Tetrahedron Vol. 47, No. 18/19, pp. 3183-3206, 1991
Printed in Great Britain

0040-4020/91 $3.00+.00

SOLUTION STRUCTURE OF PENTAMERIC AND HEPTAMERIC BRANCHED-RNA MODELLING THE LARIAT STRUCTURE OF GROUP II OR NUCLEAR m-RNA INTRONS STUDIED BY ONE- AND TWO-DIMENSIONAL NMR SPECTROSCOPY AT 500 MHz

L. H. Koole, P. Agback, C. Glemarec, X.-X. Zhou and J. Chattopadhyaya*

Department of Bioorganic Chemistry, Box 581, Biomedical Center,
University of Uppsala, S-751 23 Uppsala, Sweden

(*Received in UK* 9 *January* 1991)

Abstract : A detailed 500 MHz ^1H NMR study of the branched RNA systems 3 (pentamer) and 4 (heptamer) is reported. The conformational properties of 3 and 4 were compared with those of the previously studied smaller branched RNAs 1 (trimer) and 2 (tetramer). Compounds 1 - 3 are structural constituents of 4, which corresponds to the sequence at the branch site of Group II intron b11 from yeast mitochondria. The present data provide further substantiation to our previous preliminary conclusion that trimer 1 and pentamer 3 show very similar conformational features, which are remarkably distinct from tetramer 2 and heptamer 4. It is found that the absence of a nucleotide unit at the 5'-site of the branch point (A) leads to a conformation at the branch point which is characterized by C2'-endo (S) ribose, syn orientation of the A-base, and A2'→5'G base stacking. This conformation is encountered for trimer 1, and -to a larger extent- also for pentamer 3. The introduction of a mono- or dinucleotide unit at the 5'-site of the branch point introduces a marked conformational transition at the branch point, in such a way that the conformation becomes C3'-endo (N) ribose, anti orientation of the base, and disruption of A2' → 5'G base stacking. This conformation is found for tetramer 2 and heptamer 4. The structural differences between 1,3 and 2,4 are highly reminiscent of the 5'-conformational transmission effect for a series of linear trimer and tetramer RNA molecules. We therefore propose that the 5'-conformational model is also of great interest with respect to our understanding of the structural preferences of branched RNAs. However, our chemical shift vs. temperature profiles and NOE data on 3 and 4 show several additional conformational features. We propose a model for the secondary structure of pentamer 3, which entails peculiar backfolding of the 3'-linked chain in such a way that stacking is achieved between the remote bases C+2 (in the 3'-linked chain), and branch point A. Also, our NOE data on heptamer 4 led to the suggestion that a stack exists between the bases U+2 (in the 2'-linked chain), and U-1 (upstream of A). All our experimental observations on 3 and 4 were integrated into preliminary structural models, shown in Figure 10. It was impossible, however, to obtain strong support for these models from NOE data, since virtually no interresidual NOE crosspeaks appeared in our 2D NOESY spectra. Only in the case of 4, we could observe four weak interresidual NOEs at mixing times of 300, 500 and 900 ms in the aromatic spectral region, provided that a sample temperature of 292 K was used; elevation of the sample temperature to 303 K rendered these NOEs invisible.

Many genes are interrupted by stretches of non-coding DNA called intervening sequences or introns.[1] Transcription of these genes results in pre-mRNA, pre-rRNA, or pre-tRNA, i.e., RNA strands in which both the coding sequences (exons), and introns are still present. The process of excission of the introns and the subsequent ligation of the exons is called RNA splicing. Thus, the splice process generates a mature RNA molecule which subsequently can serve as a template in protein synthesis using the ribosomal machinery in the cytoplasm.

To date, the splice mechanisms of four major groups of precursor RNAs have been characterized.[2] Splicing of Group I precursor RNA requires guanosine, or a guanosine derivative as a cofactor.[3] All Group I introns entail essentially the same conserved sequence elements, leading to a typical secondary structure. A key feature of Group I splicing is the binding of the cofactor to a specific binding site in the intron. Consequently, the guanosine (derivative) becomes oriented in such a way that nucleophilic attack of its 3'-OH group on the phosphorus atom at

the 5'-splice site is greatly facilitated. Self splicing of *Tetrahymena thermophila*, as originally discovered by Cech et al., is an example of Group I splicing.

Group II introns possess a different secondary structure which allows the splicing reaction to proceed even without any cofactor or any external source of energy such as ATP or GTP, but Mg^{2+} ion is necessary.[1e,4] In the total inventory of the reaction two phosphodiester bonds are formed and two are broken, and the reaction is self-driven. Splicing of Group II introns is initiated by nucleophilic attack of the 2'-OH group of an adenosine residue in the intron sequence, to the phosphorus atom at the 5'-splice site next to guanosine. In this way, a 2'→5' phosphodiester bond is formed, and the intron is ultimately excised in a typical lariat form. The splice mechanism of nuclear mRNA, the third class of precursor RNAs, is mechanistically related to Group II splicing, i.e. the intron is also released in the lariat form. However, nuclear mRNA does not show self-splicing, since an assembly of the substrate with a large complex of proteins and small ribonucleoproteins is required to form a spliceosome.[5] The fourth major group of precursor RNAs are the nuclear tRNAs. Splicing of these structures occurs through involvement of ATP and at least four different enzymes (i.e., an endonuclease, a kinase, a ligase, and a phosphatase).

Our interest in RNA splicing is focussed on the exact role of lariat formation during splicing of Group II and nuclear mRNA precursors. Our approach comprises the chemical synthesis of branched RNA structures that mimic the branch point, and conformational analysis of these systems by spectroscopic techniques, in particular high-field NMR.[6] Over the last years, we have developed synthetic routes to branched RNA model systems of increasing structural complexity, i.e. branched trimers (e.g. **1**),[6a,m] tetramers (e.g. **2**),[6j,l,n] pentamers (e.g. **3**),[6d] and one heptamer structure, **4**[6d] which corresponds to the sequence at the branch site of Group II intron b11 from yeast mitochondria. It should be noted that the compounds **1** - **3** are structural constituents of **4** (see Figure 1; the numberings of constituent nucleoside residues used throughout this paper are also shown). In our previous work, we have already shown that the branched structures **1** - **4** can be divided in two groups on the basis of their conformational characteristics.[6n] The first group comprises trimer **1** and pentamer **3**, which show essentially the same type of conformation, with predominant C2'-*endo* (S) conformation of the ribose ring at the branch point, and A2'→5'G base stacking. The second group contains tetramer **2** and heptamer **4**, which show a conformational bias toward C3'-*endo* (N) for the ribose ring at the branch point, along with $U_{(-1)}\rightarrow A\rightarrow U_{(+1)}$ base stacking. Presently, we wish to report detailed NMR-spectroscopic data that reveal entirely new conformational information about our pentameric and heptameric branched RNAs. We used several homo- and heteronuclear correlated 2D NMR experiments in order to arrive at an unambiguous assignment for all proton resonances of **3** and **4**. Then, we performed three essentially different classes of experiments in order to investigate the molecular conformations: (i) double quantum filtered (DQF) COSY with and without ^{31}P decoupling was used in order to estimate most of the vicinal ^{1}H - ^{1}H and ^{1}H - ^{31}P coupling constants; (ii) 1D NMR spectra, recorded in the temperature range 8 - 85 °C yielded chemical shift (δ) vs. temperature profiles for the non-exchangeable base protons, and the anomeric protons. These curves provided very important information concerning stack-destack equilibria which shift upon variation of the sample temperature; (iii), inspection of the 2D NOESY spectra provided qualitative information about spatial proximities of protons. Although most of the NOE cross peaks that could be detected are due to intraresidual NOE contacts, these experiments provided important information about the glycosidic conformations, especially those of the branch point A residue, and the neighbouring 2'→5' linked G residue. Above all, the present conformational data on **3** and **4** show the importance of the so-called 5'→3' conformational transmission effect with respect to the secondary structure of branched RNAs. Alternatively stated, it was found that the presence of at least one nucleotide linked to the 5'-end

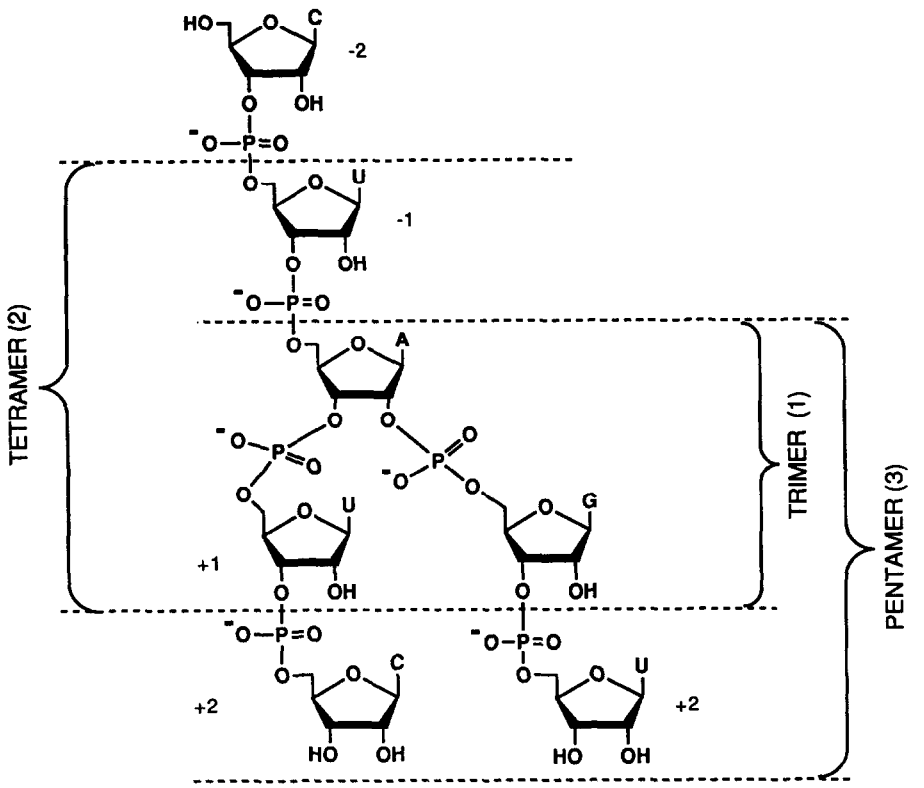

Figure 1. Detailed structural formula of the largest model system studied in this work i.e., the branched heptamer **4**. C = 1-cytosine-yl, U = 1-uracil-yl, A = 9-adenine-yl, G = 9-guanine-yl. The numbering scheme indicated in this structural formula is used throughout this work. Note that trimer **1**, tetramer **2**, & pentamer **3** shown in this figure are structural constituents of heptamer **4**; the numbering scheme as defined for **4** also applies to **1** - **3**.

of the central adenosine residue is necessary to mimic the conformation of the lariat branch point. Based on the present data we propose preliminary structural models for branched pentamer **3** and heptamer **4**.

RESULTS AND DISCUSSION

(A) *Assignment of resonances.* We were able to assign all sugar protons and non-exchangeable base protons in pentamer **3** and heptamer **4**. Our assignments were based on the combined interpretation of several correlated 2D NMR experiments, in particular Hartmann-Hahn (HOHAHA)[7], DQF-COSY[8], and ^{1}H - ^{31}P correlated[9] spectra. Figures 2 and 3 show expansions of the HOHAHA and the ^{1}H - ^{31}P correlated spectra of **4**. In addition we used 2D NOESY in order to obtain unambiguous assignments in the aromatic spectral region (i.e. H2 of A, H6 of U and C, H8 of A and G). Figure 4 shows the detailed assignments in the aromatic (δ 8.2 - 7.5 ppm) and anomeric/aromatic (δ 6.2 - 5.5 ppm) spectral regions of **3** and **4**. Our detailed assignments of proton resonances are summarized in Tables 1 and 2.

(B) *J-coupling analysis.* This study was aimed particularly at the determination of the vicinal ^{1}H - ^{1}H and ^{1}H - ^{31}P J coupling constants for **3** and **4**, since these parameters contain a wealth of information concerning the molecular conformation in solution.[10] The conformational equilibria of all individual ribose rings and the

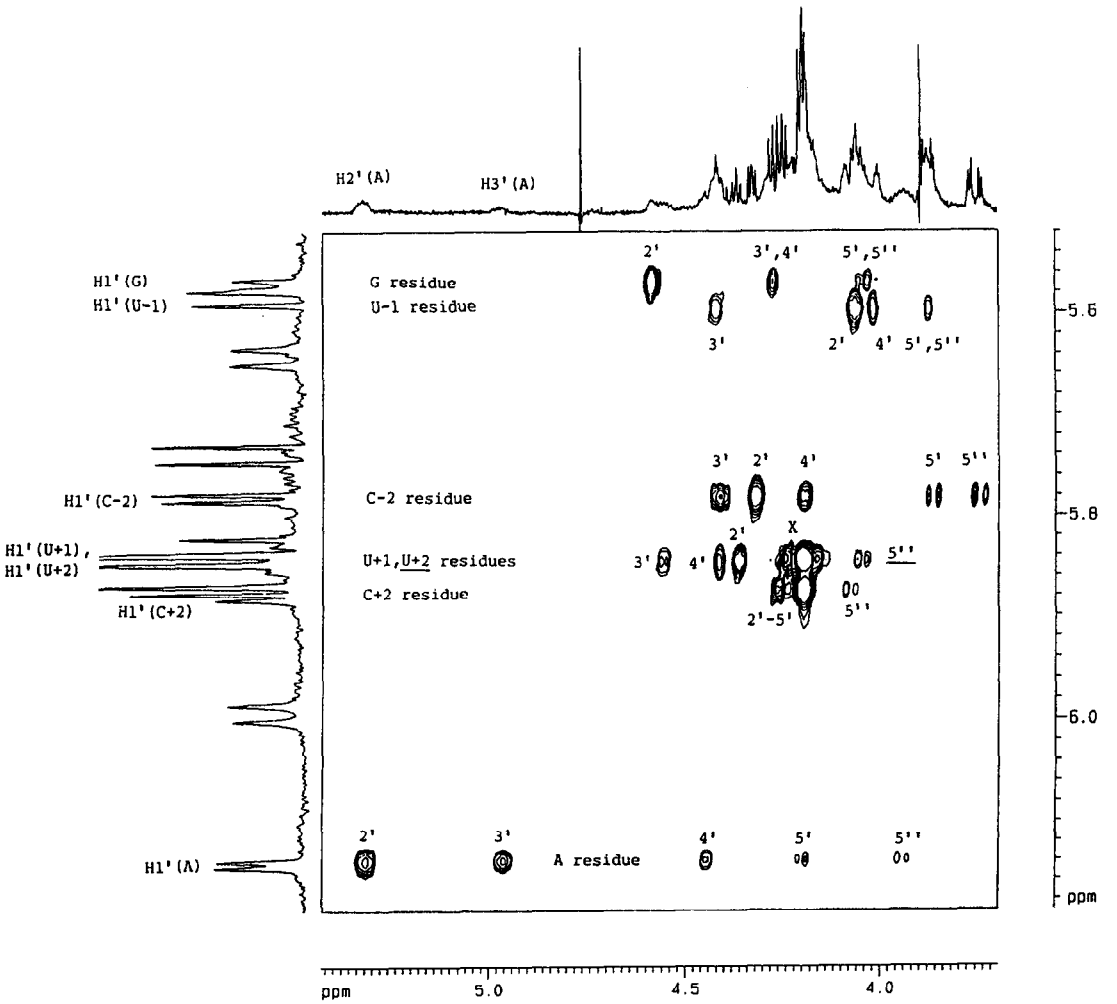

Figure 2. Expansion of the 500 MHz homonuclear Hartmann-Hahn spectrum of **4**. Sample temperature 19 °C, oligomer concentration approximately 7 mM. The crosspeak marked "x" is due to H5' of residue U+1, and H2'-H5' of residue U+2. The assignments as indicated in this figure were derived also on the basis of the ^{1}H - ^{31}P correlated spectrum (Figure 3), and the DQF-COSY spectra (vide infra).

backbone bonds C4' - C5' (γ) and C5' - O5' (β) can be determined from J coupling data, using relatively straightforward methods.[10] Most of our J-couplings determined on **3** and **4** were abstracted from high resolution DQF COSY spectra (resolution 1.1 Hz/point). DQF COSY spectra of **3** and **4** were recorded both with and without ^{31}P decoupling.[8] This technique proved to be very powerful, since it enabled us to obtain detailed resonance identifications and good estimations for chemical shifts and vicinal ^{1}H - ^{1}H and ^{1}H - ^{31}P coupling constants despite considerable crowding in the spectral region δ 4.6 - 3.8 ppm. It should be noted that DQF COSY was also used successfully by Tinoco et al. in their recent 500 MHz ^{1}H-NMR conformational analysis of the RNA hairpin loop 5'-(GGAC(UUCG)GUCC)-3'.[8c] Figure 5 shows a typical expansion of a set of crosspeaks in the ^{31}P decoupled DQF COSY spectrum of **4**. This figure also illustrates how the active and passive J coupling constants were actually abstracted from each crosspeak. Estimates for the vicinal ^{1}H - ^{31}P J couplings were obtained by comparing the total width of relevant crosspeaks in the ^{31}P coupled and ^{31}P

Table 1. Assignments of chemical shifts (ppm) for pentamer **3** at 19 °C (oligomer concentration 4 mM)

Chemical Shifts (δ)	A residue	G residue	U+1 residue	U+2 residue	C+2 residue
H-1′	6.12	5.23	5.85	5.85	5.80
H-2′	5.23	4.51	4.37	4.18	4.12
H-3′	4.86	4.55	4.57	4.18	4.21
H-4′	4.45	4.13	4.41	a	a
H-5′	3.80	3.95	4.27	a	a
H-5″	3.71	3.67	4.12	a	a
H-2	7.71	-	-	-	-
H-5	-	-	5.75	5.63	5.90
H-6	-	-	7.78	7.73	7.76
H-8	8.14	7.61	-	-	-

a: could not be determined

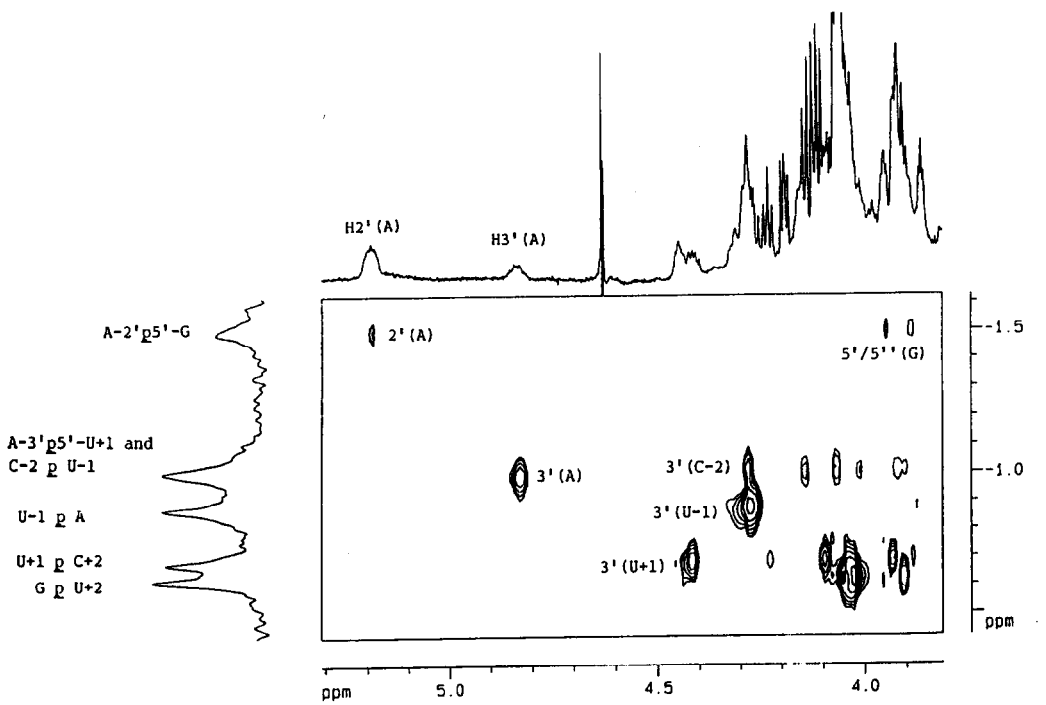

Figure 3. Heteronuclear ^{1}H - ^{31}P correlation spectrum of **4**. Sample temperature 25 °C, oligomer concentration approximately 7 mM.

decoupled DQF COSY spectra. This procedure is delineated in Figure 6. The sets of vicinal ^{1}H - ^{1}H and ^{1}H - ^{31}P J coupling constants measured for **3** and **4** at 19 and 40(45) °C are summarized in Tables 3 and 4, respectively. For the conformational analysis of the ribose rings in **3** and **4** there are several approaches available. A detailed pseudorotational analysis in terms of phase angles, puckering amplitudes, and equilibrium constant

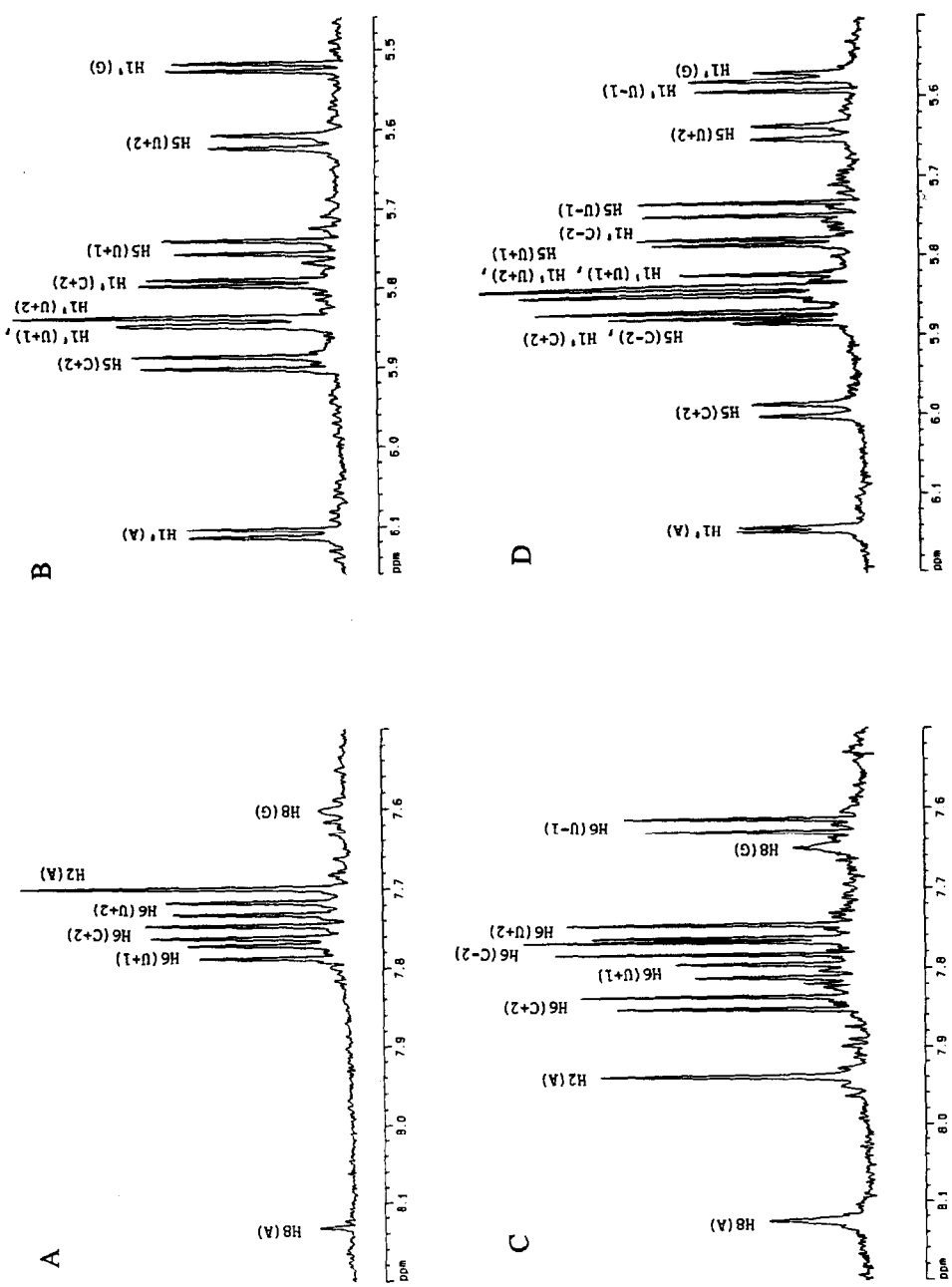

Figure 4. A: H2, H6, H8 region of the 500 MHz ^{1}H NMR spectrum of **3** in D_2O. Sample temperature 19 °C, oligomer concentration approximately 2-2.5 mM. Note that H8A and H8G are almost completely exchanged by solvent deuterium. The H1', H5 region of the same spectrum is shown in panel B. C: H2, H6, H8 region of the 500 MHz ^{1}H NMR spectrum of **4** in D_2O. Sample temperature 19 °C, oligomer concentration approximately 3 mM. The H1', H5 region of the same spectrum is shown in panel D.

Table 2: Assignments of chemical shifts for heptamer **4** at 19 °C (oligomer concentration 7 mM)

Chemical Shifts (δ)	C-2 residue (ppm)	U-1 residue (ppm)	A residue (ppm)	G residue (ppm)	U+1 residue (ppm)	U+2 residue (ppm)	C+2 residue (ppm)
H-1′	5.80	5.61	6.17	5.59	5.86	5.85	5.92
H-2′	4.35	4.08	5.35	4.61	4.39	4.22	4.22
H-3′	4.44	4.45	5.02	4.31	4.58	a	a
H-4′	4.22	4.03	4.47	4.31	4.44	a	a
H-5′	3.91	3.91	4.23	4.12	4.31	4.20	4.26
H-5"	3.79	3.91	3.96	4.05	4.18	4.07	4.10
H-2	-	-	7.94	-	-	-	-
H-5	5.87	5.74	-	-	5.83	5.66	6.00
H-6	7.78	7.63	-	-	7.81	7.77	7.85
H-8	-	-	8.13	7.66	-	-	-

a: could not be determined.

Table 3. J couplings (Hz) for pentamer **3** at 19 °C and 45 °C (oligomer concentration 4 mM).

J couplings	A residue		G residue		U+1 residue		U+2 residue		C+2 residue	
	19 °C	45 °C	19 °C	45 °C	19 °C	45 °C	19 °C	45 °C	19 °C	45 °C
$J_{1',2'}$	5.2	(5.3)	4.4	(5.1)	5.2	(5.1)	5.3	(5.1)	3.6	(4.1)
$J_{2',3'}$	5.1	(4.9)	5.6	(5.1)	5.6	(6.1)	5.1	(5.1)	< 4	(5.1)
$J_{3',4'}$	3.1	(3.3)	4.9	(5.6)	3.6	a	4.1	a	< 4	a
$J_{4',5'}$	3.4	(3.1)	2.3	(2.0)	3.1	(4.2)	3.0	a	a	a
$J_{4',5"}$	4.2	(4.1)	3.1	(3.1)	3.1	(3.2)	3.0	a	a	a
$J_{2',P}$	8.1	(8.1)	-	-	-	-	-	-	a	a
$J_{3',P}$	6.0	(7.7)	8.5	(7.1)	7.0	(8.1)	-	-	a	a
$J_{4',P}$	-	-	1.5	a	1.5	a	a	a	a	a
$J_{5',P}$	-	-	3.7	(4.1)	4.8	(4.0)	a	a	a	a
$J_{5",P}$	-	-	2.9	(3.3)	4.3	(2.0)	a	a	a	a

a: could not be determined.

of the C2'-*endo* (S) / C3'-*endo* (N) equilibria requires the accurate determination of $J_{1'2'}$, $J_{2'3'}$, and $J_{3'4'}$ for several sample temperatures. In view of the fact that $J_{2'3'}$ and $J_{3'4'}$ had to be abstracted from DQF COSY (i.e. accuracy ± 1 Hz), we decided to use merely the $J_{1'2'}$ values to calculate the populations of the C3'-*endo* (N) type ribose conformer according to the formula[10a]:

$$\%(\text{C3'-}endo\ (\text{N})) = 100 \times (7.9 - J_{1'2'}) / 6.9$$

For most ribose rings in **3** and **4**, it was possible to measure $J_{1'2'}$ with high accuracy (± 0.2 Hz) directly from the one-dimensional ^{1}H-NMR spectrum. In this way, we arrived at accurate $J_{1'2'}$ values for the residues A, G, and C+2 in pentamer **3**, and C-2, U-1, A, G, and C+2 in heptamer **4** (see also Figure 4). Tables 5 and 6 summarize the C3'-*endo* (N) populations as derived for **3** and **4** at 19 °C; Figure 7 shows these data in a graphical representation. Inspection of Figure 7 allows one to compare the conformations of the ribose units in the branched systems **1** - **4**, for a sample temperature of 292 K. Note that data points in the centre of these graphs

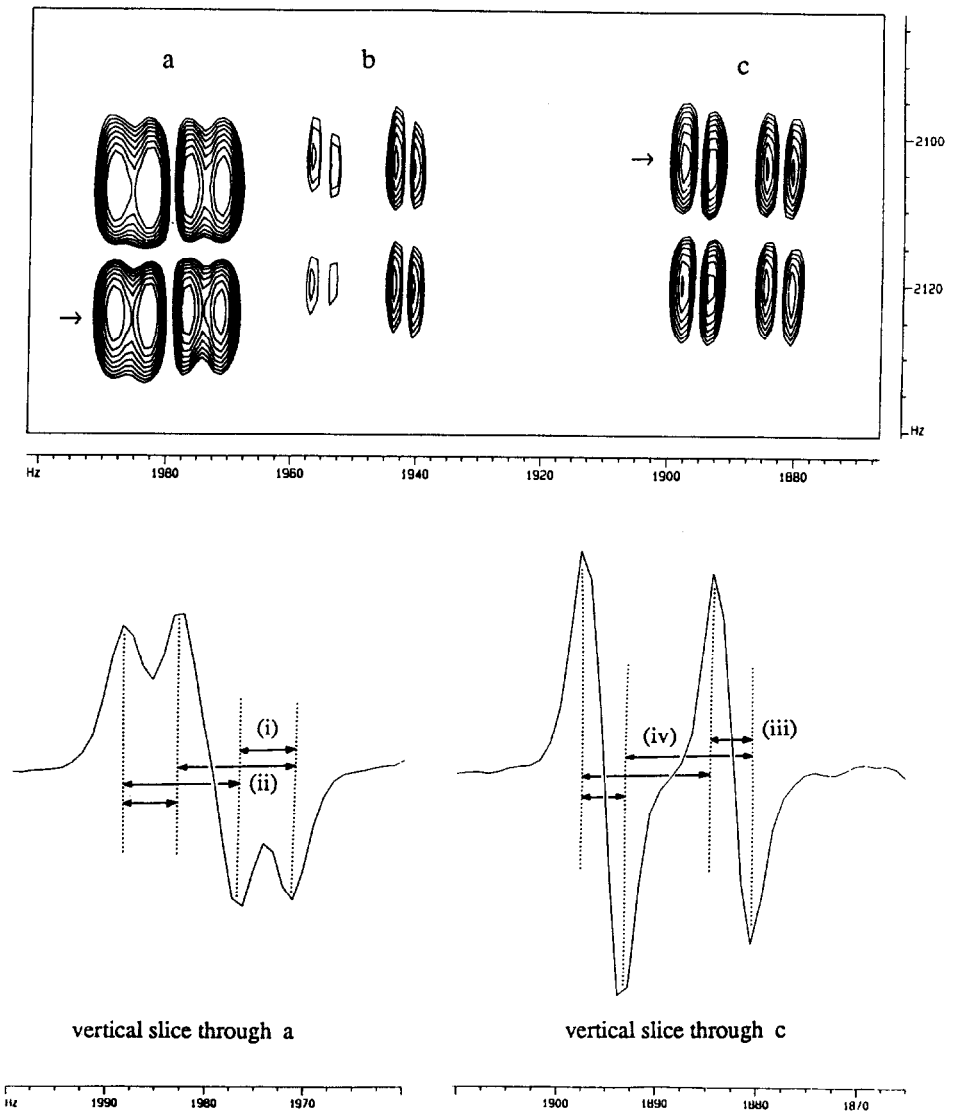

Figure 5. Detailed expansion of the contour plot of the ^{31}P decoupled DQF COSY spectrum of **4** (top). Sample temperature 19 °C, oligomer concentration 7 mM. Crosspeak a: H5' (low field) - H5" (high field) of the A-residue; b: H4' (low field) - H5' (high field) of residue C-2; c: H4' (low field) - H5" (high field) of residue C-2. Note that the active and passive J-couplings can not be distinguished directly from these contour plots. The lower part of this figure represents vertical slices through crosspeaks a and c, taken at the sites indicated with an arrow. Subspectrum a is composed of a small passive (in phase) J coupling marked (i), i.e. $J_{4'5'}$ of A = 0.9 ± 1 Hz, and a large active (out of phase) J coupling marked (ii), i.e. $J_{5'5''}$ of A = 11.6 ± 1 Hz. Subspectrum c is composed of a small active J coupling (i.e. $J_{4'5''}$ of C-2 = 6.0 ± 1 Hz), and a large passive J coupling (i.e. $J_{5'5''}$ of C-2 = 11.5 ± 1 Hz). The active and passive J-couplings in subspectrum c are indicated by the arrows (iii) and (iv), respectively. By careful analysis of all crosspeaks in the DQF COSY spectrum, confirmation of all prior assignments (from HOHAHA and ^{1}H - ^{31}P correlated spectra) was obtained.

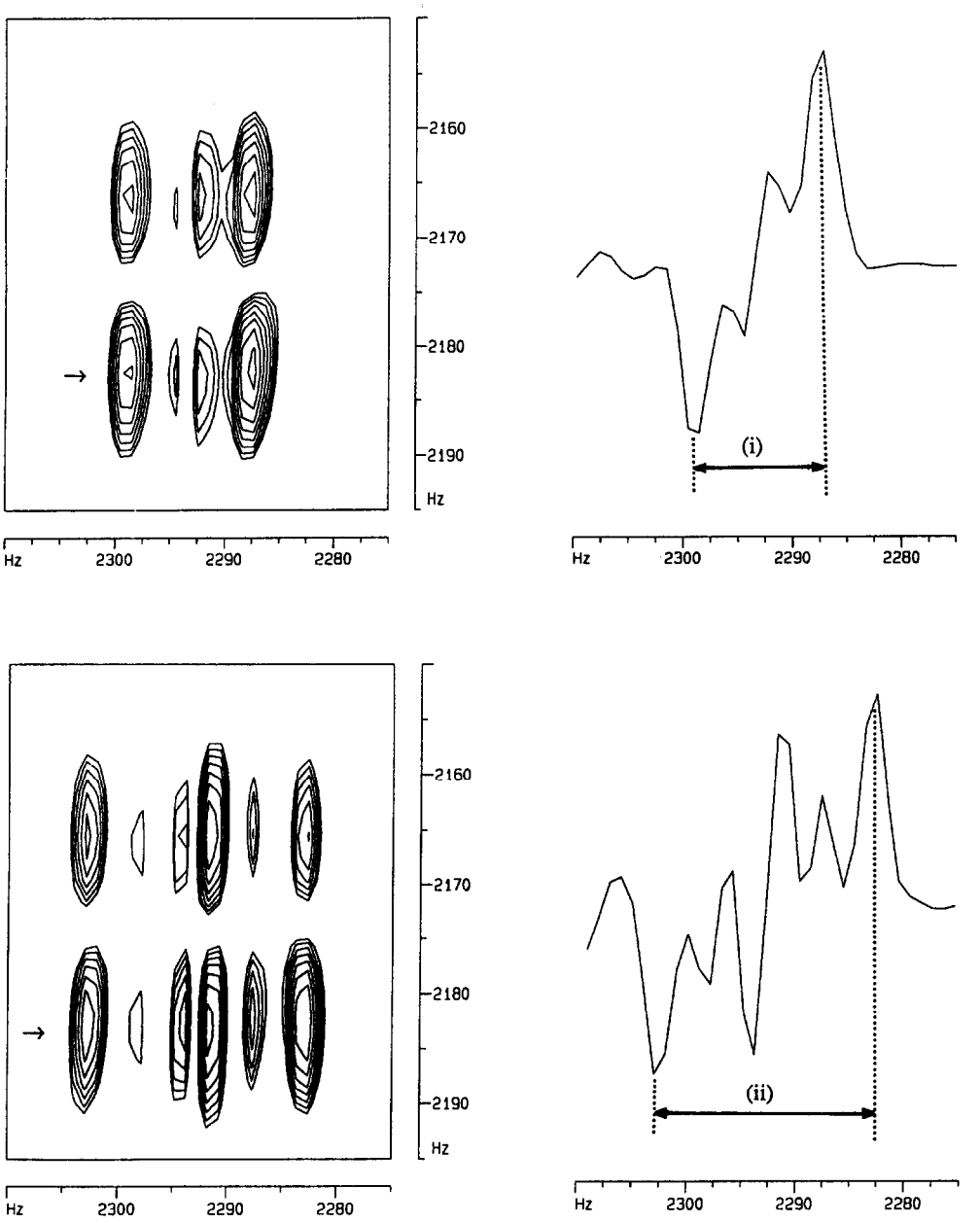

Figure 6. Expanded crosspeak in the ^{31}P decoupled (top) and ^{31}P coupled (bottom) DQF COSY spectra of **4**. Vertical slices through these patterns are shown on the righthand side, slices were taken at the sites indicated by an arrow. The difference in the total width of these spectra (substraction of the arrow lengths (i) and (ii)) provides a means to estimate $^{3}J_{HP}$.

Table 4. J couplings (Hz) for heptamer **4** at 19 °C and 40 °C (oligomer concentration 7 mM)

J couplings	C-2 residue 19 °C (40 °C)	U-1 residue 19 °C (40 °C)	A residue 19 °C (40 °C)	G residue 19 °C (40 °C)	U+1 residue 19 °C (40 °C)	U+2 residue 19 °C (40 °C)	C+2 residue 19 °C (40 °C)
$J_{1'2'}$	3.6 (4.1)	6.1 (6.7)	3.1 (4.1)	4.1 (5.1)	5.2 (6.1)	3.8 (5.1)	3.6 (4.6)
$J_{2'3'}$	5.2 (5.1)	5.5 (5.2)	5.3 (5.3)	4.8 (5.6)	5.6 (5.1)	a (5.1)	a (4.6)
$J_{3'4'}$	5.8 (6.1)	2.9 (4.2)	5.5 (5.4)	4.4 (3.6)	4.0 (3.9)	(a) (a)	a (a)
$J_{4'5'}$	4.6 (4.7)	4.0 (c)	0.9 (1.0)	3.8 (d)	3.9 (4.7)	<4 (a)	< 4 (a)
$J_{4'5''}$	6.0 (5.7)	4.0 (c)	5.6 (6.7)	4.2 (d)	4.0 (4.2)	<4 (a)	< 4 (a)
$J_{2'P}$	(-) (-)	(-) (-)	8.9 (9.0)	-	- (-)	- (-)	- (-)
$J_{3'P}$	9.1 (8.6)	7.9 (7.2)	8.5 (8.6(	10.4 (7.1)	7.9 (8.4)	- (-)	- (-)
$J_{4'P}$	- (-)	1.7 (1.0)	<1 (1.0)	2.1 (a)	0.6	a (a)	a (a)
$J_{5'P}$	- (-)	b (b)	a (1.0)	3.7 (e)	4.7 (4.0)	a (a)	a (a)
$J_{5''P}$	- (-)	b (b)	3.1 (3.1)	2.5 (e)	2.4 a	a a	a (a)

a: Could not be determined; c: < 3 Hz; e: $J_{5'P} + J_{5''P} = 3.0$ Hz
b: $[J_{5',P} + J_{5'',P}] = 4.1$ Hz; d: $J_{4'5'} + J_{4'5''} = 6.6$ Hz

correspond with a C3'-*endo* (N) / C2'-*endo* (S) conformational equilibrium close to 50-50 %. Data points in the upper left or lower right regions point out that the corresponding ribose conformation is biased toward C3'-*endo* (N) or C2'-*endo* (S), respectively. The top graph shows our data points for trimer **1** (riboses A, G, and U+1) and pentamer **3** (riboses A, G, U+1, C+2, and U+2), i.e. *this graph visualizes how the trimer conformation is stabilized through the addition of the nucleotide units C+2 and U+2*. The ribose ring of U+1 shows a clear change (from 52% N / 48% S in **1** toward 41% N / 59% S in **3**), while the ribose conformations of A and G are virtually not affected. The added nucleotides C+2 and U+2 show conformational equilibria of 62% N / 38 %S, and 41% N / 59% S, respectively (see also Table 5). These data reveal clearly that **1** and **3** adopt the same type of conformation, which is characterized by a preferred C2'-*endo* (S) conformation at the branch point.[11] Our subsequent experiments substantiated the conformational similarity of trimer **1** and pentamer **3** (vide infra). The bottom graph in Figure 8 shows the data for tetramer **2** (riboses U-1, A, G, and U+1) and heptamer **4** (riboses C-2, U-1, A, G, U+1, C+2, U+2), i.e., here we see *the stabilization of the entire core of the branch, that is achieved via the addition of the extra nucleotides* C-2, C+2, and U+2. It is clearly seen that the riboses of A and G move substantially toward C3'-*endo* (N) (A: 58% N / 42 % S in **2**, and 69% N / 31% S in **4**; G: 38% N / 62

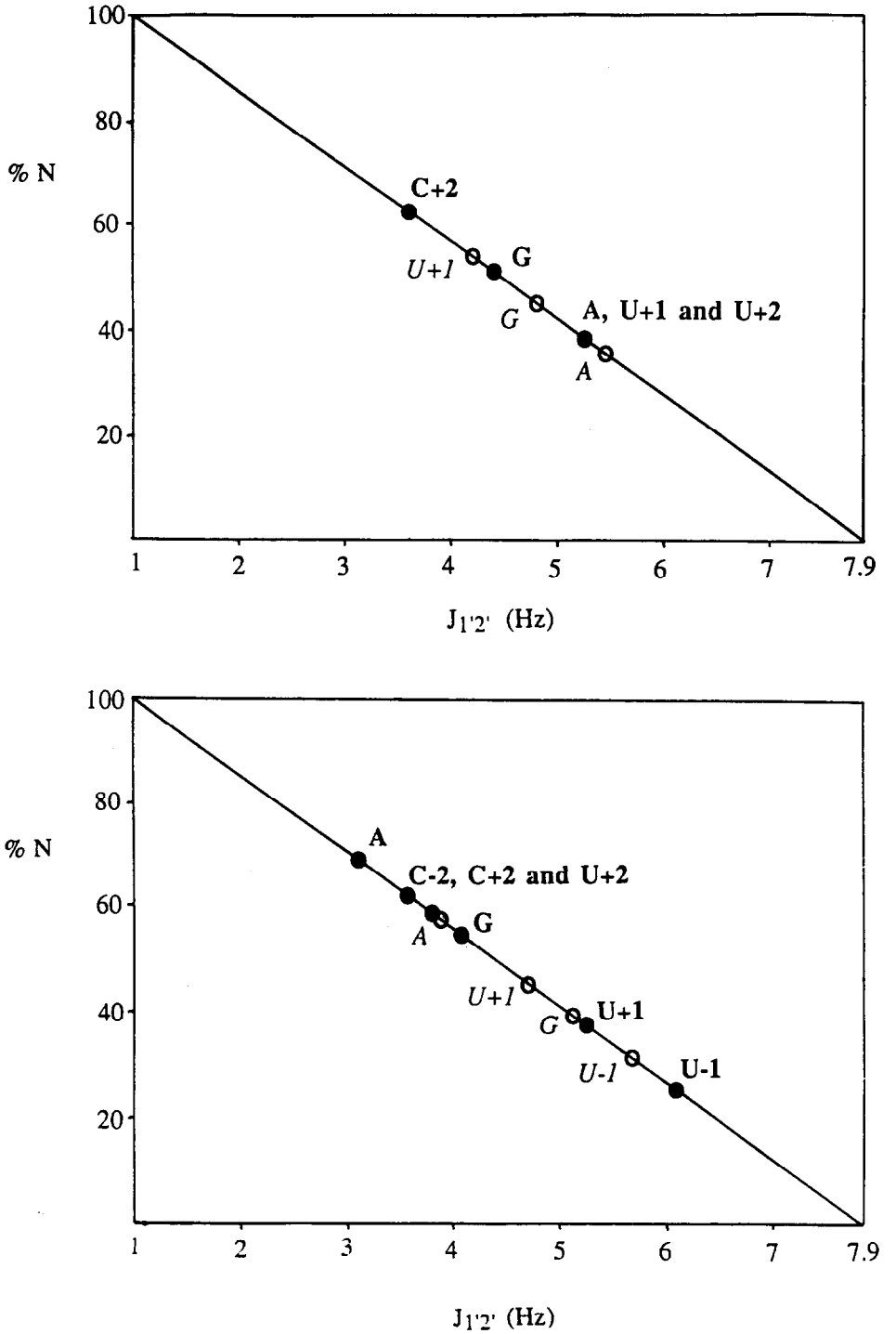

Figure 7. Graphical representation of the experimental $J_{1'2'}$ values and the corresponding populations of the C3'-*endo* (N) type puckered ribose rings for **1** - **4**. The top graph shows a comparison of pentamer **3** (filled circles, bold labels) with trimer **1** (open circles, italic labels). The lower graph analogously compares heptamer **4** (filled circles, bold labels) with tetramer **2** (open circles, italic labels).

Table 5. Calculated time-averaged populations of the predominant C4'-C5' (γ) and C5'-O5' (β) backbone rotamers, and the C3'-*endo* type puckered ribose rings in pentamer **3** at 19 and 45 °C.

Residue	Temperature (°C)	% γ+ (a)	% β t (b)	% 3'-*endo* (% N) (c)
A	19	68	-	40
	45	63	-	38
G	19	84	92	47
	45	72	86	41
U+1	19	74	80	41
	45	d	79	41
C+2	19	> 55	d	62
	45	d	d	55
U+2	19	75	d	62
	45	61	d	41

a: calculated using the equation: $\%(\gamma^+) = 100 \times (13.3 - \Sigma) / 9.7$, with $\Sigma = J_{4'5'} + J_{4'5''}$

b: calculated using the equation: $\%\ (\beta^t) = 100 \times (23.9 - \Sigma') / 18.9$, with $\Sigma' = J_{5'P} + J_{5''P}$

c: calculated using the equation: %(C3'-*endo*) = 100 x (7.9 - $J_{1'2'}$) / 6.9

d: could not be determined

%S in **2**, and 52 %N / 48 %S in **4**). The ribose rings of U-1 and U+1, on the other hand, move toward the C2'-*endo* (S) form (U-1: 30% N / 70% S in **2**, 23% N / 77% S in **4**; U+1 : 43% N / 57 %S in **2**, 37% N/63% S in **4**). The added nucleotides C-2, C+2, and U+2 are clustered around 60% N / 40% S conformational equilibria for the ribose rings. *These data show that tetramer* **2** *and heptamer* **4** *adopt essentially the same type of conformation which is distinct from the conformation encountered for* **1** *and* **3**. The elongation of the 5'-, 3'-, and 2'- tails of the tetramer has markedly stabilized the preferred conformation at the core of the branch in the heptamer **4**, i.e., outspoken preferences for N-type riboses are found for the residues A and G, and preferences for S-type riboses for the residues U-1 and U+1. Evidently, heptamer **4** is at present the best available model for the conformation at the branch-point in the lariat-introns of Group II and nuclear mRNA splicing. Our present data are therefore of interest with respect to the construction of a structural model for branched RNA. The conformational analysis of the C4' - C5' backbone bonds rests upon the coupling constants $J_{4'5'}$ and $J_{4'5''}$. Detailed analysis in terms of a rotamer populations for all three staggered rotamers around C4' - C5' requires an unambiguous assignment for H5' and H5". However, according to Altona it is possible to correlate the sum $J_{4'5'} + J_{4'5''}$ directly to the population of the γ^+ rotamer (i.e. the rotamer in which O5' is *gauche* with respect to O4' and C3'): $\%\ (\gamma^+) = 100 \times (13.3 - \Sigma) / 9.7$, with $\Sigma = J_{4'5'} + J_{4'5''}$.[10a] The calculated γ^+ populations for **3** and **4** are also listed in Tables 5 and 6, respectively. It can be clearly seen from these data that the γ^+ rotamer is preferred for all residues in **3** and **4**. It is of great interest to compare the γ^+ rotamer populations of the 2'-linked G unit for the compounds **1** - **4**, especially since it is known that the population of γ^+ is related to the extent of base-base stacking along the 2'→5' axis.[12] Our data of $x(\gamma^+)$ for G unit in **1** - **4** are as follows: **1**: $x(\gamma^+)$ = 82% at 27 °C [6p]; **2**: $x(\gamma^+)$ = 76% at 27 °C [6p]; **3**: $x(\gamma^+)$ = 84% at 19 °C ; **4**: $x(\gamma^+)$ = 55% at 19 °C. These data lead to two important conclusions: (i) $x(\gamma^+)$ is much lower for **2** and **4** than for **1** and **3**, i.e., A2'→5'G base stacking is a predominant structural factor in **1** and **3**, but not for **2** and **4**; (ii) heptamer **4** shows a clearly

Table 6. Calculated time-averaged populations of the predominant C4'-C5' (γ) and C5'-O5' (β) backbone rotamers, and the C3'-*endo* type puckered ribose rings in heptamer **4** at 19 °C and 45 °C

Residue	Temperature (T °C)	% γ^+ (a)	% β^t (b)	% 3'-*endo* (c) (% N)
C-2	19	28	-	62
	40	5	-	55
U-1	19	55	100	28
	40	> 50	100	17
A	19	70	d	70
	40	d	100	55
G	19	55	94	57
	40	69	100	41
U+1	19	56	90	40
	40	45	d	26
C+2	19	> 50	d	62
	40	d	d	48
U+2	19	> 50	d	61
	40	d	d	41

a: calculated using the equation: $\%(\gamma^+) = 100 \times (13.3 - \Sigma) / 9.7$, with $\Sigma = J_{4'5'} + J_{4'5''}$
b: calculated using the equation: $\%\ (\beta^t) = 100 \times (23.9 - \Sigma') / 18.9$, with $\Sigma' = J_{5'P} + J_{5''P}$
c: calculated using the equation: %(C3'-*endo*) = 100 x (7.9 - $J_{1'2'}$) / 6.9
d: could not be determined

diminished γ^+ population in comparison with tetramer **2**, showing that the importance of A2'→5G base stacking is reduced further upon the addition of the three extra nucleotides C-2, U+2, and C+2.

The conformational equilibrium around the C5' - O5' (β) backbone bonds can be characterized on the basis of the J coupling constants $J_{P5'}$ and $J_{P5''}$. We used a sum rule, which directly leads to an estimation of the β^t rotamer population: $\%\ \beta^t = 100 \times (23.9 - \Sigma') / 18.9$; with $\Sigma' = J_{P5'} + J_{P5''}$[10a]; β^t is the rotamer in which P is oriented *trans* with respect to C4'. For all β bonds in **3** and **4** that could be analyzed, we found that the populations of β^t by far exceed those of ($\beta^+ + \beta^-$) (Tables 5 and 6). This is commonly encountered for oligonucleotides in which partial or complete base stacking occurs.

It should be noted that the *simultaneous* occurrence of γ^+ around C4' - C5' and β^t around C5' - O5' leads to a planar "W" conformation for the coupling path P - O5' - C5' - C4' - H4', resulting in a long range $^4J_{PH4'}$ coupling constant.[10] This $^4J_{PH4'}$ coupling is observed for the residues G and U+1 in pentamer **3**, and also for the residues U-1, A, G, and U+1 in heptamer **4**, thus proving a marked preference for regular (γ^+, β^t) conformation for these units.

(C) *Chemical shift vs. temperature profiles.*

The next step in our conformational study of **3** and **4** comprised the measurement of the δ vs. T °C profiles of the non-exchangeable base protons and the anomeric protons. As is well-known, the δ-values of these protons are highly sensitive to ring current effects exerted by spatially proximate bases. Therefore, the aromatic and anomeric

chemical shifts provide a suitable probe to monitor stacking phenomena in oligonucleotides. In particular, protons residing above or below a purine base plane experience considerable shielding. These protons are expected to show substantial $\Delta\delta$ values upon raising the sample temperature, which causes gradual disruption of the stacking interactions. Especially H2A, H5C, and H5U are convenient markers for changes in stack-destack equilibria; interpretation of $\Delta\delta$ for H6(C), H6(U), H8(A), or H8(G) is much more difficult, since the chemical shifts of these protons also depend on the conformation around the glycosidic bond.[10b]

The δ vs. temperature profiles for **3** and **4** were measured on slightly diluted samples; concentrations for these experiments were approximately 2 - 2.5 mM, both for pentamer **3** and heptamer **4**. It was found that this dilution had almost no effect on the chemical shifts at 19 °C, which indicates that intermolecular association of **3** and **4** is not of great importance under our experimental conditions. The δ vs. temperature data for all non-exchangeable base protons in **3** and **4** are summarized in Tables 7 and 8, respectively. Figure 8 shows the δ vs. temperature profiles for the protons H2(A), H8(A), and H8(G) of the branched RNA models **1** - **4**. Figure 9 shows the δ vs. temperature profiles for the H5 protons of the U and C residues in **3** and **4**. The δ vs. temperature profiles of **1** and **2** were abstracted from our previous work, and merely included to facilitate comparison of the different systems. Figure 8 shows that enlargement of trimer **1** to pentamer **3** by the addition of the residues C+2 and U+2 leads to *upfield* shifting of H2(A), H8(A) and H8(G). This also holds for the enlargement of tetramer **2** to heptamer **4** by addition of the residues C-2, C+2 and U+2. Apparently, enlargement of the branched structures through elongation of the tails leads to larger contributions of base-base stacked states, and hence to more defined secondary structures. These observations are completely in line with our data on the ribose conformations (vide supra). Further inspection of Figures 8 and 9 also reveals information about the *mode of base stacking* in **1** - **4**. It is of particular interest to compare δ(H2(A)) at 20 °C for the four branched RNA models **1** - **4** (δ(H2(A)) = 7.80, 7.97, 7.70, and 7.93 ppm for **1** - **4**, respectively). These data show that H2A in pentamer **3** resonates 0.2 - 0.3 ppm *more upfield* than H2(A) in **2** or **4**, which is also clearly visible in Figure 4. This shielding of H2A in **3** must reflect the occurrence of A2'→5'G base stacking in **3**, rather than in **2** and **4**. The ring current of the G base then explains the upfield shift of H2(A) in **3**. Three independent observations support this reasoning: (i) at 20 °C, H8(G) in **3** resonates upfield from H8(G) in **2** and **4**, which is most likely due to the ring current effect of A; (ii)

Table 7. Chemical shifts of the non-exchangeable base protons in pentamer **3** for different sample temperatures. Chemical shifts were referenced against a trace of CH_3CN (δ 2.000 ppm), which was added as an internal standard (oligomer concentration 3 mM).

Temp. (°C)	H2(A)	H8(A)	H8(G)	H5(U+1)	H6(U+1)	H5(C+2)	H6(C+2)	H5(U+2)	H6(U+2)
10	7.691	8.145	7.600	5.750	7.791	5.853	7.757	5.612	7.741
20	7.705	8.138	7.613	5.752	7.783	5.880	7.760	5.625	7.732
30	7.719	8.127	7.627	5.756	7.773	5.898	7.765	5.658	7.727
40	7.761	8.118	7.643	5.762	7.763	5.921	7.768	5.690	7.722
50	7.796	8.113	7.655	5.768	7.756	5.940	7.768	5.711	7.716
60	7.828	8.110	7.667	5.773	7.750	5.955	7.767	5.729	7.711
70	7.861	8.109	7.678	5.778	7.743	5.975	7.765	5.744	7.704
80	7.891	8.109	7.686	5.782	7.739	5.982	7.762	5.755	7.699

elevation of the sample temperature to 80 °C leads to substantially smaller differences for H2(A) in **2** - **4**. It must be noted that trimer **1** represents some intermediate situation; A2'→5'G stacking occurs to a lesser extent than in pentamer **3**, as is evidenced by the relatively small shielding of H2A in **1**; (iii) a substantial chemical shift

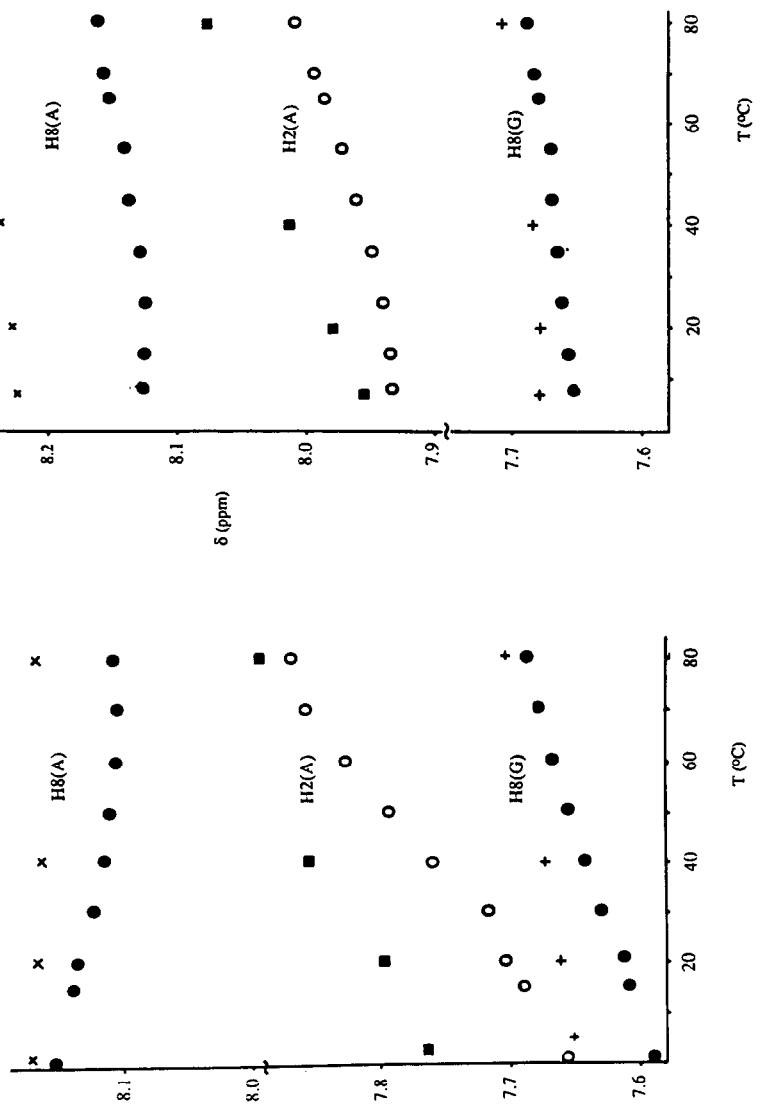

Figure 8. Left: Chemical shift (δ) vs. temperature profiles for H2(A), H8(A), and H8(G) of pentamer **3** and trimer **1**. Filled circles: H8(A) (top) and H8(G) (bottom) of **3**; open circles: H2(A) of **3**; x: H8(A) of **1**, filled squares: H2(A) of **1**, +: H8(G) of **1**. Data on **1** were abstracted from ref. 6g.
Right: Chemical shift (δ) vs. temperature profiles for H2(A), H8(A), and H8(G) of heptamer **4** and tetramer **2**. Filled circles: H8(A) (top) and H8(G) (bottom) of **4**; open circles: H2(A) of **4**; x: H8(A) of **2**, filled squares: H2(A) of **2**; x: H8(G) of **2**. Data on **2** were abstracted from ref. 6j.

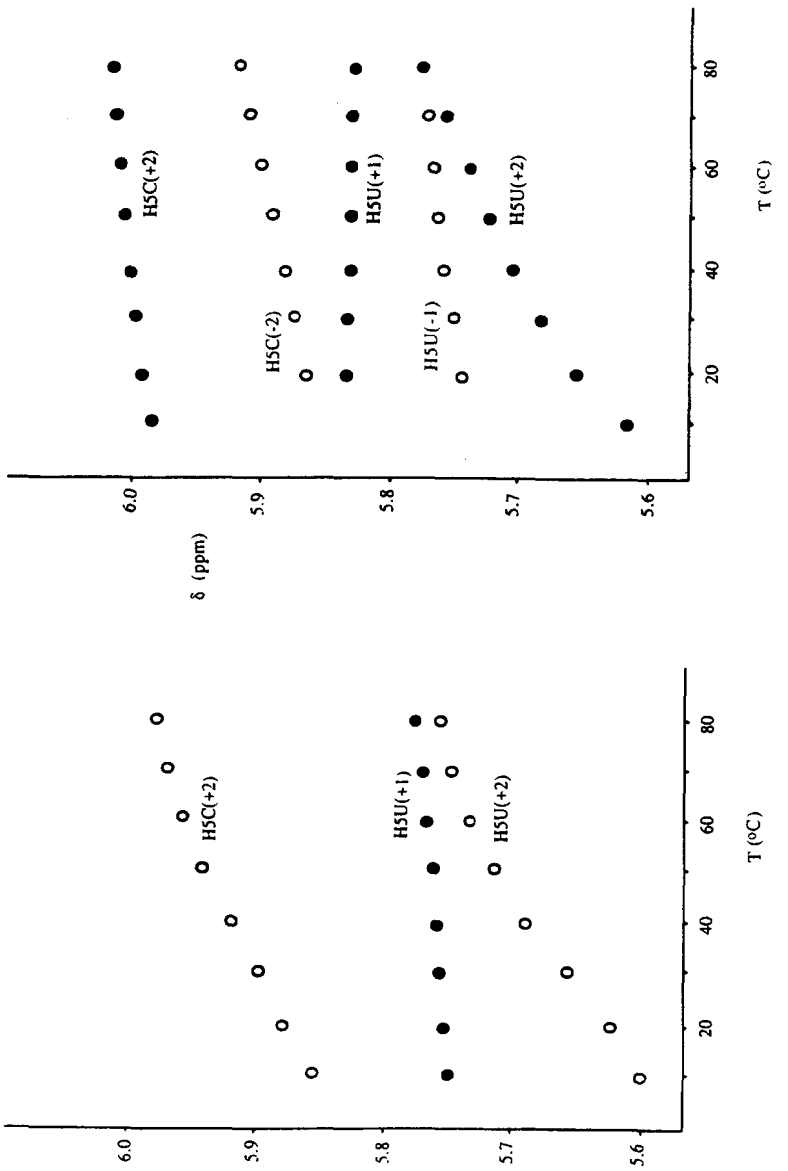

Figure 9. Left: Chemical shift (δ) vs. temperature profiles for H5 of the bases U and C in pentamer **3**. Filled circles: H5(U+1), open circles: H5(C+2) (top), and H5(U+2) (bottom). Right: Chemical shift (δ) vs. temperature profiles for H5 of the bases U and C in heptamer **4**. Filled circles: H5(C+2) (top), H5(U+1) (middle), H5(U+2) (bottom). Open circles: H5(C-2) (*top*), H5(U-1) (*bottom*).

Table 8. Chemical shifts of the non-exchangeable base protons in heptamer **4** for different sample temperatures. Chemical shifts were referenced against a trace of CH_3CN (δ 2.000 ppm), which was added as an internal standard (oligomer concentration 3 mM)

Temp. (T °C)	H5(C-2)	H6(C-2)	H5(U-1)	H6(U-1)	H2(A)	H8(A)	H8(G)
10	-	7.806	5.727	7.632	7.933	8.127	7.653
20	5.872	7.778	5.741	7.627	7.938	8.125	7.659
30	5.880	7.757	5.749	7.625	7.945	8.127	7.660
40	5.885	7.740	5.757	7.625	7.955	8.134	7.666
50	5.896	7.725	5.763	7.627	7.967	8.141	7.670
60	5.902	7.712	5.768	7.628	7.980	8.149	7.677
70	5.908	7.700	5.773	7.629	7.995	8.159	7.683
80	5.914	7.690	5.775	7.630	8.010	8.163	-

Temp. (T °C)	H5(U+2)	H6(U+1)	H5(C+2)	H6(C+2)	H5(U+2)	H6(U+2)
10	-	7.823	5.990	7.869	5.618	7.775
20	5.833	7.807	5.995	7.850	5.659	7.767
30	5.827	7.795	5.998	7.833	5.688	7.760
40	5.823	7.786	6.001	7.821	5.713	7.755
50	5.824	7.779	6.004	7.810	5.734	7.749
60	5.823	7.774	6.006	7.801	5.751	7.774
70	5.823	7.770	6.008	7.793	5.764	7.738
80	5.824	7.766	6.010	7.787	5.775	7.733

difference Δδ (Δδ = [δ(^{31}P) at 8 °C - δ(^{31}P) at 80 °C]) of approximately 0.55 ppm is found for the ^{31}P resonance of the A2'→5'G phosphodiester linkage in **3**; this Δδ(^{31}P) value definitely shows that there is considerable constraint in the structure of **3** around the 2'→5' axis[6k], which is in agreement with the proposed model. Comparison of this result with the Δδ(^{31}P) value for trimer **1** (0.59 ppm)[6k] confirms that a constraint exists around the 2'→5' linkage both for the trimer and the pentamer structure.

In Figure 9, it can be easily seen that H5U(+2) in **3** and **4** shows substantial deshielding upon increasing the sample temperature. This effect strongly points at regular 3'→5' stacking between G and U(+2). The δ vs. temperature profiles for H5U(+2) are almost identical for compounds **3** and **4**, which reveals that the mode and extent of G3'→5'U stacking in pentamer **3** and heptamer **4** must be very similar. Figure 9 also shows a large Δδ vs. temperature effect for H5C(+2) in pentamer **3**, but not in heptamer **4**. A possible explanation for the strong Δδ vs. temperature effect in **3** could be that the C+2 base experiences shielding from the A base by virtue of a back-folded conformation as shown in Figure 10C. The absence of any Δδ vs. temperature effect for H5U(+1) (viz. Figure 9) would support such a rough and qualitative model. It should also be noted that the presence of a nucleotide moiety on the 5'-site of A should preclude stacking between A and C+2. Indeed, H5C(+2) is considerably deshielded in tetramer **2** and heptamer **4**, if compared to pentamer **3** (20 °C: δH5C(+2) = 5.88 ppm for **3**, and 6.01 and 5.99 ppm in **2** and **4**, respectively). However, our NOESY experiments on **3** did not provide any support for the model in Figure 10C, as no interresidual NOEs could be observed (vide infra). The data points for heptamer **4** in Figures 8 and 9 show that only H5U(+2) exhibits a substantial Δδ vs temperature effect. Much smaller values were encountered for H2A and H5C(-2), while the other stack markers remain virtually at t he same δ-values over the entire temperature range. Remarkably, the bases U-1 nor U+1 appear not to be stacked

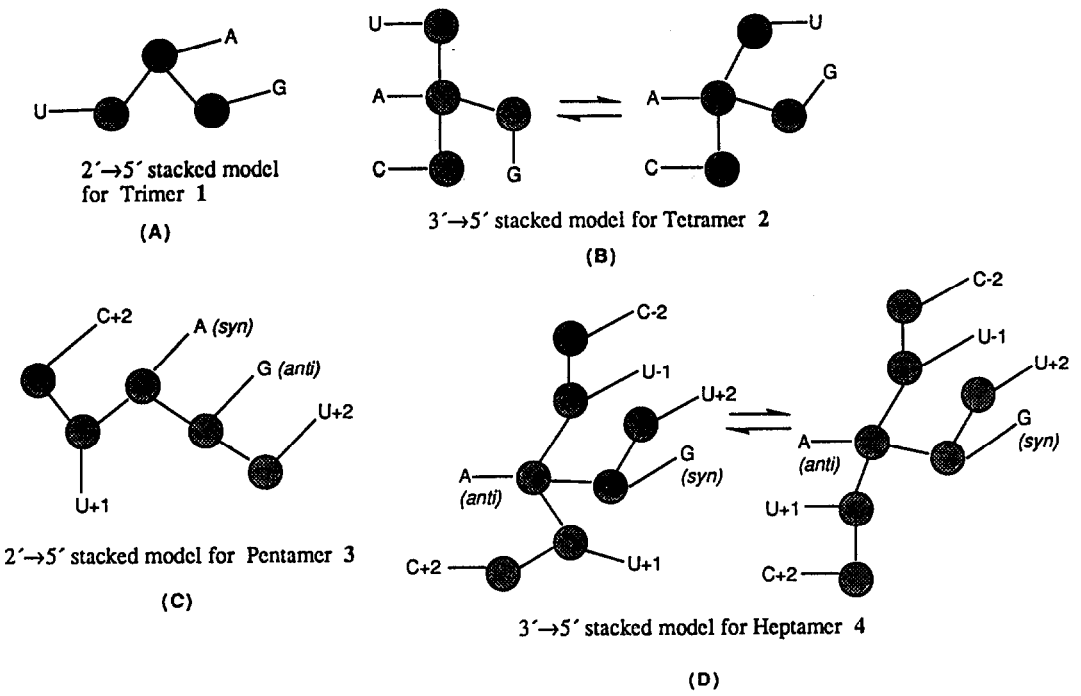

Figure 10. Preliminary working models of the secondary structure of the branched trimer **1** (A), tetramer **2** (B), pentamer **3** (C), and heptamer **4** (D). Models A and B are discussed in detail in ref. 6p.

to either A or G. These observations lead to the following proposals concerning base-base stacking in heptamer **4**: (i) U+2 is stacked to G (vide supra); (ii) C-2 is stacked to A in such a way that U-1 is bulged out; (iii) no strong stacking occurs between A and U+1. These structural features were essential for the construction of a preliminary structural model for heptamer **4**. Since this model was bolstered by NOE data, it will be presented and discussed in the next section (vide infra).

In summary, the δ vs. temperature profiles of the non-exchangeable base protons in **1** - **4** have revealed that A2'-5'G stacking predominates if the branch point a has a 5'-terminal OH group (structures **1** and **3**). The introduction of U-1 (structures **2** and **4**) changes the molecular conformation, and A2'→5'G stacking is disrupted or at least substantially weakened. Perhaps, the secondary structure of pentamer **3** is stabilized via a backfolding of C+2 in such a way that stacking with the A base at the branch point is achieved (Figure 10C). *We strongly feel that the conformational differences between (trimer* **1**, *pentamer* **3**), *and (tetramer* **2**, *heptamer* **4**) *constitute a new example of the so-called 5'-conformational transmission effect.* This effect was investigated previously by Tinoco et al.[13], Borer et al.[14], and more recently also by Altona et al.[15] The latter workers have studied the linear RNA oligomers A'UA (**5**), AA'UA (**6**), and UA'UA (**7**) [Compounds **5** - **7** have 6-N,N dimethylaminoadenine, which is denoted as A] in detail, and by comparing residue A' in **5** with residue A' in **6** and **7**, they concluded that introduction of an extra 5'-substituent A or U leads to strong A-A' stacking in **6**, and moderate U-A' stacking in **7**, resulting in the following changes in the conformational preferences: (i) ribose from S-type to N-type conformation; (ii) base from *syn* to *anti* orientation; (iii) ε (C3' - O3') from ε^- to ε^t. As a consequence of these changes, the U-residue which is bulged-out in **5** takes more part in regular stacking in **6** and **7**. *We now wish to point out that the branch point conformation of our systems follows the same pattern but in a more sophisticated manner: the absence of a 5'-substituent (as in* **1** *and* **3**) *leads to a branch-point*

conformation which is characterized by S-type ribose and syn orientation of the base, whereas the presence of one or two 5'-nucleotides (as in **2** *and* **4***) corresponds with a branch-point conformation with N-type ribose, and anti orientation of the base. However, the implication of the 5'-conformational transmission in our branched structures is unique in that the mode of stacking changes from A2'→5'G stack in* **1** *and* **3** *towards A3'→5'U stack in* **2** *and* **4**. *Apparently, the conformation at the branch point serves as an on-off switch for A2'→5'G versus A3'→5'U base stacking, and it is the 5'-U base (U-1) that drives this switch. The predominance of an S-type conformation at the branch-point allows for stacking along the 2'→5' axis, and it should be noted that the two known X-ray crystal structures of diribonucleoside-(2'→5')monophosphates indeed show an S-type conformation for the 5'-terminal nucleotide. (See X-ray crystal structures of A(H+)2'→5'U* [16,18] *and A2'→5'C(H+).* [17,18] *Also, several NMR studies on 2'→5' linked dinucleotides have shown that a preference exists for S-conformation in the 5'-terminal residue,*[19] *although for A2'- 5'A, the mode of stacking could not be defined.*[12] *On the other hand, the N-type conformation of the ribose ring at the branch-point allows for normal stacking along the 3'→5' axis, and it must be recalled that N-type ribose rings are encountered in native RNA structures. We propose therefore that a direct look at the ribose conformation at the branch-point can in fact tell where the stacking is: a large $J_{1'2'}$ coupling constant (approximately 6 Hz), showing an S-type ribose at the branch-point would indicate stacking along the 2'→5' axis, whereas a small $J_{1'2'}$ coupling constant (2-4 Hz) showing an N-type ribose ring at the branch-point can serve as a diagnostic for stacking along the 3'→5' axis.*

The 5'-conformational transmission effect also entails a rotation around the C3' - O3' (ε) bond of the A' residue.[15] Such a rotation is not encountered for our branched RNAs. The most explicit data on the conformations of the C3' - O3' (ε) and C2' - O2' (ε') are available for trimer **1** and tetramer **2'** (in which U+1 is replaced by a cytidine residue, which does virtually not affect the molecular conformation), for which high resolution ^{13}C spectra were recorded.[6p] These spectra provided the $^{3}J_{PC}$ coupling constants $^{3}J_{P2'-C1'}$ $^{3}J_{P2'-C3'}$ $^{3}J_{P3'-C2'}$ and $^{3}J_{P3'C4'}$. For trimer **1**, these values were respectively 7.6, 3.0, 6.6, and < 1 Hz, showing a preference for the (ε^-, ε'^-) conformation.[6p,20] For tetramer **2'** , we found: 2.8, 5.6, 4.6, and 2.0 Hz, respectively, leading to the conclusion of a (ε^-,ε'^t) conformation.

(D) *2D NOESY experiments.* We recorded 2D NOESY spectra of pentamer **3** and heptamer **4** at a sample temperature of 292 K, using mixing times (τ_m) of 300 and 900 ms. For heptamer **4**, a 2D NOESY experiment was also run at 303 K, and τ_m = 500 ms. Both 2D NOESY spectra of pentamer **3** showed only intraresidual NOE contacts. From these data we inferred *syn* orientation of the A base around the glycosidic bond (NOE between H8A and H1'A), *anti* orientation of the G base (NOEs H8G - H2'G and H8G - H1'G), and *anti* orientation of the U and C bases (NOEs: H6 - H2', and H6 - H3').

Interestingly, the 2D NOESY spectra of heptamer **4**, recorded at 292 K, also showed some *inter* residual NOE contacts, besides the expected intraresidual NOEs. The intraresidual NOEs of **4** (Figure 11) clearly revealed that the A base is in the *anti* orientation (NOEs H8A - H2'A and H8A - H1'A), the G base is in the *syn* range (NOE only between H8G and H1'G), and the five pyrimidine bases are all *anti* (NOEs between H6 and H2', H3'). Interestingly, four interresidual NOEs were observed in the aromatic spectral region 8.2 - 7.4 ppm, and for a sample temperature of 292 K (Figure 11). The first interresidual crosspeak (a) reveals regular stacking between C-2 and U-1. The second one (b) points out that a unique spatial proximity exists beween the remote bases U-1 and U+2. The third interresidual NOE (c) shows stacking between the A base and U+1, while it must be concluded from the fourth crosspeak (d) that A is also involved in stacking with next neighbour C+2. It should be noted that the four interresidual NOE crosspeaks are visible in both 2D NOESY spectra that were recorded at 292 K (τ_m 300 and 900 ms). A slight increase of the sample temperature to 303 K, however, leads to total

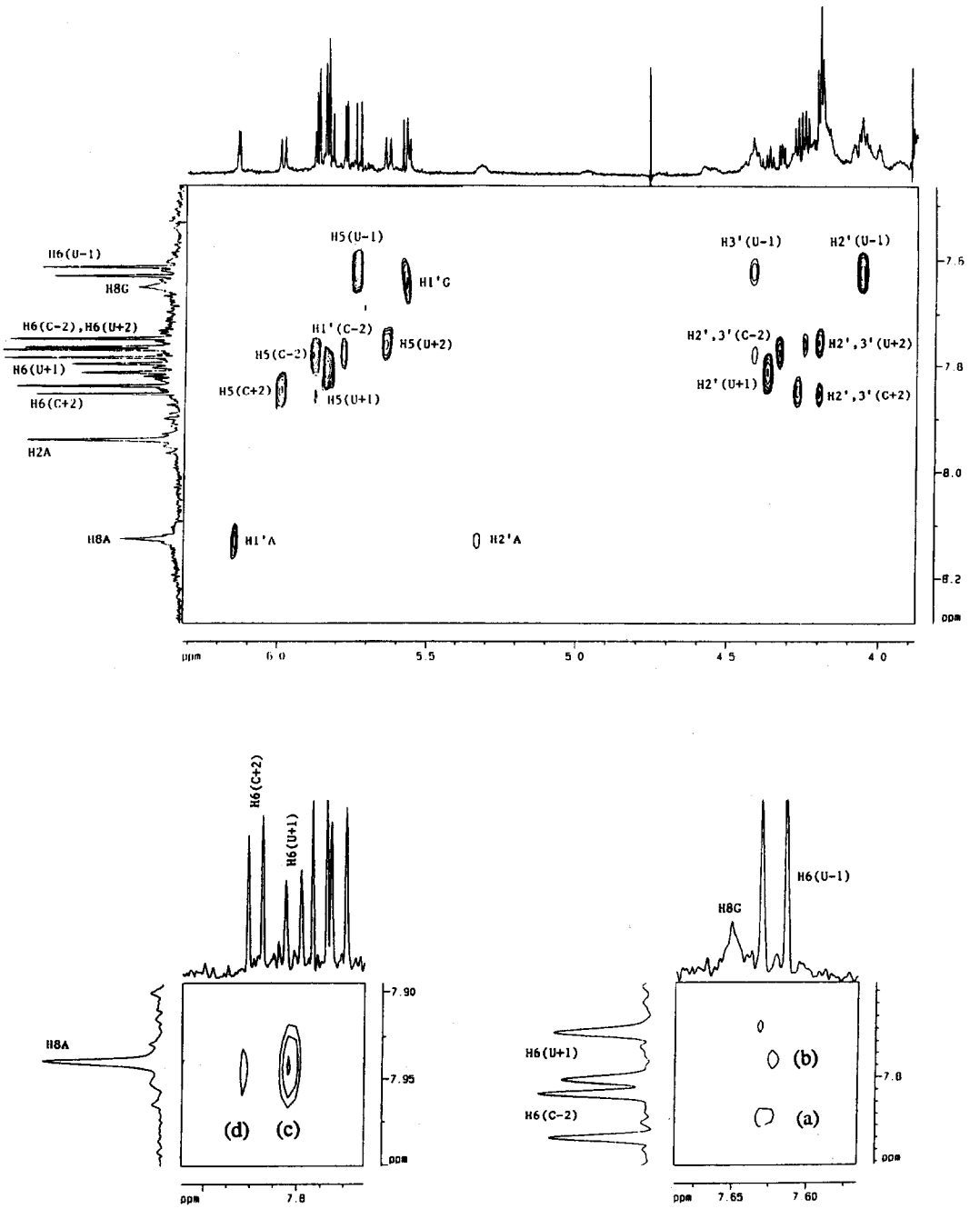

Figure 11. Expansion of the 2D-NOESY spectrum of heptamer **4** (sample temperature 19 °C, oligomer concentration approximately 7mM, mixing time 900 ms). Top: NOE contacts between the aromatic and sugar protons (intraresidual). Bottom: the four interresidual NOE contacts in the anomeric spectral region (see text).

disappearance of these signals (2D NOESY experiment with τ_m 500 ms). This shows that the secondary structure of heptamer **4** must be rather loose, which agrees well with our J-coupling data (vide supra).
We have integrated all experimental data on **4** (J couplings, $\Delta\delta$ vs. temperature experiments, 2D NOESY) into a dynamic preliminary structural model that is shown in Figure 10D. In fact, this rough model comprises two structures, one with A→U(+1) → C(+2) stacking, and the other with stacking between A and C+2, with U+1 being bulged out. *The most characteristic feature of this model is that it incorporates a backfolding of the 2'-linked chain in such a way that U+2 is stacked to U-1, as is indicated by 2D NOESY.* Thus, U+2 becomes sandwiched in between the base planes of U-1 and G. This unique structure is compatible with the NOE data, and also with the $\Delta\delta$ vs. temperature experiments. The backfolding could be facilitated via the *syn* orientation of the G base. An independent piece of support for the proposed model was obtained from variable temperature ^{31}P NMR spectra of **4**. These showed that the ^{31}P chemical shift of the 2'→5' phosphodiester group has a $\Delta\delta$ of approximately 0.60 ppm, which indicates considerable constraint around the 2'→5' axis in the heptamer structure, while no stacking occurs between A and G. Clearly, these ^{31}P NMR observations are fully in line with the model as shown in Figure 10D. Spatial proximity of U-1 with U+2 must be regarded as a unique structural feature of the heptamer structure; in tetramer **2** we have previously observed a proximity of U-1 and G. It remains to be seen how the stacking interactions of U-1 will change when the 2'-linked chain is elongated further. Current work in our laboratory is therefore aimed at the synthesis of larger branched RNA structures.

CONCLUDING REMARKS

Our 500 MHz ^1H NMR study of pentamer **3** and heptamer **4** has provided new conformational data on branched RNA structures. The experimental data that we obtained for pentamer **3** and heptamer **4** point at distinct conformational properties. Interestingly, the conformational differences between (trimer **1**, pentamer **3**), and (tetramer **2**, heptamer **4**) constitute an entirely new example of 5'-conformational transmission dictating RNA conformation.We found that the introduction of at least one residue at the 5'-site of the branch residue changes the conformation at the branch-point from [C2'-*endo* (S) sugar, *syn* orientation of the A base, and A2'→ 5'G base stacking] toward [C3'-*endo* (N) sugar, *anti* orientation of the A base, and disruption of A2'→5'G stacking]. It has been shown that the 5'-conformational transmission model is of great interest for our understanding of the conformational properties of branched RNAs, although other typical features, such as the possible occurrence of 2'→5' base stacking and backfolding (e.g. C+2 toward A in pentamer **3**, and U+2 toward U-1 in heptamer **4**) are of great importance as well. The scarce observation of interresidual NOE contacts due to the inherently mobile single stranded nature of branched RNA, however, precludes the formulation of detailed structural models, which would normally be based on NMR-derived proton-proton distance constraints, complemented with fully-relaxed NOESY matrix calculations followed by energy-minimization by appropriate molecular mechanics programs such as AMBER or CHARMm. However, our data have led to some important conclusion regarding the structural models for pentamer **3** and heptamer **4**. This study of four branched RNA models has enabled us to group them into two conformational categories: (trimer **1**, pentamer **3**), and (tetramer **2**, heptamer **4**) based on the characteristic *presence* of A2'→5'G-stack in the former, and the *absence* of this A2'→5'G-stack in the latter. It is the 5'-U (i.e. U-1) that acts as a switch for the conformational transition from the A2'→5'G-stacked structures in **1** and **3** to non-A2'→5'G-stacked structures such as **2** and **4**. Since any realistic description of the branch-point geometry in any model branched RNA molecule should include nucleotide residues at all three ends (i.e. 2'-, 3'- & 5'-terminii) of the branch-point A, it is therefore the tetramer **2** which constitutes the simplest branched RNA

model while the trimer **1** and pentamer **3** do *not*. Clearly, as the branched RNA model tetramer **2** grows to be the heptamer **4**, the core branch-point structure stabilizes remarkably. It is clear that the conformational description for the heptamer **4** *represents the free-energy minimized structure of the final-product of the splicing reaction.* A perusal of the events of the splicing reaction however shows that the lariat structure is formed at the penultimate step which involve both the bond-breaking between 3′-exon and the lariat and bond-formation (ligation) between two exons. Since no external energy such ATP or GTP is required in Group II splicing reaction, and, in total, two bonds are broken and two bonds are formed, hence the total energy expenditure is equal to zero, but the free energy of activation would still be required for the transeserification reactions. It is most likely that the free-energy of activations in the third and the fourth transesterification reactions come from within the branch-RNA moiety due to the cascade of conformational transitions from a higher energy state to a lower energy state originated at or around the branch-point. We contend that the heptamer **4**, *modelling the lariat at the penultimate step of ligation,* adopts a higher free energy transition state with the help of Mg^{2+} that mimicks the A2'→5'G-stacked state found in the conformationally locked structure such as the pentamer **3**. Now the conformational transition in the heptamer **4** from *the higher free energy A2'→5'G-stacked state in the penultimate step of ligation to the free energy minimum A2'→5'G-destacked state (found after ligation)* should provide the extra energy required as the free energy of activation in the third transesterification reaction (Conformation-driven-energy-pump), while the free energy of activations for the first and second transesterification reactions presumably come from the shift of stabilization ⇄ destabilization equilibrium in hydrogen bonding in double stranded intron-RNA part.

MATERIALS & METHODS

Samples. Synthesis of the branched RNA systems **3** and **4** was reported earlier by us; the present study was conducted with the original samples which had been stored at -20 °C in lyophilized form. Purity was assured to be > 95 %, based on ^1H NMR. A 4 mg sample of pentamer **3** was lyophilized several times from 99.8 % D_2O, and finally taken up in 0.6 mL of 99.95 % D_2O. On this sample (oligomer concentration approximately 4 mM), we recorded all one- and twodimensional NMR spectra that are presented in this work. The δ vs. temperature experiments on **3**, however were run with a slightly diluted sample with an oligomer concentration of approximately 3 mM. Analogously, we used a 7 mg sample of heptamer **4**, leading to an oligomer concentration of approximately 7 mM for all one- and twodimensional experiments. The δ vs. temperature experiments on **4** were run on a diluted sample with an oligomer concentration of approximately 3 mM.

NMR Spectroscopy. All ^1H NMR spectra were run at 500 MHz on a Bruker AMX 500 spectrometer, interfaced to an X-32 computer system. A trace of dry acetonitril was used for internal δ referencing; δ (CH_3CN) = 2.000 ppm. ^{31}P chemical shifts are relative to 3',5'-cyclic AMP as an external reference (δ -2.1 ppm).

The following 2D NMR techniques were employed:

(i) *Homonuclear Hartmann-Hahn Spectroscopy (HOHAHA).*[7] These spectra were recorded in the phase-sensitive mode according to the method of Bax et al. Thus, the MLEV17 sequence is applied for mixing, and two different power levels were used for exchange and spin-lock. A mixing time of 350 ms was used is all our HOHAHA spectra. During each experiment, we collected 512 spectra of 2K, using a sweep width of 4000 Hz. Quadrature detection in t_1 was achieved with the TPPI method.[21] The t_1 domain was zero-filled to 2K, and the FIDs were multiplied with a sine-bell window, before Fourier transformation.

(ii) *Double Quantum Filtered COSY.* [8] These experiments were performed in the phase-sensitive mode at 292 K. During the relaxation delay period, i.e., prior to the DQF COSY pulse sequence, we selectively irradiated the

HDO residual peak. The DQF COSY experiments on **3** and **4** were run in pairs, i.e., with and without ^{31}P broad band decoupling during the relaxation delay period. In each case, we collected 512 spectra of 4K, using a sweep width of 4000 Hz. Quadrature detection in t_1 was achieved using the TPPI method.[21] The multiple quantum evolution period was 3 msec. The t_1 domain was zero-filled to 2K, and apodization with a sine-bell window was applied in both dimensions before Fourier transformation.

(iii) *Proton-phosphorus Correlated Spectroscopy.*[9] These experiments were run in the phase-sensitive mode using ^1H detection. The correlation is based on heteronuclear zero- and double quantum coherence. We 256 t_1 increments; for each t_1 we accumulated 80 FIDs (2K, sweep width 4000 Hz). Quadrature detecttion in t_1 was realized by application of the TPPI method.[21] The t_1 domain was zero-filled to 2K, and apodization with a square sine-bell window was applied in both dimensions before Fourier transformation.

(iv) *Nuclear Overhauser Enhancement Spectroscopy (NOESY).* We recorded five phase-sensitive 2D NOESY spectra on pentamer **3** and heptamer **4**, using different sample temperatures and mixing times (τ_m). On **3**: τ_m = 300 ms, T = 292 K; τ_m = 900 ms, T = 292 K. On **4**: τ_m = 300 ms, T = 292 K; τ_m = 900 ms, T = 292 K; τ_m = 500 ms, T = 303 K. In all these experiments we used 512 t_1 increments (2K data points, sweep width 4000 Hz). The residual HDO peak was selectively irradiated during the relaxation delay period. For each t_1 value we accumulated 88 FIDs. Quadrature detection in t_1 was achieved using the TPPI method.[20] Data manipulation comprised zero-filling to 2K in the t_1 dimension, multiplication of the FIDs with a squared sine-bell window in both dimensions, and Fourier transformation.

Acknowledgements

Authors thank Swedish Board for Technical Development and Swedish Natural Science Research Council for generous financial support. Authors also thank Wallenbergs Stiftelse, University of Uppsala and Swedish Research Council (FRN) for funds toward the purchase of 500 MHz NMR spectrometer. Financial support from the European Molecular Biology Organization (EMBO) through a two-year EMBO fellowship to L.H.K. is gratefully acknowledged.

References and Notes

1. (a) T.R. Cech, A.J. Zaug, P.J. Grabowski, *Cell* **27**, 487 (1981). (b) M.M. Konarska, P.J. Grabowski, R.A. Padgett, P.A. Sharp, *Nature (London)* **313**, 552 (1985). (c) T.R. Cech, B.L. Bass, *Ann. Rev. Biochem.* **55**, 599 (1986). (d) P.A. Sharp, *Science* **235**, 766 (1987). (e) T.R. Cech, *Angew. Chemie, Int. Ed. Engl.* **29**, 749 (1990).
2. F. Michel, B. Dujon, *EMBO Journal* **2**, 33 (1983).
3. (a) B.L. Bass, T.R. Cech, *Nature (London)* **308**, 820 (1984). (b) C.Guerrier-Takada, K. Gardiner, T. Marsh, N Pace, S. Altman, *Cell* 35, 849 (1983). (c) G. Garriga, A.M. Lambowitz, *Cell* **39**, 631 (1984). (d) G. van der Horst, H. Tabak, *Cell* **40**, 759 (1985). (e) K. Ehrenman, J. Pedersen-Lane, D. West, R. Herman, F. Maley, M. Belfort, *Proc. Natl. Acad. Sci. USA* **83**, 5875 (1986).
4. (a) C.L. Peebles, P.S. Perlman, K.L. Mecklenburg, M.L. Petrillo, J.H. Tabor, K.A. Jarell, H.-L. Cheng, *Cell* **44**, 213 (1986). (b) R. van der Veen, A.C. Arnberg, G. van der Horst, L. Bonen, H. Tabak, L.A. Grivell, *Cell* **44**, 225 (1986).
5. (a) E. Brody, J. Abelson, *Science* **228**, 963 (1985).(b) P.J. Grabowski, S.R. Seiler, P.A. Sharp, *Cell* **42**, 345 (1985). (c) T. Marriatis, R. Reed, *Nature(London)* **325**, 673 (1987).
6. (a) J.-M. Vial, N. Balgobin, G. Remaud, A. Nyilas, J. Chattopadhyaya, *Nucleosides Nucleotides* **6**, 209 (1987). (b) G. Remaud, X.-X. Zhou, B. Oberg, J. Chattopadhyaya, *Rev. Heteroat. Chem.* **1** (S. Oae, Ed.), MYU Publishing Inc., Tokyo, p 340 (1988). (c) X.-X. Zhou, A. Nyilas, G. Remaud, J. Chattopadhyaya, *Tetrahedron* **43**, 4685 (1987). (d) X.-X. Zhou, G. Remaud, J. Chattopadhyaya, *Tetrahedron* **44**, 6471 (1988). (e) N. Balgobin, A. Foldesi, G. Remaud, J. Chattopadhyaya, *Tetrahedron*

44, 6929 (1988) (f) A. Foldesi, N. Balgobin, J. Chattopadhyaya, *Tetrahedron Lett.* **30** 881 (1989). (g) G. Remaud, J.-M. Vial, A. Nyilas, N. Balgobin, J. Chattopadhyaya, *Tetrahedron* **43**, 947 (1987). (h) J.-M. Vial, G. Remaud, N. Balgobin, J. Chattopadhyaya, *Tetrahedron* **43**, 3997 (1987). (i) L.H. Koole, N. Balgobin, H.M. Buck, W.H.A. Kuijpers, A. Nyilas, G. Remaud, J. Chattopadhyaya, *Recl. Trav. Chim. Pays-Bas* **107** 663 (1988). (j) X.-X. Zhou, A. Nyilas, G. Remaud, J. Chattopadhyaya, *Tetrahedron* **44**, 571 (1988). (k) A. Sandstrom, G. Remaud, J.-M. Vial, X.-X. Zhou, A. Nyilas, N. Balgobin, J. Chattopadhyaya, *J. Chem. Soc., Chem. Commun.* 542 (1988). (l) G. Remaud, N. Balgobin, C. Glemarec, J. Chattopadhyaya, *Tetrahedron* **45**, 1537 (1989). (m) G. Remaud, N. Balgobin, A. Sandstrom, J.-M. Vial, H.M. Buck, A.F. Drake, X.-X. Zhou, J. Chattopadhyaya, *J. Biochem. Biophys. Meth.* **18**, 1 (1989). (n) L.H. Koole, X.-X. Zhou, G. Remaud, H.M. Buck, J. Chattopadhyaya, *J. Chem. Soc., Chem Commun.*, 759 (1989). (o) G. Remaud, J.-M. Vial, N. Balgobin, L.H. Koole, A. Sandstrom, A.F. Drake, X.-X. Zhou, C. Glemarec, J. Chattopadhyaya In "Structure and Methods, RNA and DNA", Vol. 3, R.H. Sarma and M.H. Sarma, Eds., Adenine Press, p. 319 (1990). (p) C. Glemarec, M. Jaseja, A. Sandstrom, L.H. Koole, P. Agback, J. Chattopadhyaya, *Tetrahedron* ..**47**., preceding paper in this issue.

7. (a) D.G. Davies, A. Bax, *J. Am. Chem. Soc.* **107**, 2802 (1985). (b) D.G. Davies, A. Bax, *J. Am. Chem. Soc.* **107**, 7197 (1985).
8. (a) D. Neuhaus, G. Wagner, M. Vasak, J.H.R. Kagi, K. Wuthrich, *Eur. J. Biochem.* **151**, 257 (1985). (b) K. Wuthrich In "NMR of Proteins and Nucleic Acids", Wiley, New York (1986). (c) C. Cheong, G. Varani, I. Tinoco, Jr. *Nature (London)* **346**, 680 (1990).
9. A. Bax, R.H. Griffey, B.L. Hawkins, *J. Magn. Res.*, **55**, 301 (1983).
10. See, for instance: (a) C. Altona, Recl. Trav. Chim. Pays-Bas 101, 413 (1982). (b) D.B. Davies In "Progress in NMR Spectroscopy", J.W. Emsley, J. Feeney, L.H. Sutcliffe, Eds., Vol. 12, p. 135 (1978).
11. Two papers on the conformational properties of branched RNA trimers have been published by other workers: (a) M.J. Damha, K.K. Ogilvie, *Biochemistry* **27**, 6403 (1988). (b) M. Lee, S. Huss, G. Gosselin, J.-L. Imbach, J.A. Hartley, J.W. Lown, *J. Biomol. Struct. Dyn.* **5**, 651 (1987).
12. J. Doornbos, J.A.J. den Hartog, J.H. van Boom, C. Altona, *Eur. J. Biochem.*, **116**, 403 (1981).
13. C. H. Lee, I. Tinoco, Jr. *Biophysical Chem.*, **11**, 283 (1980).
14. M.P. Stone, D.L. Johnson, P.N. Borer, *Biochemistry* **20**, 3604 (1981).
15. (a) Y. Th. van den Hoogen, S.J. Treurniet, H.C.P.F. Roelen, E. de Vroom, G.A. van der Marel, J.H. van Boom, C. Altona, *Eur. J. Biochem.* **171**, 155 (1988). (b) Y. Th. van den Hoogen, PhD. Thesis, University of Leiden, Netherlands, 1988.
16. E. Shefter, M. Barlow, R.A. Sparks, K.N. Trueblood, *Acta Cryst.* **B25**, 895 (1969).
17. R. Parthasarathy, M. Malik, S.M. Fridey, *Proc. Natl. Acad. Sci. USA* **79**, 7292 (1982).
18. It should be noted that the X-ray crystal structures in refs 16 and 17 show intramolecular base stacking only to a small extent. This may be due to the fact that both dinucleotides systems crystallized in a protonated form.
19. M.M. Dhingra, R.H. Sarma, *Nature (London)* **272**, 798 (1978). (b) N.S. Kondo, H.M. Holmes, L.M. Stempel, P.O.P. Ts'O, *Biochemistry* **9**, 3479 (1970).
20. P.P. Lankhorst, C.A.G. Haasnoot, C. Erkelens, C. Altona, *J. Biomol. Struct. Dyns.* **1**, 1387 (1984).
21. D. Marion, K. Wuthrich, *Biochem. Biophys. Res. Commun.*, **113**, 967 (1983).