A Comparative High-field N.M.R. Study on Tri-, Tetra-, Penta-, and Hepta-mer Ribonucleotides modelling the Branch Site of Group II and Nuclear Messenger RNA spliced Lariat Introns

L. H. Koole,[a] G. Remaud,[b] X.-X. Zhou,[b] H. M. Buck,[a] and J. Chattopadhyaya[b]

[a] *Department of Organic Chemistry, Eindhoven University of Technology, The Netherlands*
[b] *Department of Bio-organic Chemistry, Uppsala University, Uppsala, Sweden*

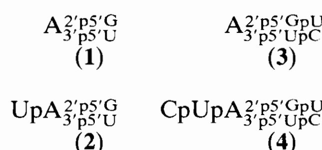
Conformational analysis of the branched RNA systems (**1**)—(**4**) using variable-temperature 600 MHz proton and 223 MHz phosphorus-31 n.m.r. spectroscopy revealed that (i), generation of a lariat structure, typical for Group II and nuclear messenger RNA splicing, can occur with preservation of the A-geometry of the RNA strand; (ii), mimicry of the lariat branch site requires the presence of at least one nucleotide linked to the 5′-site of the branch point [as in tetramer (**2**) and heptamer (**4**)], since the absence of this nucleotide [as in trimer (**1**) and pentamer (**3**)] results in an unnatural molecular conformation which is largely dominated by adenine 2′–5′ guanine base stacking.

RNA splicing entails excision of non-coding fragments (introns) from pre-messenger RNA, and subsequent ligation of coding fragments (exons) to form a functional messenger RNA as a template for protein biosynthesis.[1] In the case of Group II or nuclear messenger RNA splicing, the introns are released as *lariat* structures.[2] Chemical and enzymatic analyses of these lariats have shown that the branch point is inevitably adenosine (A), which is linked to guanosine (G) *via* a 2′–5′ phosphodiester, and to cytidine (C) or uridine (U) *via* a 3′–5′ phosphodiester group. As the first phase in answering the question why RNA lariats are formed in splice reactions, we have recently synthesized a variety of branched RNA models.[3] For some of these systems on the trimer and tetramer level, the conformational properties in solution were studied using high field n.m.r. and circular dichroism.[4—6] In this communication, we report the first results of a structural comparison of the branched tri-, tetra-, penta-, and hepta-mer RNAs (**1**)—(**4**), which correspond to the branch site of the Group II intron bI1 from yeast mitochondria (*cf.* Figure 1). Structural information was obtained from high field variable temperature proton (600 MHz), and phosphorus-31 (223 MHz) n.m.r. spectroscopy.

Table 1. Experimental values of $J_{1'2'}$ at different sample temperatures for the branch point (A), and the 2′–5′ linked G residue in (**3**) and (**4**),[a] along with the calculated populations of the North ribose conformation [x(North)], and ratios x(North)/x(South).[b]

T/°C	$J_{1'2'}$/Hz[c]	A-residue x(North)/%	x(North)/x(South)	$J_{1'2'}$/Hz[c]	G-residue x(North)/%	x(North)/x(South)
			Compound (**3**)			
4	4.92	42	0.73	4.33	51	1.04
20	5.21	38	0.62	4.62	47	0.88
40	5.47	34	0.52	4.71	45	0.83
60	5.48	34	0.52	4.80	44	0.79
80	5.46	34	0.52	4.96	42	0.72
			Compound (**4**)			
8	—[d]	—	—	5.48	34	0.52
20	2.40	79	3.85	5.52	34	0.52
40	3.20	68	2.09	5.71	31	0.44
60	3.39	65	1.85	5.88	28	0.39
80	4.07	55	1.22	6.04	26	0.35

[a] These data are a selection of all available information; $J_{1'2'}$ was actually measured at temperature intervals of 5 °C. [b] Formula used: $x(\text{North}) = 100 \times [J_{1'2'}(\text{exp}) - J_{1'2'}(\text{South})]/[J_{1'2'}(\text{North}) - J_{1'2'}(\text{South})]$. With $J_{1'2'}$(North) = 1.0 Hz, and $J_{1'2'}$(South) = 7.8 Hz, the relation $x(\text{North}) = 100 \times [7.8 - J_{1'2'}(\text{exp})]/6.8$ is obtained. [c] The accuracy of $J_{1'2'}$ values is estimated to be 0.03 Hz (spectral window 2000 Hz, 64 K data points). The spectra were recorded on the Bruker AM 600 spectrometer of the Dutch National high-field n.m.r. facility at Nijmegen. Sample temperature was controlled using a carefully calibrated thermocouple system with an accuracy of 1 °C. [d] Could not be determined accurately owing to severe spectral crowding.



Concerning proton n.m.r. analysis, we have restricted ourselves to the resonances which could be assigned unequivocally for all 4 systems. These are the doublets of $H_{1'}$(A) and $H_{1'}$(G), and the singlets of H_8(G), H_2(A), and H_8(A). From the signals of $H_{1'}$(A) and $H_{1'}$(G) we determined the spin–spin coupling constants $J_{1'2'}$ (Table 1). These parameters provided essential information on the conformation of the respective ribose rings, which are involved in a rapid $C_{2'}$-*endo* ⇄ $C_{3'}$-*endo* equilibrium. Figure 2 shows the ratio $x(C_{3'}\text{-}endo)/x(C_{2'}\text{-}endo)$ for the ribose ring at the branch point, plotted as a function of temperature. The data points of trimer (**1**)[4] and pentamer (**3**) practically fit in the same curve, showing that a preference exists for $C_{2'}$-*endo* ribose. The data for tetramer (**2**)[5] and heptamer (**4**), on the other hand, reveal a conformational bias towards $C_{3'}$-*endo* ribose; raising the sample temperature leads to a clearly diminished preference for $C_{3'}$-*endo*. It should be mentioned that the systems (**1**)—(**4**) show virtually the same ribose conformation for the G-residue [slight preference for $C_{2'}$-*endo*;[4,5] data on (**3**) and (**4**) not shown].

Figure 3 shows the chemical shift (δ) *versus* temperature profiles of H_8(G), H_2(A), and H_8(A) measured for (**3**) and (**4**). Comparison with the data on (**1**)[4] and (**2**)[5] again shows a close resemblance for the systems (**1**) and (**3**) [slightly decreasing δH_8(A) and markedly increasing δH_2(A) profiles] on one hand, and (**2**) and (**4**) [moderately increasing δ-profiles for H_8(G), H_2(A), and H_8(A)] on the other hand. At this point, we recall previous experiments on (**1**) and related branched trimer systems.[4] Primarily on the basis of δ *versus* temperature profiles, it was deduced that adenine 2′–5′ guanine base stacking actually dictates the conformation of trimer systems.[4] By contrast, tetramer (**2**) and other tetramers showed an A RNA geometry with U3′–5′A3′–5′U base stacking and a more or less free (*i.e.* unstacked) guanine base.[5] We can now conclude that adenine 2′–5′ guanine base stacking also has a great impact on the conformation of pentamer (**3**). *Mutatis mutandis*, heptamer (**4**) most probably assumes an A RNA helical structure dominated by C3′–5′U3′–5′A3′–5′U3′–5′C stacking and a dangling G3′–5′U side chain.

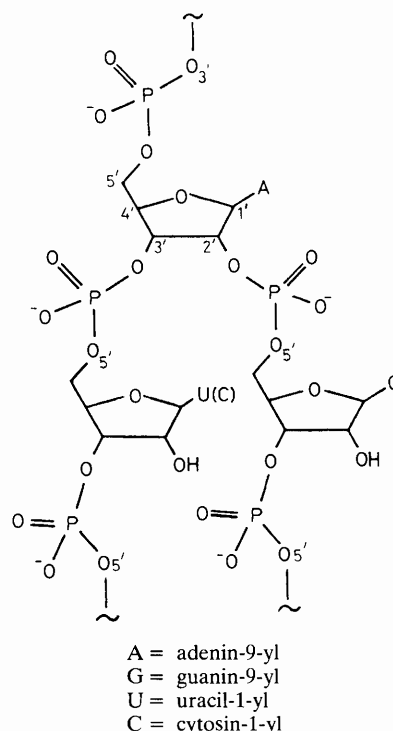


Figure 1. Structure of the branch site of Group II and nuclear messenger RNA spliced lariat introns.

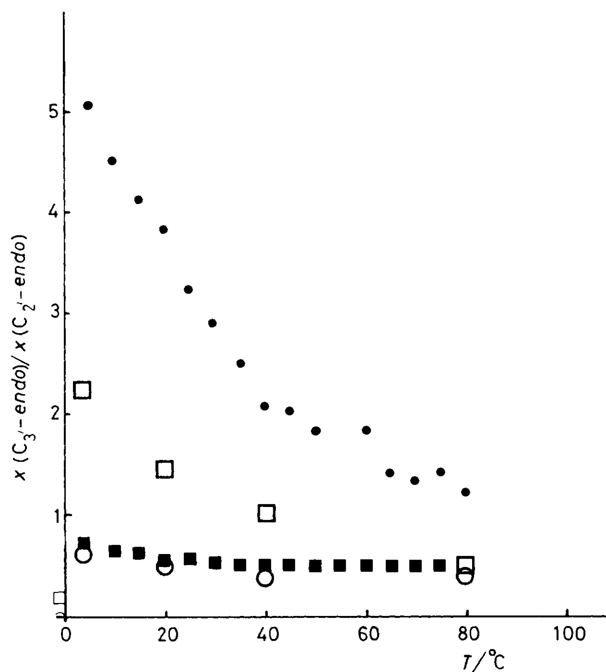


Figure 2. Ratio $x(C_{3'}\text{-}endo)/x(C_{2'}\text{-}endo)$ for (1) (○), (2) (□), (3) (■), and (4) (●), plotted as a function of the sample temperature.

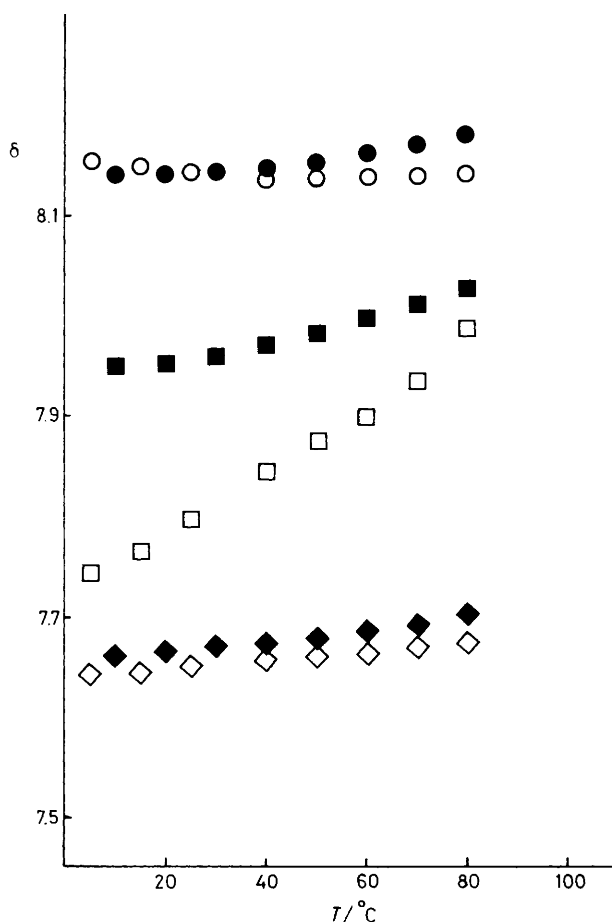


Figure 3. Chemical shift (δ) *versus* temperature profiles for $H_8(G)$, $H_2(A)$, and $H_8(A)$, measured for (3) and (4). (3): $H_8(G)$ (◇), $H_2(A)$ (□), $H_8(A)$ (○); (4): $H_8(G)$ (◆), $H_2(A)$ (■), $H_8(A)$ (●).

Evidence in support of these conclusions was inferred from variable temperature ^{31}P n.m.r. spectroscopy focused on the resonances of the A2′–5′G and A3′–5′U phosphodiester groups.[7] For pentamer (3), it was found that the parameter $\Delta\delta(^{31}P)$, *i.e.*, the difference between $\delta(^{31}P)$ values at 8 and 85 °C,[7] amounts to 0.63 and 0.28 p.p.m. for the 2′–5′ and 3′–5′ phosphate groups, respectively. This implies that the 2′-linked phosphate is far more constrained than the 3′-linked phosphate,[7] consistent with adenine 2′–5′ guanine base stacking, as was also deduced from the proton n.m.r. analysis (*vide supra*). For the heptamer (4), only the A2′–5′G phosphate resonance could be assigned; the corresponding $\Delta\delta(^{31}P)$ was found to be 0.38 p.p.m. Since this value is even lower than the value of 0.50 p.p.m. measured previously for tetramer (2),[7] we conclude that the 2′-linked phosphate is relatively flexible, *i.e.*, there is no apparent adenine 2′–5′ guanine stacking in the heptamer system. This fits with the A RNA geometry which was proposed on the basis of proton n.m.r. data (*vide supra*).

The present results show for the first time that formation of a lariat *via* generation of a branch with an A2′–5′G phosphodiester linkage (as in Group II or nuclear messenger RNA splice reactions) can occur without substantial distortion of the conformation around the branch point. In order to mimic the branch, it is apparently necessary to have at least one nucleotide linked to the 5′-site of the branch point. Thus, tetramer (2) is a representative model for the local conformation of the branch, whereas pentamer (3) is not. Comparing pentamer (3) with tetramer (2), we are faced with the remarkable fact that a larger fragment of the naturally occurring branched system provides a poorer picture of the molecular structure. Of course, the heptamer (4) must be regarded as the best available model. Spectroscopic studies in order to solve the molecular conformation of this system in detail are currently in progress.

Received, 28th February 1989; Com. 9/00874H

References

1 K. Jarell, C. Peeble, R. Dietrich, S. Romitri, and P. Perlman, *J. Biol. Chem.*, 1988, **263**, 3432; M. R. Green, *Nature (London)*, 1988, **336**, 716; P. Sharp, *Science*, 1987, **235**, 766; T. R. Cech and B. L. Bass, *Ann. Rev. Biochem.*, 1986, **55**, 599; R. van der Veen, A. Arnberg, G. Horst, L. Bonen, H. Tabak, and L. Grivell, *Cell*, 1986, **44**, 225; M. M. Konarska, P. J. Grabowski, R. A. Padgett, and P. A. Sharp, *Nature (London)*, 1985, **313**, 552.

2 G. A. Freyer, J. Arenas, K. K. Perkins, H. M. Furneaux, L. Pick, D. Young, R. J. Roberts, and J. Huinitz, *J. Biol. Chem.*, 1986, **262**, 4267; H. Hornig, M. Aebi, and C. Weissmann, *Nature (London)*, 1986, **324**, 589.

3 X.-X. Zhou, G. Remaud, and J. Chattopadhyaya, *Tetrahedron*, 1988, **44**, 6471; N. Balgobin, A. Foldesi, G. Remaud, and J. Chattopadhyaya, *ibid.*, 1988, **44**, 6929; X.-X. Zhou, A. Nyilas, G. Remaud, and J. Chattopadhyaya, *ibid.*, 1987, **43**, 4687; J.-M. Vial, N. Balgobin, G. Remaud, A. Nyilas, and J. Chattopadhyaya, *Nucleosides Nucleotides*, 1987, **6**, 209; G. Remaud, X.-X. Zhou, B. Oberg, and J. Chattopadhyaya, in 'Reviews in Heterocyclic Chemistry,' ed. S. Oae, M.Y.U. Publishing Inc., Tokyo, 1987.

4 G. Remaud, J.-M. Vial, N. Balgobin, and J. Chattopadhyaya, *Tetrahedron*, 1987, **43**, 947.

5 X.-X. Zhou, A. Nyilas, G. Remaud, and J. Chattopadhyaya, *Tetrahedron*, 1988, **44**, 571.

6 M. J. Damha and K. K. Ogilvie, *Biochemistry*, 1988, **27**, 6403; L. H. Koole, N. Balgobin, H. M. Buck, W. Kuijpers, A. Nyilas, G. Remaud, J.-M. Vial, and J. Chattopadhyaya, *Recl. Trav. Chim. Pays-Bas*, 1988, **107**, 663.

7 A. Sandstrom, G. Remaud, J.-M. Vial, X.-X. Zhou, A. Nyilas, N. Balgobin, and J. Chattopadhyaya, *J. Chem. Soc., Chem. Commun.*, 1988, 542.