Tetrahedron Vol. 45, No. 5, pp. 1537 to 1548, 1989
Printed in Great Britain.

0040-4020/89 $3.00+.00

WHY DO ALL LARIAT RNA INTRONS HAVE ADENOSINE AS THE BRANCH POINT NUCLEOTIDE ? CONFORMATIONAL STUDIES ON THE IMPLICATION OF THE BRANCH-POINT MODIFICATION BY GUANINE, URACIL OR CYTOSINE IN THE NATURALLY-OCCURRING BRANCHED TETRANUCLEOTIDE BY ^{1}H- & ^{31}P-NMR SPECTROSCOPY.

G. Remaud, N. Balgobin, C. Glemarec & J. Chattopadhyaya*

Department of Bioorganic Chemistry, Biomedical Centre, University of Uppsala, Box 581, S-75123 Uppsala, Sweden

(*Received in UK* 27 *September* 1988)

Abstract : *Conformational studies with three analogues of the naturally occurring branched tetraribonucleotide* ***4****, where the branch-point was either cytidine* ***1****, uridine* ***2****, or guanosine* ***3****, have been performed using ^{1}H- & ^{31}P-NMR spectroscopy. It emerged that* ***1*** *and* ***2*** *are in loose conformational states at room temperature. The possibility for C and especially U at the branch-point to be bulged out have been proposed to describe the temperature-dependent ^{1}H chemical shifts. On the other hand,* ***3*** *adopts a distorted A-type RNA helix with the distortion at the 3′ terminal U residue instead of the 5′ terminal A residue as in* ***4****. The problem of the choice of a suitable model to understand the conformational requirements in splicing is also discussed.*

The regiospecific scission of introns from the pre-mRNA and ligation of the 3′- and 5′-coding regions (exons) to form a functional RNA for protein biosynthesis is generally called splicing[1, 2]. A branched RNA (lariat) intermediate is found to be formed in Group II and nuclear mRNA splicing[1-11]. These lariats have adenosine at the "branch-point" which is linked *via* a 2′ → 5′ phosphodiester to a guanosine residue, and a 3′ → 5′ phosphodiester to a pyrimidine residue[1-11]. The 5′-terminal residue to the branch-point has been found to be either an uridine or an adenosine moiety. Specific nucleobase requirements of the branch-accepting point for the splicing in eukaryotes have been recently reported[12-14]. Mutation experiments have shown that all four nucleotides can serve as branch-acceptors, but it is adenosine and cytidine residues which are prefered to guanosine and uridine residues[12, 13]. These studies clearly address to an important question if there is any structural, conformational and/or thermodynamic basis for the evolutionary choice of adenosine as the branch-accepting nucleotide in the splicing reaction. And if so, how the conformation of the lariat would be changed when the "branch-point" adenosine is replaced with cytidine, guanosine or uridine? A plausible means to approach these questions is to study the conformation of model lariats with substituted nucleotides (cytidine, guanosine and uridine) at the branch-point by NMR spectroscopy and compare them with the corresponding naturally-occurring branched RNA. As regards the choice of a suitable model "branched" RNA molecules which would be representative of the biologically-occurring lariat structure, we ruled out a simple branched triribonucleotide since it has been shown by ^{1}H-NMR spectroscopy that the conformation of branched triribonucleotides is completly different from the one of naturally occurring branched tetramers[15]. Furthermore no similarities, regarding the conformational state were found between branched trimers with adenosine and with cytidine at the branch-point[16]. We therefore chose the model system of a tetraribonucleotide for our conformational studies, and herein report the conformation of all four possible "branched" RNA structures **1** - **4** with identical substituents on the 2′ → 5′, 3′ → 5′ and 5′ → 3′ termini from the "branch-point" which is either cytidine, uridine, guanosine or naturally-occurring adenosine.

^{1}H-NMR spectroscopy

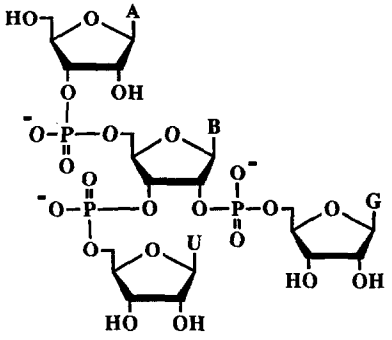
(a) Assignment of the non-exchangeable ^{1}H-NMR signals of the branched tetramers ***1*** *-* ***4*** *and the linear trimer* ***5*** *:* A previous assignment of **1** to **3** has been proposed[17] on the basis of 2D NMR experiments such as DOSE-COSY, DOSE-SECSY and DQFSECSY, allowing the determination of their structures unequivocally. We used, herein, TAYCOSY experiments[15, 18, 19]. Thus, through three relays, subspectra can be obtained showing the J-network from H1′ up to H5′/5″. By comparison with corresponding branched trimers[16] and linear trimers[20-22], ^{1}H resonances were incrementally assigned. Fig.1 shows the

TABLE 1: Temperature-dependent ^1H chemical shifts[a] of aromatic and anomeric protons of **1** - **3** and 5.

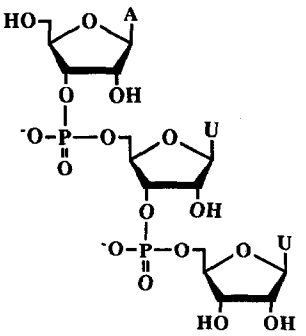
Compound	Fragment	Proton	7 °C	30 °C	80 °C
		H8	8.166	8.203	8.213
1	A3′p	H2	8.003	8.081	8.148
		H1′	5.940	5.980	5.980
		H6	7.640	7.659	7.661
	C*	H5	5.740	5.733	5.790
		H1′	6.055	6.103	6.110
		H6	7.769	7.738	7.730
	p5′U	H5	5.740	5.749	5.812
		H1′	5.824	5.806	5.817
		H8	7.837	7.835	7.842
	p5′G	H1′	5.705	5.706	5.722
		H8	8.179	8.200	8.213
2	A3′p	H2	8.053	8.091	8.150
		H1′	5.988	5.986	5.984
		H6	7.721	7.686	7.659
	U*	H5	5.588	5.595	5.646
		H1′	6.062	6.042	6.026
		H6	7.721	7.726	7.723
	p5′U	H5	5.686	5.730	5.797
		H1′	5.779	5.790	5.808
		H8	7.863	7.857	7.854
	p5′G	H1′	5.700	5.706	5.726

Compound	Fragment	Proton	7 °C	30 °C	80 °C
		H8	8.138	8.148	8.160
3	A3′p	H2	8.071	8.097	8.142
		H1′	5.881	5.857	5.837
		H8	7.815	7.830	7.861
	G*	H1′	5.984	5.971	5.981
		H6	7.829	7.789	7.753
	p5′U	H5	5.797	5.793	5.809
		H1′	5.801	5.836	5.887
		H8	7.707	7.723	7.756
	p5′G	H1′	5.602	5.621	5.663
		H8	8.278	8.263	8.237
5	A3′p	H2	8.118	8.141	8.177
		H1′	6.080	6.002	5.995
		H6	7.714	7.723	7.719
	5′pU3′p	H5	5.570	5.644	5.746
		H1′	5.720	5.770	5.815
		H6	7.742	7.732	7.726
	5′pU	H5	5.691	5.732	5.792
		H1′	5.798	5.804	5.832

[a] CH_3CN as internal reference, set at 2.000 ppm.



1: B = C
2: B = U
3: B = G
4: B = A



5

TABLE 2: Percentage of N pseudorotamers[a] as a function of temperature for **1** - **5**.

Compound	A3′p			p3′N*2′p			3′→5′pU			2′→5′pG		
	7 °C	30	80	7 °C	30	80	7 °C	30	80	7 °C	30	80
1	26	26	26	37	23	26	53	50	50	26	17	20
2	20	22	26	31	29	25	48	45	46	32	26	29
3	47	43	48	62	44	34	28	26	35	29	28	28
4	34	40[b]	35	62	44[b]	34	45	43[b]	45	44	38[b]	34
5	53	45	34	50	44	35	53	52	45[c]			

[a] see text, [b] at 40 °C see ref. 15, [c] at 70 °C due to overlap.

TABLE 3: Temperature-dependent ^{31}P chemical shifts[a] and $\Delta\delta^{31}P$ for **1** - **5**.

Compound	Type of phosphate	10 °C	33 °C	62 °C	81 °C	$\Delta\delta^{31}P$
	A3′→5′C*	1.07	1.18	1.28	1.37	0.30
1	C*3′→5′U	0.76	0.86	0.96	1.02	0.26
	C*2′→3′G	0.45	0.62	0.76	0.85	0.40
	A3′→5′U*	1.05	1.14	1.26	1.35	0.30
2	U*3′→5′U	0.74	0.83	0.91	0.97	0.23
	U*2′→5′G	0.64	0.72	0.87	0.95	0.31
	A3′→5′G*	0.71	0.91	1.13	1.27	0.56
3	G*3′→5′U	0.65	0.79	0.93	1.01	0.36
	G*2′→5′G	0.36	0.48	0.63	0.75	0.39
	A3′→5′A*	0.47	0.75	1.03	1.17	0.70
4	A*3′→5′U	0.69	0.82	0.94	1.01	0.32
	A*2′→5′G	0.19	0.37	0.58	0.70	0.51
5	A3′→5′U	0.66	0.84	1.06	1.19	0.53
	U3′→5′U	0.84	0.96	1.13	1.23	0.39

[a] c-AMP as secondary reference.

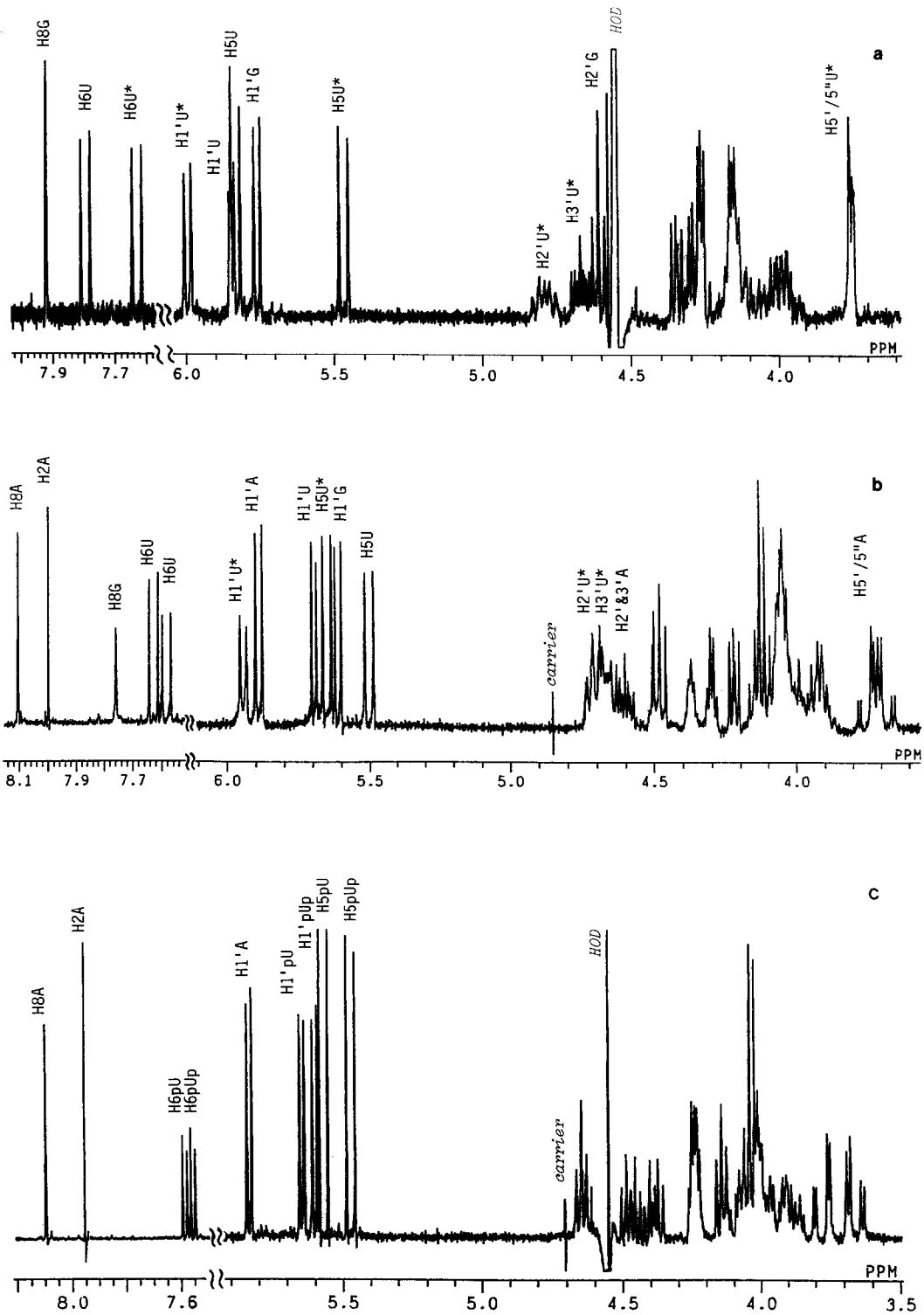


FIGURE 1: 1D 270 MHz ^{1}H-NMR spectra in ^2H_2O of: (a) $U^{*\,2'p5'G}_{3'p5'U}$ (40°C), (b) **2** (35°C) and (c) **5** (30°C).

spectra of $U^{2'p5'G}_{3'p5'U}$ [16], **2** and **5** in order to delineate the substitution effects on U* [* denotes the branch-point]. The connection between the sugar and the base moiety was determined by 2D NOE experiments. The ^1H chemical shifts of aromatic and anomeric protons are shown in table 1. The assignment of **5** was done according to the reports in the literature[20]. *(i) Nucleobase protons* : Compound **1** is constituted with four different nucleobases which possess their own characteristic absorptions: (i) H8 of A [adenine] appears usually at low field, (ii) H8 of G [guanine] is usually broader due to partial exchange with ^2H_2O, (iii) H6 and H5 of pyrimidines are coupled to each other with a coupling constant of ~7.5 Hz for C [cytosine] and ~8.0 Hz for U [uracil] independent of temperature. Interestingly, H6 of the branch-point pyrimidines (C in **1** and U in **2**) are more upfield than H6U of the 3′ → 5′ linkage. On the other hand, H8G* resonates at higher frequencies than H8G in **3**. H8A in **3** has a noticeably higher field shift than in **1** and **2**. These resonances have been assigned through NOE connectivities between the base residue and the sugar moiety. Thus in **3**, H8G (2′ → 5′) shows a clear NOE (Fig. 2) with H1′ G which is, as usual,[15, 16, 18, 23-25] at a high field and morever does not belong to the branch-point G* sugar network (*vide infra*). *(ii) Anomeric protons* : The H1′ of the branch-point residue is always found at the lowest field among the anomeric protons (this is true in branched trimers,[16, 18, 23-25] tetramers,[15] pentamers[26] and heptamer,[26] independent of the nature of the branch-point nucleotide). Fig. 3 shows a 2D Triple Relayed COSY spectrum for **1** which shows that the H1′A is easily detectable since it belongs to the J-network in which the H5′/5′′ signals do not show ^{31}P-^1H couplings and the H1′G is found at the highest field (found also for the naturally occurring branched tetramers[15]). In **1** and **2**, H2′ C* (H2′ U*), H3′ C* (H3′ U*), H2′ A and H3′ A are strongly overlapped at 270 MHz. However, it is clear from TAYCOSY experiments (see Fig. 3 for example) that H3′A is at a higher field than H2′ and H3′ of U* or C*. This is very important in view of assigning ^{31}P chemical shifts (*vide infra*). This overlap does not exist in **3** where H2′ G* and H3′ G* are well separated and strongly shifted to high frequencies[17]. This feature was found in naturally occurring branched tetramers[15] where A constituted the branch-point and also in the corresponding branched trimers[16, 23]. This can be rationalized by a stronger ring-current effect in purines than in pyrimidines[27]. Finally 2D NOE experiments confirmed these assignments by determining the connectivities between the base and the sugar protons. Fig. 4 shows a NOE spectrum of **2** as an example.

(b) Effect of temperature on ^1H chemical shifts in 1 - 3 and 5 : The evolution of ^1H chemical shifts of anomeric and aromatic protons as a function of temperature is a classical technique[28] to detect stackings between nucleobases. Indeed theoretical calculations[27] have shown that among the aromatic protons H2A, H5U and H5C are good probes for detection of stacking since their chemical shifts are solely determined by the ring-current effect of the neighbouring nucleobases. In a stacked form, at a low temperature they experience an upfield shift while in a destacked form at a higher temperature they are shifted to higher frequencies. On the other hand H8A, H8G, H6U, H6C and H1′ (H1′ shows also a dependence on the ring-current effect) have a very strong dependence on the glycosidic bond conformation. Table 1 summarized the variation of ^1H chemical shifts of anomeric and aromatic protons from 7 ° to 80 °C in **1** - **3** and **5**. The 2′ → 5′ linked H8G in **2** and **3** behave in a similar manner as in the corresponding branched trimers. In **2**, H8G goes downfield by ~0.05 ppm when the temperature decreases from 80 ° to 7 °C while in **3**, the H8G has a downfield shift of ~0.037 ppm (~0.01 ppm in the corresponding branched trimers[16]). On the other hand, H8G in **1** undergoes an upfield shift of ~0.005 ppm while a downfield shift of ~0.03 ppm was noticed in $C^{2'p5'G}_{3'p5'U}$ [16] upon the same temperature variation (80 ° to 7 °C). The 3′ → 5′ linked H5U experiences a larger upfield shift in **1** (~0.072 ppm) and **2** (~0.111 ppm) than in the corresponding branched trimers (~0.028 ppm and ~0.011 ppm, respectively[16]). On the contrary, the H5U in **3** has a smaller upfield shift (~0.012 ppm) than in $G^{2'p5'G}_{3'p5'U}$ [16] (~0.025 ppm) from 80 ° to 7 °C. Such results were also found[15] for **4** compared to $A^{2'p5'G}_{3'p5'U}$ [23]. This clear difference between **1** and **2**, on the one hand, and **3** and **4**, on the other hand, is also reflected in the temperature-dependent chemical shift of H6U. Thus in **3** the H6U has a tendency to go more downfield (~0.076 ppm) than in $G^{2'p5'G}_{3'p5'U}$ [16] (~0.072 ppm), which is also found[15] in **4** (~0.1 ppm against ~0.05 ppm in $A^{2'p5'G}_{3'p5'U}$ [23]). The H6U in **1** experienced a smaller downfield shift of ~0.039 ppm [in **2**, H6U is nearly invariant] than in $C^{2'p5'G}_{3'p5'U}$ and in $U^{2'p5'G}_{3'p5'U}$ (~0.069 and ~0.065 ppm, respectively[16]). In **5** the H5 of 5′pU (3′ C 5′ terminal) moves upfield by ~0.1 ppm while the H6 goes downfield by ~0.016 ppm over a temperature change from 80 ° to 7 °C. The H2 and H8 of the 5′-terminal A in branched tetramers **1** - **4** have a temperature dependency which does not reflect the expected[29] order of the strength of stacking ApA > ApG > ApC > ApU. Indeed, it is **1** which possesses similar shifts as in **4**: H2A and H8A move upfield by ~0.145 ppm and ~0.047 ppm in **1**, while in **4** it was found[15] to be ~0.136 ppm and ~0.053 ppm respectively (from 80 ° to 7 °C). Suprisingly the shifts upon temperature variation in **3** are the smallest among the branched tetramers **1**, **2**, **3** and **4**. A sharp difference also emerged between **2** and **5**. In the latter the H8A moves to high frequencies upon lowering of the temperature (~0.041 ppm) and H2A has an upfield shift of ~0.059 ppm. In **2** not only

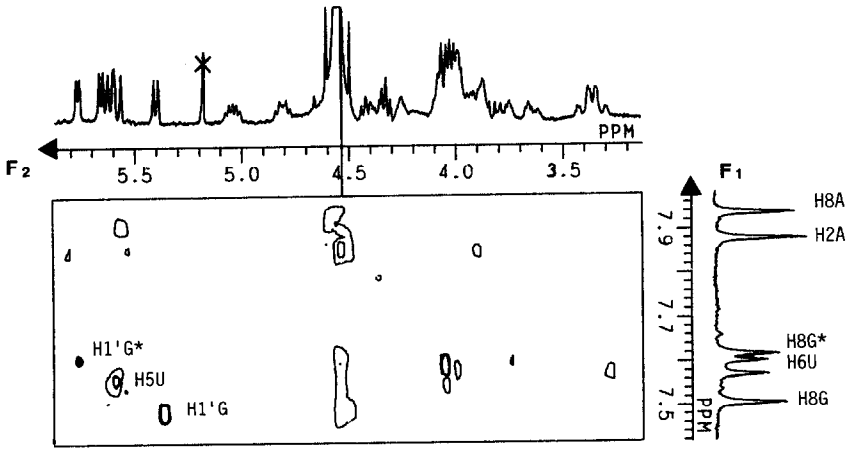


FIGURE 2: NOESY spectrum of **3** (~4mM in ^2H_2O, T = 20°C) showing the connectivities between the aromatic region (right side) and the upfield region (top), x denotes solvent impurities. Experimental conditions: $F_1 = F_2 = 2100$ Hz, resolution 4.1 Hz/pt for a zero-filled matrix up to 512 x 1K, mixing time $\tau_m = 500$ ms, recycle time of 3 s, $\Delta t_1 = 0.476$ ms, 208 scans were recorded for every t_1 value. A sine-bell function was applied in both directions and a symmetrization was done.

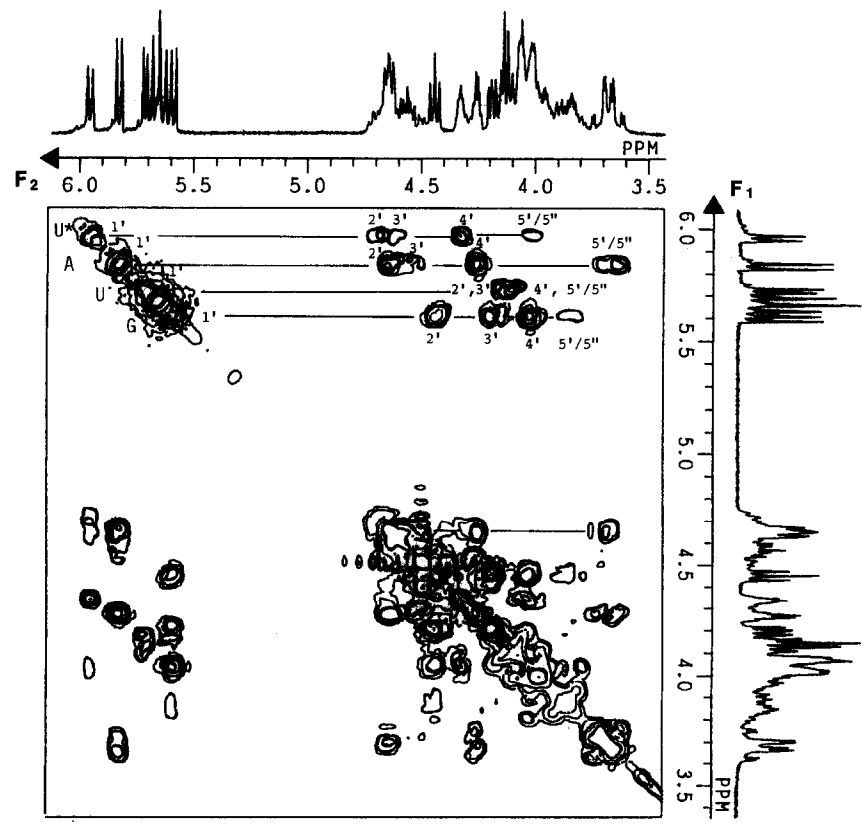


FIGURE 3: TAYCOSY spectrum of **1** (~25 mM in ^2H_2O, T = 35°C). The subspectra for each sugar are shown from H1´. Experimental conditions: $F_1 = F_2 = 2700$ Hz, the relay time was 110 ms, resolution 5.3 Hz/pt. A non-shifted sine-bell function was applied and the spectrum was symmetrized in a 512 x 1K matrix. 96 FIDs were accumulated for each 256 experiment, $\Delta t_1 = 0.370$ ms.

H8A goes upfield when the temperature changes from 80 ° to 7 °C but H2A has a larger upfield shift ~0.1 ppm. These features suggest that **2** possesses a molecular geometry which is different from **5** (A-type stacking[28]). Regarding the branch-point, again **3** has a similar behaviour as **4** which does not resemble the corresponding branched trimers[16]. Thus the H8G* in **3** and H8A* in **4** experience a noticeable upfield shift of ~0.046 ppm and ~0.076 ppm respectively[15] (80 ° → 7 °C). Downfield shifts of ~0.011 ppm in $G^{2'p5'G}_{3'p5'U}$ [16] and ~0.004 ppm in $A^{2'p5'G}_{3'p5'U}$ [23] (80 ° → 7 °C) have been observed earlier. The H5C* in **1** moves upfield by ~0.05 ppm and the H6C* by ~0.021 ppm, while in $C^{2'p5'G}_{3'p5'U}$ [16] the upfield shift of H5C* was ~0.18 ppm and the H6C* moved downfield by ~0.009 ppm (80 °→ 7 °C). The features of **2** are closer to those of $U^{2'p5'G}_{3'p5'U}$ [16] (but to a lesser extent)than in **5**: (i) H5U* has an upfield shift of ~0.058 ppm in **2**, of ~0.193 ppm in $U^{2'p5'G}_{3'p5'U}$ [16] and H5U of ~0.176 ppm in **5**, (ii) H6U* has a downfield shift of ~0.062 ppm in **2** and of ~0.043 ppm in $U^{2'p5'G}_{3'p5'U}$ [16] while H6U has a slight upfield shift of ~0.005 ppm in **5** (80 ° →7 °C).

(c) Population of pseudorotamers N ⇄ S of sugar residues : The ribose ring conformation can be described as an equilibrium mixture of two puckered modes S and N, undergoing rapidinterconversion[30]. An approximate population of sugar pseudorotamers (N or S), at various temperatures has been estimatedfrom $J_{1',2'}$ coupling constants, using the procedure described by Altona and coworkers[31]. The percentage of N pseudorotamers at a given temperature can be calculated from $J^{Exp}_{1',2'} = x_N J^{N}_{1',2'} + (1-x_N) J^{S}_{1',2'}$ where x_N is the molar fraction of N-type conformers, and $J^{Exp}_{1',2'}$ represents the experimental coupling constant between H1′ and H2′. The normal ranges of pseudorotation[31] are defined as P_S = 162 °, f_S = 38 °, P_N = 9 ° and f_N = 38 ° leading to $J^{S}_{1',2'}$ = 7.8 Hz and $J^{N}_{1',2'}$ = 1 Hz (N and S are the subscripts for N- and S-type conformers).

The percentage of N conformers have been calculated for each nucleotide and are shown in table 2. It is clear from table 2 that G* in **3** prefers N conformation at low temperature (~62%). The A* in **4** adapted a similar conformation[15]. On the contrary, C* and U* in **1** and **2** prefer predominantly an S-type even at a low temperature as found in the corresponding branched trimers[16]. Although it should be noted that not only the amount of N-type is higher in **1** and **2** but they also exhibit a larger variation in the decrease of S-type pseudorotamer at a low temperature in **1, 2** and **3**. The 2′ → 5′ G residue also adopts preferentially S-type conformers with a very small variation over a change of temperature (80 ° → 7 °C). A larger amount of N-type existed in **4** especially at a low temperature[15]. On the other hand, the 3′ → 5′ linked U residue has a slight preference of N-type conformer in **1** and **2** which is very close to the features found in **4** and in the corresponding branched trimers. Suprisingly the 3′ → 5′ U residue in **3** displays a large S-type population which increases upon lowering the temperature. The 5′-terminal A in **1** and **2** have a sugar conformation which is biased to the S-type with nearly no change when the temperature decreases. In **4** a similar data was found[15], with however a higher population of the N-type. The 5′-terminal residue in **3** is again singularized by adopting almost 50 % of N conformation at a low temperature (the 5′-terminal A in **4** exists in ~35 % of N conformation). The sugar conformations of **5** are N-type at a low temperature. The population is displaced to S-type when the temperature is increased as in normal right handed stack single helix (A-type).

(d) 2D NOE studies : The connectivities between a nucleobase and its own sugar residue are the only useful data extractable (*vide supra*) in compounds **1, 2** and **3**. This is presumably due to the lack of rigidity in **1, 2** and **3**. In contrast, the naturally-occurring **4** showed inter-residue NOE connectivities which could be conveniently used in a qualitative manner to understand its secondary structure[15]. We believe that **1, 2** or **3** do not possess any strongly preferred conformational state at room temperature.

^{31}P-NMR spectroscopy

(a) Assignment of ^{31}P NMR chemical shifts : We have shown[33] that a correct assignment of ^{31}P NMR resonances of branched oligonucleotides can give an independent view of the overall conformation of such structures. The ^{31}P chemical shifts of **1, 2, 3** are shown in table 3 together with those of **4** and **5** for comparison. The two dimensional ^{31}P/^{1}H correlation is one of the best tools to assign ^{31}P chemical shifts in small oligonucleotides[16]. The 2D ^{31}P/^{1}H correlation spectra of **1, 2** and **3** have been already presented[17] to show the exact location of the 2′ → 5′ phosphate linkage with respect to other 3′ → 5′ phosphate linkages (see Figs. 5 & 6, for example). Indeed it was expected that, as in naturally branched tetramers[15], the A3′p5′ N* (where N= C, U, G in **1, 2, 3**) phosphate resonances should absorb between the (N*3′→ 5′) and (N*2′ → 5′) phosphates. However, a careful analysis of 2D ^{31}P/^{1}H spectra in conjunction with the TAYCOSY spectra (*vide supra*) showed

that A3′→ 5′ N* phosphate absorbed the most downfield amongst all other phosphate resonances. We attributed this to both the ring-current effect of each nucleobase (A > G > C > U) and the conformational state of different branched tetramers. Present temperature-dependent ^{31}P-NMR studies have confirmed earlier observations that the 2′ → 5′ linked phosphate always has the most upfield absorption[16, 33, 34]. The separation of H5′ and H5′′ of the 2′ → 5′ phosphate linked G has been associated[15, 16, 18, 23, 24] with a specific conformation[35, 36] around the 2′ → 5′ phosphate backbone, corresponding to a favoured stacking between the branch-point nucleobase and the 2′ → 5′ linked nucleobase. The 2D ^{31}P/^{1}H correlation spectra in **1** and **2** (Fig. 5) clearly show the absence of this splitting of H5′ and H5′′ at 270 MHz while for **3** (Fig. 6) this splitting exists to a much smaller extent than in the corresponding branch trimer $G^{2'p5'G}_{3'p5'U}$ [16].

(b) ^{31}P chemical shifts temperature-dependence: In previous studies[16, 33] we have shown that the variation of ^{31}P chemical shifts ($\Delta\delta^{31}$P) in branched trimers and tetramers over the same range of temperature (81 ° → 10 °C) can indeed reflect the relative conformational constraint of the different phosphate linkages within the molecule. Such $\Delta\delta^{31}$P have been interpreted[16, 27, 33, 37] as a change in the bond angle and the torsion angle O-P-O, stacking ⇄ destacking equilibria, a change in the torsion angle of the bond C3′-C4′. It is difficult to separate these different contributions to $\Delta\delta^{31}$P. It is commonly accepted[38] that the stability in a single stranded helix, originates from base stacking and the flexibility of the helix comes from the sugar conformation which transfers this flexibility (or rigidity) to the phosphate backbone through the torsion angle around C3′-C4′ bond. Therefore we interpreted[16, 33] changes in $\Delta\delta^{31}$P as being due to a large influence of shifts of N ⇄ S and stack ⇄ destack equilibria. $\Delta\delta^{31}$P values of compounds **1** - **5** are listed in table 3. In the naturally occurring branched tetramer **4**, a decrease of $\Delta\delta^{31}$P of the 2′ → 5′ phosphate linkage has been noticed[33] compared with that in the branched trimer $A^{2'p5'G}_{3'p5'U}$, while the 3′ → 5′ phosphate linkage possesses a steady $\Delta\delta^{31}$P. The latter statement is respected in **1**, but a clear difference can be seen in **2** and especially in **3**. Thus Δd^{31}P in **3** of G*3′p5′ U is much larger than in $G^{2'p5'G}_{3'p5'U}$ [16] , while Dd^{31}P in **2** of U*3′p5′ U is smaller than in $U^{2'p5'G}_{3'p5'U}$ [16]. Although Δd^{31}P of the 2′ → 5′ linkage in **1**, **2**, **3** are smaller than in the corresponding branched trimers[16] (a large difference is found in **2** and **3**, while in **1** this difference is comparable to that in **4**). Interestingly $\Delta\delta^{31}$P of 2′ → 5′ in **1** is the largest amongst the compounds **1**, **2**, and **3**. The $\Delta\delta$ of A3′ → 5′ N* linkage is very small in **1** and **2** (0.30). In a normal A-type stack as found in **5** $\Delta\delta$ ^{31}P has a higher value (0.53). The difference in $\Delta\delta$ of A3′ → 5′ N* in **3** and in **4** can be explained in part by a stronger stacking in ApA than in ApG.

DISCUSSION

Previous studies[15, 16, 18, 23, 24, 33, 39] from this and other laboratories[25, 34] have shown that the overall conformation of branched trimers could be described as a strong preference for the branch-point nucleobase to stack with the 2′ → 5′ nucleobase which dictates its secondary structure. Such a structure yields the following characteristic features: (i) the conformation of the branch-point is biased to a high percentage of S-type conformers, (ii) a strong temperature-dependent ^{1}H chemical shift for the branch-point nucleotide and the 2′ → 5′ nucleobase protons, and (iii) the specific phosphate backbone conformation is in accordance with a 2′ → 5′ stacking. This preferential 2′ → 5′ stacking results in a close proximity of the nucleobase in such structures, and which is favoured by a repulsion between the 2′ → 5′ and 3′ → 5′ phosphates. The stacking N*3′ → 5′ M is therefore much less favoured and sometimes nearly non-existent. The strength of the stacking (2′ → 5′ and 3′ → 5′) however depends on the nature of the constituent nucleobases. The following classification could be established according to the conformational constraint within the molecule delineated on the basis of ^{1}H, ^{31}P-NMR and CD studies:

$$A^{2'p5'U}_{3'p5'G} \geq A^{2'p5'C}_{3'p5'G} > A^{2'p5'G}_{3'p5'U} > A^{2'p5'G}_{3'p5'G} > A^{2'p5'G}_{3'p5'C} > A^{2'p5'G}_{3'p5'U} > A^{2'p5'A}_{3'p5'A} > G^{2'p5'G}_{3'p5'U} \geq G^{2'p5'G}_{3'p5'C} > U^{2'p5'G}_{3'p5'U} \geq C^{2'p5'G}_{3'p5'U} \geq U^{2'p5'G}_{3'p5'C} \geq C^{2'p5'G}_{3'p5'C}$$

These general features do not exist, however, in the naturally occurring branched tetramers[15]. A study of four branched tetramers with A* at the branch-point has shown that their conformational states can be described as a distorted A-type stacking, the distortion coming mainly from the 5′-terminal nucleoside (A or U)[15]. But in any case the A* adapted predominantly N-type conformation, as the U at the 3′-terminus does to a small extent (it is noteworthy that when C is at the 3′ terminus, percentage of N is larger)[15]. Therefore a 3′ → 5′ stacking is preferred, the 5′-terminal nucleoside is oriented away from A*3′p5′ U axis in order to obtain the additional stability by stacking with G at the 2′-terminus. None of branched tetramers in the present study display such clear features. It emerges that each of them possesses its own structure.

Compound 1 : The 5′-terminal A in **1** prefers the stacking with G due to: (i) high percentage of S conformers for A and G, (ii) upfield shift of H8G (80 ° → 7 °C), and (iii) large upfield shifts for H2 and H8 of A. The C* and 3′pU interact through S-N

stackings. This interaction is relatively weak due to a poor capacity for stacking of both U and C. This weakness does not impose any strong constraint on the phosphate backbone C*3′ → 5′ U (small $\Delta\delta^{31}P$). Therefore other stable conformations could be found (especially at room termperature or above) such as the conformation mimicking the branched trimer $C^{2'p5'G}_{3'p5'U}$ or a conformation similar to a "linear stacking" along the axis A3′p5′C*2′p5′G. The NMR time average conformation of **1** is probably a blend of these three conformations, although we believe that the first one is predominant, at least at a low temperature. In any case, these conformations involve the 2′ → 5′ residue to a large extent, exerting an unexpected large conformational constraint on the 2′ → 5′ phosphate linkage ($\Delta\delta^{31}P = 0.40$).

Compound 2 : An examination of $\Delta\delta^{31}P$ of **2** reveals that no strong constraint exists in **2**. The conformation of **2** could be described by three components of almost equal importance at room temperature: (i) as in **1**, a stacking interaction between 5′-terminal and 2′pG is counterbalanced by an interaction between U* and 3′pU, (ii) a possible interaction exists between 5′-terminal A and 3′pU while U* is bulged out to interact with G and (iii) S-S stacking mode is prevalent between 5′-terminal A and U* but weaker than in **5**, such a B-type short helix is further stabilized by a weak interaction with G. In such a case 3′pU can also undergo an interaction with the 5′-terminal A. The large shielding effect of H5 of 3′pU and the relativly smaller shielding of H5U* are in agreement with a bulge out of U*. Such behaviour is well described in the literature but it occurs, however, only in the sequence Pu-U-Pu[32, 40]. In **2** it is likely that this possible conformation is favoured (even weakly) due to the repulsive interaction between the 2′- and the 3′-phosphates.

Compound 3 : It showed a similar overall conformation as in **4** due to a similar sequence $Pu3'p5'Pu^{2'p5'Pu}_{3'p5'Py}$ a clear difference can be seen However, some sharp differences could be pointed out: (i) $\Delta\delta^{31}P$ indicate that A3′ → 5′ G* and G*2′ → 5′ G phosphate backbones are less constrained than in **4** while G*3′ → 5′ U in **3** is significantly more constrained than A*3′ → 5′ U in **4**, (ii) The 5′-terminal A has a much larger percentage of N conformers in **3** than in **4** and (iii) conversely, 3′-terminal U exhibits an S-type population which is much larger in **3** than in **4**. To rationalize these differences it is likely that the A-type stacked mini helix along the A3′p5′ G*3′p5′ U axis is distorted at the 3′pU level in **3** instead of the 5′-terminal level in **4**. In this hypothetical scheme the main axis of stabilization is A3′p5′ G*2′p5′ G while the 3′pU can interact weakly with A and it is nearly coplanar with G (table 1). This is supported by a clear deshielding of H5 and H6 of 3′pU in **3** compared with **1**, **2** and **4** due to the paramagnetic effect of G. These factors increase the constraint of G*3′ → 5′ U (large $\Delta\delta^{31}P$). The presence of G*2′ → 5′ G stacking has already been pointed out when describing the splitting of H5′/5′′ of 2′ → 5′G in **3** (*vide supra*).

CONCLUDING REMARKS

The absence of noticeable NOEs between nucleobases indicates a loose conformational state for **3** and especially for **1** and **2** as in the corresponding branched trimers[16]. The conformations of **1** and **2** do not resemble an A-type stack helix (as in compound **5**), but possibly the branch point U* or C* could be bulged out. It also emerged that the conformation of such branched tetramers is strongly dependent on the constituent nucleobases. Recently we reported[16] that the branch core triribonucleotides are not suitable models to understand the structural require-ments of the splicing intron. Indeed large conformational differences between branched trimers and tetramers could be observed. The branched trimers, independent of the nucleotide at the branch-point, adopt a conformation biased to the 2′ → 5′ axis leading to an unusually high population of S-type conformers for the sugar residue at the branch-point nucleotide. The strength of the 2′ → 5′ interaction is weak when C or U constitutes the branch-point. An addition of a nucleotide at the 5′ terminal perturbs these conformational features considerably. Thus with A, or to a lesser extent G, at the branch point a 3′ → 5′ stabilization is achieved which is however distorted at the 5′ terminus for A and at the 3′ terminus for G. Finally, enzymes recognized branched tetrameters where A was the branch-point as a substrate in enzymatic 5′-phosphorylation reaction but not the corresponding branched trimers[41, 42]. Splicing is achieved with A, and to a lesser extent with C, at the branch-point in *in vitro* experiments . However, the present work does not reveal any strong similarities between branched tetramers with A or with C at the branch-point . It is however noteworthy that in **1** and **4** the conformational constraint along the 2′ → 5′ phosphate linkage is strongest amongst the studied branched tetramers (in **3** a similar constraint is found but it is competing with the one along the 3′ → 5′ phosphate linkage). This common structural feature in **1** and **4** may serve as a driving force and a recognition signal for the second step in splicing. The current study clearly addresses that the length of oligonucleotide residues at the 3′ and 5′ termini is important for the stabilization owing to the stacking of the central core branched structure with the neighbouring residues. It becomes obvious that by the addition of nucleotides at the 3′ and 5′ termini would promote the stabilization of the 3′ → 5′ mini helix which in turn would promote the spatial arrangements of the neighbouring nucleotides. For these reasons work on the conformation of a branched heptameter (two nucleotides at the 2′, 3′ and 5′ termini[26]) is in progress in this laboratory.

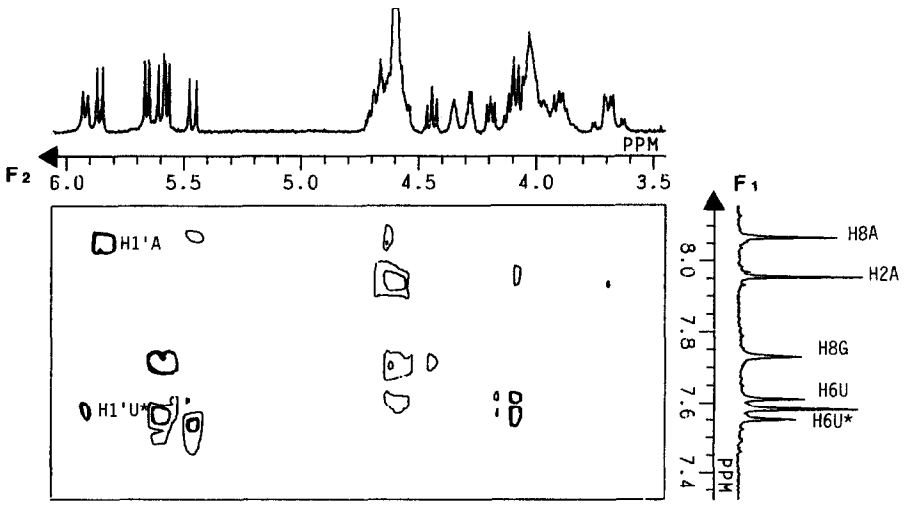


FIGURE 4: NOESY spectrum of **2** (~5 mM in $^2H_20$, T = 23°C). NOEs connectivities are shown between the aromatic part (right side) and the upfield region (top). Experimental conditions: $F_1 = F_2 = 2100$ Hz, resolution 4.1 Hz/pt for a zero-filled matrix (512 x 1K), mixing time $\tau_m = 500$ ms, recycle time of 3 s, $\Delta t_1 = 0.476$ ms, 176 scans were recorded for every t_1 value. A sine-bell function was applied in both directions and a symmetrization was done.

MATERIALS AND METHODS

Samples: The title compounds **1** - **3** were prepared according to a new synthetic method already reported[17].The synthesis of **5** was accomplished following standart phosphotriester methods[43]. The samples were treated with Dowex Na^+ exchange resin and lyophilized to dryness. *^1H NMR*: The samples were coevaporated twice in small amounts of 99.8% $^2H_20$. Finally the samples were dissolved in 0.5 ml $^2H_20$ "100 atom% D" (Aldrich). The concentration for the temperature study was < 5 mM. The actual concentrations used for 2D NMR are notified for each spectrum. For all the experiments the pH was found to be ≈ 7. The spectra are referenced to the acetonitrite signal which was added as an internal reference (set at 2.000 ppm).The triple relay COSY experiment (TAYCOSY) has already been described[15, 18, 19], wherein we used a version with 32 phase cycle steps instead of 16[19]. Thus the quadrature detection is better achieved. The 2D NOE experiments have been performed using the basic NOESY pulse sequence[44] where the mixing time t_m was systematically changed as t_1 increments with a maximum change of $t_m/10$[45].*^{31}P-NMR*: The $^{31}P/^1H$ chemical shift correlations spectra were recorded as already reported[16, 33]. The ^1H decoupled ^{31}P-NMR spectra were recorded at 109.4 MHz using 10 mm o.d. tubes. The ^{31}P chemical shifts are referenced to 10% H_3PO_4 in $^2H_20$ in a capillary. The samples were dissolved in H_20 to give a concentration below 5 mM. The 3′,5′-cAMP was used as secondary reference to probe both the bulk susceptibility and the hydration phenomena[16, 33, 46]. The NMR experiments described in this work were performed on a Jeol JNM-GX 270 spectrometer.

ACKNOWLEDGEMENTS

Authors thank the Wallenbergs Stiftelsen for funds for the purchase of a 270 MHz NMR spectrometer, the Swedish Board for Technical Development and the Swedish Natural Science Research Council for generous research grants and Caroline Martin for excellent secretarial assistance.

REFERENCES

1. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. and Sharp, P. A (1986) *Ann. Rev. Biochem.* **55**, 1119-1150.
2. Cech, T. R. and Bass, B. L. (1986) *Ann. Rev. Biochem.* **55**, 599-629.
3. Konarska, M. M., Grabowski, P. J. , Padgett, R. A. and Sharp, P. A. (1985) *Nature* **313**, 552-557.
4. Ruskin, B., Krainer, A. R., Maniatis, T. and Green, M. R. (1984) *Cell* **38**, 317-331.
5. Ruskin, B. and Green, M. R. (1985) *Science* **229**, 135-140.
6. Wallace, J. C. and Edmonds, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 950-954.
7. Sharp, P. A. (1987) *Science*, **235**, 766-771.

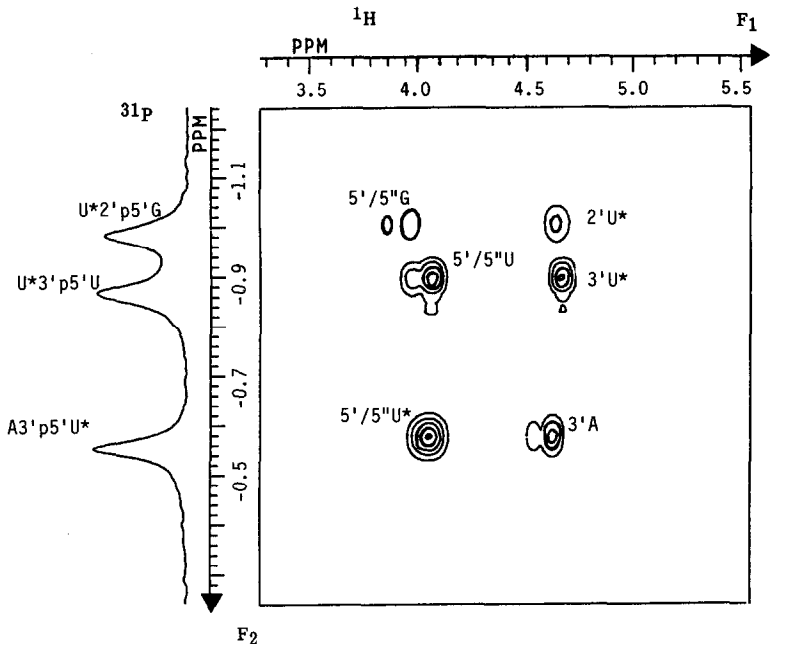


FIGURE 5: ^{31}P/^{1}H chemical shift correlation spectrum of **2** (~9 mM in ^2H_2O, T = 35°C). The ^{31}P spectrum is displayed on the left side. Experimental parameters F_2 (^{31}P dimension) = 500 Hz, F_2 (^{1}H dimension) = 2200 Hz, a zero-filled to 1024 x 0.5K matrix yielded a resolution of 2 Hz/pt for F_2. 96 FIDs were accumulated for the 128 t_1 incre-ments (0.227 ms), Δ_1 = 58 ms and Δ_2 = 29 ms. An exponential function was used in both directions. ^{31}P chemical shifts are referenced against 85 % H_3PO_4.

8. Ruskin, B., Pikielny, C. W., Rosbash, M. and Green, M. R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2022-2026.
9. van der Veen, R., Arnberg, A., van der Horst, G., Bonen, L., Tabak, H. and Grivell, L. (1986) *Cell* **44**, 225-234.
10. Peebles, C., Perlmann, P., Meckenburg, K., Petrillo, M., Tabor, J., Jarrell, K. and Cheng, H. (1986) *Cell* **44**, 213-223.
11. Jarrell, K., Peebles, C., Dietrich, R., Romiti, S. and Perlmann, P. (1988) *J. Biol. Chem.* **263**, 3432-3439.
12. Aebi, M., Horning, H., Padgett, R. A., Reiser, J. and Weissmann, C. (1986) *Cell* **47**, 555-565.
13. Horning, H., Aebi, M. and Weissmann, C. (1986) *Nature* **324**, 589-591.
14. Freyer, G. A., Arenas, J., Perkins, K. K., Furneaux, H. M., Pick, L., Young, D., Roberts, R. J. and Huinitz, J. (1987) *J. Biol. Chem.* **262**, 4267-4273.
15. Zhou, X.-X., Nyilas, A., Remaud, G. and Chattopadhyaya, J. (1988) *Tetrahedron* **44**, 571-589.
16. Remaud, G., Balgobin, N., Sandström, A., Koole, L. H., Drake, A., Zhou, X.-X. and Chattopadhyaya, J. *J. Biochem. Biophys. Methods*, in press. .
17. Balgobin, N., Földesi, A., Remaud, G. and Chattopadhyaya, J. *Tetrahedron*, in press.
18. Remaud, G., Zhou, X.-X., Öberg, B. and Chattopadhyaya, J. (1988) in: *Reviews of Heteroatom Chemistry* (Oae, S. ed.) MYU Publishing Inc., Tokyo.Vol. 1, 340-366.
19. Remaud, G., *J. Biochem. Biophys. Methods*, in press.
20. Lee, C.-H. and Tinocco, I. (1980) *Biophys. Chem.* **11**, 283-294.
21. Everett, J. R., Hughes, D. W., Bell, R. A., Alkema, D., Neilson, T. and Romaniuk, P. J. (1980) *Biopolymers* **19**, 557-573.
22. Stone, M. P., Johnson, D. L. and Borer, P. N. (1981) *Biochemistry* **20**, 3604-3610.
23. Remaud, G., Vial, J.-M., Balgobin, N. and Chattopadhyaya, J. (1987) *Tetrahedron* **43**, 947-958.
24. Vial, J.-M., Remaud, G., Balgobin, N. and Chattopadhyaya, J. (1987) *Tetrahedron* **43**, 3997-4006.
25. Lee, M., Huss, S., Gosselin, G., Imbach, J.-L., Hartley, J. A. and Lown, J. W. (1987) *J. Biomol. Struct. Dyn.* **5**, 651-668.
26. Zhou, X.-X., Remaud, G. and Chattopadhyaya, J. *Tetrahedron*, in press.
27. Giessner-Prettre, C. and Pullman, B. (1987) *Quart. Rev. Biophys.* **14**, 113-172.
28. Altona, C. (1982) *Recl. Trav. Chim. Pays-Bas* **101**, 413-433.
29. Saenger, W. (1984) in *Principles of Nucleic Acid Structure*. Springer-Verlag, New York Inc.
30. Altona, C. and Sundaralingam, M. (1972) *J. Am. Chem. Soc.* **94**, 8205-8212 and *ibid* (1973) *J. Am. Chem. Soc.* **95**, 2333-2344.

31a. Haasnoot, C. A. G., de Leeuw, F. A. A. M. and Altona, C. (1980) *Tetrahedron* **36**, 2783-2792.

31b. De Leeuw, F. A. A. M. and Altona, C. (1982) *J. Chem. Soc. Perkin II*, 375-384.

32. van den Hoogen, Y. Th., Erkelens, C., de Vroom, E., van der Marel, G. A., van Boom, J. H. and Altona, C. (1988) *Euro. J. Biochem.* **173**, 295-303.
33. Sandström, A., Remaud, G., Vial, J.-M., Zhou, X.-X., Nyilas, A., Balgobin, N. and Chattopadhyaya, J. (1988) *J. Chem. Soc., Chem. Commun.* , 542-544.

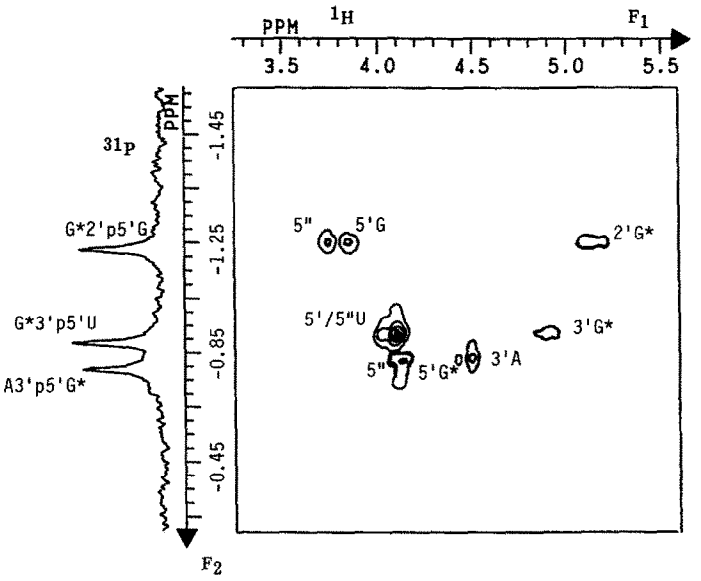


FIGURE 6: $^{31}P/^1H$ chemical shift correlation spectrum of **3** (~6 mM in ^2H_2O, T = 30°C).The ^{31}P spectrum is displayed on the left side. The spectrum has been obtained using the same parameters as in Fig. 5. ^{31}P chemical shifts are referenced against 85 % H_3PO_4.

34. Damka, M. J. and Ogilvie, K. K. (1988) *J. Org. Chem.* **53**, 3710-3722.
35. Srinivasan, A. R. and Olson, W. K. (1986) *Nucleic Acids Res.* **14**, 5461-5479.
36. Anukanth, A. and Ponnuswamy, P. K. (1986) *Biopolymers* **25**, 729-752.
37. Gorenstein, D. G. (1981) *Ann. Rev. Biophys. Bioeng.* **10**, 355-386.
38. Sundaralingam, M. and Westhof, E. (1981) *Biomol. Stereodyn. Proc. Symp.* **1**, 301-326.
39. Koole, L. H., Balgobin, N., Buck, H. M., Kuijpers, W. H. A., Nyilas, A., Remaud, G., Vial, J.-M. and Chattopadhyaya, J. *Recl. Trav. Chim. Pays-Bas*, in press.
40. van den Hoogen, Y. Th., van Beuzekom, A. A., de Vroom, E., van der Marel, G. A., van Boom, J. H. and Altona, C. (1988) *Nucleic Acids Res.* **16**, 5013-5030.
41. Edmonds, M., personal communication on July 7th, 1987.
42. Kierzek, R., Kopp, D. W., Edmonds, M. and Caruthers, M. H. (1986) *Nucleic Acids Res.* **14**, 4751-4764.
43. Reese, C. B. (1978) *Tetrahedron* **34**, 3143-3179.
44. Morris, G. A. (1986) *Magn. Reson. Chem.* **24**, 371-403.
45. Wider, G., Macura, S., Kumar, A., Ernst, R. R. and Wütrich, K. (1984) *J. Magn. Reson.* **56**, 207-234.
46. Haasnoot, C. A. G. and Altona, C. (1979) *Nucleic Acids Res.* **6**, 1135-1149.