, p. 109.

Sherm, B.W.-K. (1977) Ph.D. Thesis, Yale University, New Haven, CT.

Smith, M., Rammler, D.H., Goldberg, I.H. and Khorana, H.G. (1962) J. Am. Chem. Soc. 84, 430.

Smrt, J. and Sorm, F. (1962) Coll. Czech. Chem. Comm. 27, 73.

Stannard, B.S. and Felsenfeld, G. (1975) Biopolymers 14, 299.

Stone, M.P., MeGriff, D. and Borer, P.N. (1981) Biochemistry 20, 3604.

Studnicka, G.M., Rohn, G.M., Cummings, I.W. and Salser, W.A. (1978) Nucl. Acids Res. 5, 3365.

Sundaralingam, M. (1969) Biopolymers 7, 821.

Sundaralingam, M. (1975) in "Structure and Conformation of Nucleic Acid Interactions." Sundaralingam, M. and Rao, S.T., eds., Baltimore, MD., University Park Press, p. 487.

Sussman, J.L., Seeman, N.C., Kim, S.H. and Berman, H.M. (1972) J.Mol. Biol. 66, 403.

Swartz, U., Menzll, H.M. and Gassen, H.G. (1976) Biochemistry 15, 2484.

Swartz, U. and Gassen, H.G. (1977) FEBS Letters 78, 267.

Swurkuusk, J., Alvarez, J., Freue, E. and Biltonen, R. (1977) Biopolymers 16, 2641.

Thiyagarajan, P. and Ponnuswamy, P.K. (1978) Biopolymers 17, 2143.

Thiyagarajan, P. and Ponnuswamy, P.K. (1979) Biopolymers 18, 789.

Tinoco, Jr., I., Uhlenbeck, O.C. and Levine, M.D. (1981) Nature 230, 362.

Tinoco, Jr., I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nat. New Biol. 246, 40.

Tsai, C.-C., Jain, S.C. and Sobell, H.M. (1975) Proc. Natl. Acad. Sci. USA 72, 628.

Ts'O, P.O.P., Kondo, M.S., Schweizer, M.P. and Hollis, D.P. (1969) Biochemistry 8, 997.

Ts'O, P.O.P. (1974) in "Basic Principles in Nucleic Acid Chemistry." Vol. II, Ts'O, P.O.P., ed., New York, NY., Academic Press, p. 305.

Turner, D.H., Petersheim, M., DePrisco-Albergo, D., Dewey, T.G. and Frier, S.M. (1981) in "Biomolecular Stereodynamics." Sarma, R.H., ed., Guilderland, NY., p. 429.

Urgurbil, K., Shulman, R.G. and Brown, T.R. (1979) in "Biological Applications of Magnetic Resonance." Shulman, R.G., ed., New York, NY., Academic Press, p. 537.

Uhlenbeck, O.C., Baller, J. and Doty, P. (1070) Nature 225, 508.

Uhlenbeck, O.C., Martin, F.H. and Doty, P. (1971) J. Mol. Biol. 57, 217.

Uhlenbeck, O.S., Borer, P.N., Dengler, B. and Tinoco, Jr., I. (1973). J. Mol. Biol. 73, 493.

Urbanke, C. and Mass, G. (1978) Nucl. Acids Res. 5, 1551.

van Boom, J.H., Burgers, P.M.J., van Deuran, P.H., Arentzen, R. and Reese, C.B. (1974) Tetrahedron Letters, 3785.

Viswamitra, M.A., Kennard, P., Jones, P.G., Sheldrick, G.M., Salisbury, S., Falvello, L. and Shakked, Z. (1978) Nature 273, 687.

Wang, A.H.-J., Quigley, G.J., Kolpah, F.J., Crawford, J.L., van Boom, J.H., van der Murel, G. and Rich, A. (1979) Nature 282, 680.

Wells, R.D., Klysik, J., Studivant, S.M. and Lason, J. (1981) in "Biomolecular Stereodynamics". Sarma, R.H., ed., Guilderland, NY., Academic Press, p. 76.

Werstiuk, E.S. and Neilson, T. (1972) Can. J. Chem. 50, 1283.

Werstiuk, E.S. and Neilson, T. (1973) Can. J. Chem. 51, 1889.

Werstiuk, E.S. and Neilson, T. (1976) Can. J. Chem. 54, 2689

Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakwak, S. and Dickerson, R.E. (1980) Nature 287, 755.

Woese, C. (1970) Nature 226, 817.

Woo, N.H., Roe, B.A. and Rich, A. (1980) Nature 286, 346.

Yoon, K., Turner, D.H., Tinoco, Jr., I., von der Haar, F. and Cramer, F. (1976) Nucl. Acids Res. 3, 2233.