Tetrahedron Letters, Vol. 34, No. 24, pp. 3929-3932, 1993
Printed in Great Britain

0040-4039/93 $6.00 + .00
Pergamon Press Ltd

The Self-cleavage of Lariat-RNA

Peter Agback, Corine Glemarec, Lee Yin, Anders Sandström, Janez Plavec, Christian Sund, Shun-ichi Yamakage, Garimella Viswanadham, Bertrand Rousse, Nitin Puri, & Jyoti Chattopadhyaya*

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

Abstract: Lariat-RNAs 3 and 4 undergo site-specific self-cleavage reaction at the $G^3 \rightarrow C^6/U^7$ phosphodiester bond by the nucleophilic attack of 2'-OH of G^3 sugar moiety to its 3'-phosphate to give 5'-hydroxyl terminal at C^6 or U^7 and 2', 3'-cyclic phosphodiester of G^3 whereas lariat-tetramer 1, pentamer 2, the cyclic-A(2'→5')G-tetramer 5 and the cyclic-A(3'→5')G-tetramer 6 are completely stable. The lariat-RNAs 3 and 4 are the smallest RNA known to undergo self-cleavage which is reminiscent of the RNA-hammerhead (Ribozyme) activity. The geometry of the cleavage-site in 3 and 4 has been defined by full conformational analysis by NMR and molecular dynamics calculation in water.

Most of the catalytic natural RNA molecules, including viroids, virusoids and satellite RNAs, that infect plants are large and form complex protein encapsulated tertiary structure. They undergo efficient site-specific self-cleavage[1] *in vitro*. These RNAs share a small structural domain (hammerhead) consisting of about 30 nucleotides which has been shown to be necessary and sufficient for the self-cleavage[1]. The catalytic activity of the hammerhead RNA includes the site-specific self-cleavage of a phosphodiester bond to give a 5'-hydroxyl and a 2', 3'-cyclic phosphodiester termini[1], which is clearly different from group I and II introns and RNase P, and is widespread in plant viruses and animals. We herein report on the site-specific self-cleavage reaction of small model lariat-RNAs[2] **3** and **4**, which leaves a 2',3'-cyclic phosphate at the 3'-end of G^3 and a 5'-hydroxy terminal at C^6/U^7, mimicking the hammerhead self-cleavage product[1].

The lariat-hexamer **3** exhibits two sets of proton resonances[3,5] attributed to two different forms (A-form ⇄ B-form) in slow exchange ($k = 0.1\ s^{-1}$ at 25° C) on the NMR time scale ($K = x_B/x_A = 0.1$ at 5 °C and 0.75 at 60 °C; $\Delta H = 7.1 \pm 0.6$ kcal.mol^{-1}, $\Delta S = 21 \pm 1.1$ cal.mol^{-1}.K^{-1}, $E_a = 23.8 \pm 1.8$ kcal.mol^{-1}, $\Delta G^{\#} = 18$ kcal.mol^{-1} at 25 °C)[3] whereas lariat-heptamer **4** shows only one average conformation on the NMR time scale. The lariat-hexamer **3** and heptamer **4** undergo self-cleavage reaction at room temperature whereas tetramer **1**, pentamer **2**, the cyclic-A(2'→5')G-tetramer **5** and the cyclic-A(3'→5')G-tetramer **6** do not self-cleave. The site of phosphodiester cleavage is *specific*[4] and occurs at the 3'-phosphate of the guanosine residue (G^3) to give a guanosine 2', 3'-cyclic phosphate[4] and a 5'-hydroxy terminal (5'-OH-C^6 in **3** and 5'-OH-U^7 in **4**). At 25 °C, lariat-heptamer **4** ($k = 0.16.\ 10^{-3}$ min^{-1}) cleaves six times faster than hexamer **3** ($k = 0.25.\ 10^{-4}$ min^{-1}). The rate of cleavage is temperature-dependent. Note that at higher temperature (>45°C), the rate of cleavage decreases considerably suggesting that the conformation required for the cleavage reaction only forms at narrow temperature range (10 °C < T < ~45 °C). The addition of magnesium ions did not produce any noticeable changes in the conformation, as evident by unaltered $^3J_{HH}$ couplings, despite the fact that it increased the rate of cleavage by a factor two. Thus, the coordination of Mg^{2+} to the non-bridging phosphate-oxygens might help self-cleavage by stabilizing the transition state but is not necessary for its formation. It has also been found that the triethylammonium salt of lariat-hexamer **3** did not self-cleave. The

bulky triethylammonium ion could exert some steric hindrance which inhibits or reduces the ability to form the intermediate or transition state required for self-cleavage. The important clue regarding the conclusion that the B-form of lariat-hexamer **3**[6a] is indeed involved in the self-cleavage reaction is found in its conformational similarities with that of lariat-heptamer **4**[6a] whereas the conformation of the A-form of hexamer **3**[6b] is considerably different. It is noteworthy that at low temperature, the nucleotides in the cyclic-A(2'→5')G-tetramer **5** and the cyclic-A(3'→5')G-tetramer **6** have ribose and phosphate backbone conformation[7] which are very similar to the A-form of lariat-hexamer **3**[6b] and as stated above they do not cleave. The following conclusions can be drawn on the self-cleavage reaction of the lariat-RNAs: (1) The ability to undergo site-specific self-cleavage is reminiscent of self-cleavage in hammerhead RNA structures. (2) The rate of self-cleavage depends on the size of the lariat-loop. (3) The nucleotides in the 3'-tail (U^4 or C^5), or some specific functional group(s) of the 3'-tail is(are) necessary for self-cleavage or is(are) critical for the formation of the self-cleaving transition state. (4) The conformational characteristics of the ribose ring[6] of **3** and **4** suggest that the overall conformation of the ribose-phosphate backbone may have profound influence *(vide infra)* on the facility of the self-cleavage reaction.

Since the final product of self-cleavage reaction of **3** and **4** is the same[4] as in RNA ribozyme (*i.e.* a 2',3'-cyclic phosphate and a 5'-hydroxyl termini), it is likely that in both cases the self-cleavage occurs by a transesterification reaction by the direct attack of the 2'-oxygen on the phosphorus, and by in-line displacement of the 5'-oxygen[8]. The reaction seems to occur via a trigonal bipyramidal phosphate that may be an intermediate or a transition state. No detailed chemical mechanism of the self-cleavage nor the structural information on the local conformation at the cleavage-site of the hammerhead structure is known. The presence of three helical regions have been however confirmed in three independent NMR studies through the assignment of imino protons in the hammerhead domain[9]. It is however likely that the sugar geometry and the glycosyl torsion of a trigonal bipyramidal phosphate which is *cis*-fused with the sugar moiety at the cleavage-site would be similar to the geometry of nucleoside 2',3'-O-cyclic phosphate because both suffer from the torsional constraint of the five-membered cyclic phosphate ring which forces the O2'-C2'-C3'-O3' group into near planar arrangement and favours glycosyl torsion in *syn* arrangement. The X-ray crystal structure of both nucleoside 2',3'-O-cyclic phosphates[10] and 2',3'-O-isopropylideneadenosine[10] have shown that the five-membered cyclic phosphate ring in the former and the dioxolane ring in the latter forces the O2'-C2'-C3'-O3' group into near planar arrangement which restricts the conformational freedom of the ribose moiety to one of the following forms: (1) O4'-*endo*-C4'-*exo* ($P = 82°$, $\Phi_m = 36°$), (2) O4'-*exo*-C4'-*endo* ($P = 262°$, $\Phi_m = 23°$), (3) C4'-*endo*-C3'-*exo* ($P = 215°$, $\Phi_m = 32°$), (4) C2'-*endo*-C3'-*exo* ($P = 174°$, $\Phi_m = 29°$) and (5) Planar. In our self-cleaving lariat RNAs, we indeed observe the following local geometry at the cleavage site which is consistent with what is found in the above constrained sugars: South sugar for G^3 with *syn* glycosyl bond and C^6 or U^7 with the South sugar and *anti* glycosyl torsion in the lariat-RNAs **3** and **4**. These synthetic lariat-RNAs[2] despite their small size and their slower rate of site-specific self-cleavage ($k = 0.16.10^{-3}$ to $0.25\ 10^{-4}$ min^{-1} at 25 °C) are important models for the investigation of the mechanism and conformational requirement of the fast self-cleavage reaction of the hammerhead RNA structures ($k_{cat} \geq 0.5$ min^{-1} at 37 °C)[1]. Both NMR constrained molecular dynamics[11] study in water at 284 K and subsequently free MD (altogether for 226 ps) of the heptameric lariat-RNA **4** show that both the constrained and unconstrained geometries and their potential energies are closely similar, the structures generated during these MD runs show that the cleavage-site between G^3 and U^7 has the following *local* conformation: ε^- (-87°±10°), ζ^+ (78°±10°), α^+ (97°±10°), β^t (167°±10°) and γ^+ (67°±10°). Molecular modelling studies on the

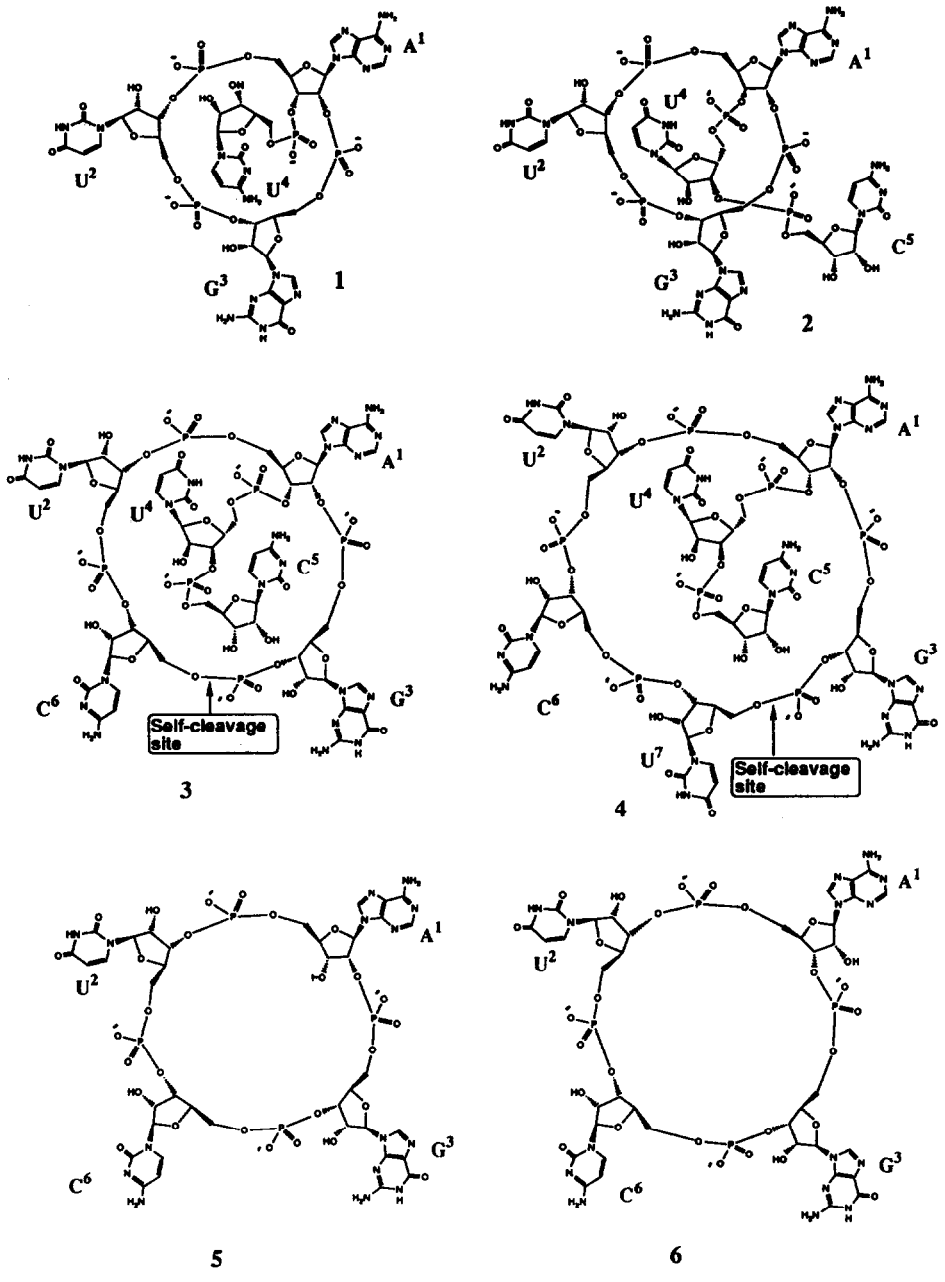

Figure 1: Synthetic lariat-RNAs **1** - **4** modelling the natural lariat counterpart which is formed as an intermediate in the RNA splicing reaction. Lariat RNAs **1** and **2** are sterically strained and are stable. Hexameric Lariat RNA **3** shows two conformations which are in slow exchange on the NMR time scale, and only one of them self-cleaves between G^3 and C^6 (shown by " → "). Heptameric lariat RNA **4** self-cleaves between G^3 and U^7 (shown by " → ") at a faster rate than **3**, while neither cyclic 2'→5' RNA **5** nor 3'→5' RNA **6** show any conformational isomerism or self-cleavage which are characteristic of hexameric lariat RNA **3**.

MD generated geometry show that a simple rotation of the *local phosphate backbone at the cleavage-site* from ε^- ($d_{O2'\text{-}3'P}$ = 3.8 Å) → ε^t ($d_{O2'\text{-}3'P}$ = 2.8 Å) and a rotation of ζ^+ → ζ^t would position the leaving 5'-termini of U^7 for a potential in-line displacement by 2'-OH of G^3 (Note that in the latter geometrical transition, α^+, β^t and γ^+ and the South-sugar of U^7 remain unchanged). Such a geometry at the cleavage site would produce an optimal *local* structure for a neighbouring nucleophilic attack by 2'-OH to give the trigonal-bipyramidal transition state/intermediate [*ab initio* optimized geometry of guanosine 2',3'-cyclic oxyphosphorane, Gaussian 92 HF/3-21G basis set: ε = +110°, ζ = -178°, α = 66°, $d_{O2'\text{-}3'P}$ = 1.8 Å], and the subsequent cleavage of the P-O(5'-U^7) bond (α-torsion) between G^3 and U^7. In an independent study[12] by the *ab initio* calculation on the cyclic oxyphosphorane as model intermediates during splicing and cleavage of RNA, it has also been shown that the ground state geometry across the α-torsion favours the α^+ mode which is consistent with our ground state geometry at the cleavage-site. We believe that this present state-of-the-art understanding of the conformational requirement of the self-cleavage reaction of RNA will be useful in the design of RNA catalyst (ribozyme).

Acknowledgements: Authors thank Swedish Board for Technical Development (NUTEK) and Swedish Natural Science Research Council (NFR), Wallenbergstiftelsen, University of Uppsala, Forskningsrådsnämnden (FRN) for financial support.

References

1. (a) Foster, A. C.; Symons, R. H. *Cell*, **1987**, *49*, 211. (b) Foster, A. C.; Symons, R. H. *Cell*, **1987**, *50*, 9. (c) Uhlenbeck, O. C. *Nature*, **1987**, *328*, 596. (d) Symons, R. H. *Trends Biochem. Sci.*, **1989**, *14*, 445. (e) Miller, W. A.; Miller, S. L. *Nucleic Acids Res.*, **1991**, *19*, 5313.
2. (a) Sund, C., Agback, P.; Chattopadhyaya, J. *Tetrahedron*, **1993**, *92*, 649.
3. K at 8 temps between 5° and 45°C, were obtained from the relative intensities of the H2A[1] at 500 MHz ^{1}H-NMR in D_2O. The ΔH and ΔS of A to B transition (chemical exchange connectivities were found between A and B forms in the NOESY spectrum) was calculated from the plot of ln(K) *vs.* 1/T. The E_a for the transition was obtained [rate (k_a) obtained from saturation transfer experiments] from the Arrhenius plot ln k_a *vs.* 1/T: $\Delta H^{\#}$ = 23 kcal.mol^{-1} and $\Delta S^{\#}$ = 15.6 cal.mol^{-1}.K^{-1}.
4. G^3H2' (δ 5.15 in the self-cleaved product of 3 & δ5.30 in the self-cleaved product of 4) and G^3H3' (δ 4.96 in the self-cleaved product of 3 & δ5.11 in the self-cleaved product of 4) are correlated with δ ^{31}P at 21.4 ppm ($J_{H2',P}$ = $J_{H3',P}$ = 8.2 Hz). HPLC (see ref 2 for conditions) of the self-cleavage reaction: R_t for 3 = 8.63 min and R_t for the self-cleavage product of 3 = 10.11 min (not separable by PAGE). ^{1}H-NMR (δ ppm in D_2O) for the self-cleavage product of 3 : A^1: 6.11 (H1'), 7.97 (H2), 8.21 (H8); U^2: 5.70 (H1'), 7.70 (H6), 5.75 (H5); G^3: 5.84 (H1'), 7.71 (H8), U^4: 5.81 (H1'), 7.76 (H6), 5.79 (H5); C^5: 5.86 (H1'), 7.77 (H6), 5.85 (H5); C^6: 5.76 (H1'), 7.84 (H6) and 5.99 (H5). ^{1}H-NMR (δ ppm in D_2O):for the self-cleavage product of 4 : A^1: 6.25 (H1'), 8.12 (H2), 8.35 (H8); U^2: 5.82 (H1'), 7.81 (H6), 5.91 (H5); G^3: 5.99 (H1'), 7.84 (H8), U^4: 5.96 (H1'), 7.91 (H6), 5.93 (H5); C^5: 6.02 (H1'), 7.97 (H6), 6.07 (H5); C^6: 6.01 (H1'), 7.98 (H6) and 6.13 (H5); U^7: 5.92 (H1'), 7.98 (H6) and 5.91 (H5). The self-cleavage product of 4 however was not separable from 4 by HPLC or PAGE.
5. The conformational analysis of 1 - 6 was performed in D_2O by measuring temperature-dependent chemical shifts of proton, phosphorous and scalar coupling constants (^{1}H-^{1}H, ^{1}H-^{31}P and ^{13}C-^{31}P by DQF-COSY, ^{1}H-^{13}C HSQC) and NOESY .
6. (a) Heptamer 4: A^1 (North-sugar, *syn*-glycosyl bond, γ^+ and β^t); U^2 (South-sugar, *anti*- glycosyl bond, γ^+, β^t and ε^-); G^3 (South-sugar, *syn*-glycosyl bond, and ε^-); C^6 (North-sugar, *anti*-glycosyl bond, γ^+ and β^t); U^7 (South-sugar, *anti*-glycosyl bond and γ^+), *which is quite similar to the conformation of B-form of hexamer 3.* (b) A-form of hexamer 3: A^1 (South-sugar, *anti*-glycosyl bond and ε^-); U^2 (South-sugar, *anti*- glycosyl bond, γ^+ and ε^-); G^3 (North-sugar, *anti*-glycosyl bond, γ^+, β^t and ε^t); C^6 (North-sugar, *anti*-glycosyl bond and ε^t).
7. (a) Lariat-tetramer 1 [A^1 (S, *anti* γ^+, β^t, ε^-); U^2 (S, *anti*, γ^+, β^t, ε^-); G^3 (S, syn, *anti*, γ or γ^t, β^t, ε^-). (b) lariat pentamer 2 [A^1 (S, *anti*, γ^+, β^t, ε^-); U^2 (S, *anti*, γ^+, β^t, ε-); G^3 (S, *anti*, γ^t or γ-, β^t). (c) Cyclic A(2'→5')G-tetramer 5 [A^1 (S, *anti*, β^t, ε^t); U^2 (S, *anti*, β^t, ε^-); G^3 (N, *syn*, γ^+, β^t, ε^t); C^6 (N, *anti*, γ^+, β^t, ε^t). (d) Cyclic A(3'→5')G-tetramer 6 [A^1 (S, *anti*, γ or γ^t, β^t, ε^t); U^2 (S, *anti*, γ^+, β^t); G^3 (N, *syn*, γ^+, β^t); C^6 (N, *anti*, γ^+, β^t, ε^t).
8. (a) van Tol, H.; Buzayan, J. M.; Feldstein, P. A.; Eckstein, F.; Bruening, G. *Nucleic Acids Res.*, **1990**, *18*, 1971. (b) Slim, G.; Gait, M. J. *Nucleic Acids Res.*, **1991**, *19*, 1183. (c) Koizumi, M.; Ohtsuka, E. *Biochemistry*, **1991**, *30*, 5145.
9. (a) Heus, H. A.; Pardi, A. *J. Mol. Biol*, **1991**, *217*, 113. (b) Odai, O.; Kodama, H.; Hiroaki, H.; Sakata, T.; Tanaka, T.; Uesugi, S. *Nucleic Acids Res.*, **1990**, *18*, 5955. (c) Pease, A. C.; Wemmer, D. E. *Biochemistry*, **1990**, *29*, 9039.
10. (a) Coulter, C. L. *J. Am. Chem. Soc.*, **1973**, *95*, 570. (b) Saenger, W. *et al. Acta Crystallogr*, **1978**, *B 34*, 1520. (c) Saenger, W.; Eckstein, F. *J. Am. Chem. Soc.*, **1970**, *92*, 4712. (d) Sundaralingam, M. *et al. Acta Crystallogr.*, **1978**, *B 34*, 2803.
11. Kollman, P *et al.*, AMBER 4.0, University of California, San Francisco **1991**.
12. (a) Taira, K. *et al Nucleic Acids Res.* **1989**, *17*, 3699. (b) Uchimaru, T. *et al FASEB J.* **1993**, *7*, 137.

(Received in UK 21 April 1993)