DOI: 10.1002/cbic.200900155

Evidence for Pseudoknot Formation of Class I preQ$_1$ Riboswitch Aptamers

Ulrike Rieder,[a] Kathrin Lang,[a] Christoph Kreutz,[a] Norbert Polacek,[b] and Ronald Micura*[a]

Riboswitches are genetic control elements within noncoding regions of mRNA.[1] They consist of a metabolite-sensitive aptamer and an adjoining expression platform. Upon binding of the metabolite to the aptamer, a structural change is induced in the expression platform and communicated as an on/off signal for gene expression. Riboswitches are currently receiving much attention since their regulation mechanism, which does not rely on the assistance of proteins, is enormously widespread among bacteria, fungi and plants.[1]

The smallest, naturally occurring known riboswitches are involved in the regulation of the biosynthesis of queuosine—a hypermodified nucleoside residing in the anticodon wobble position of certain transfer RNAs. Thereby, a stretch of only 34 nucleotides (nt) in the 5′ untranslated leader region of the respective messenger RNA specifically recognizes 7-aminomethyl-7-deazaguanine (preQ$_1$), which is an intermediate in queuosine biosynthesis.[2] Binding affinities of the ligand–aptamer complex have been reported to be in the nanomolar K_d range.[2] The minimal sequence and structure consensus refer to a hairpin comprising a 5 bp stem (P1) and a loop of 11 to 13 nucleotides (L1) together with a 3′ single-stranded nucleoside overhang, as proposed by Breaker and co-workers (Figure 1).[2] Here, we provide strong evidence that the minimal preQ$_1$ class I aptamer, in complex with preQ$_1$, forms a pseudoknot structure in solution.

The starting point for our study was the rational analysis of potential Watson–Crick base pair formation between the 3′ single-stranded nucleoside overhang and loop L1. It was remarkable that most of the annotated sequences[2] provided at least a short span of two to four consecutive nucleosides at the very 3′-end of the minimal sequence motif that was complementary to loop nucleosides, thereby meeting the demands for pseudoknot structure formation.[3]

In a first set of ^{1}H NMR experiments, we compared free aptamers (with varying loop sizes) of several preQ$_1$ motifs from *Bacillus*, *Clostridium*, *Fusobacteria*, and *Gammaproteobacteria* in response to their dedicated ligand preQ$_1$.[2,4] Indeed, upon ligand addition, all of them showed a clear increase in the number of signals in the chemical-shift region where imino proton resonances from Watson–Crick base pairs are expected (data not shown). For our further investigations, we focused on a 34 nt RNA referring to the genome of *Fusobacterium nucleatum*, which gave the best shift dispersion (Figure 2A).

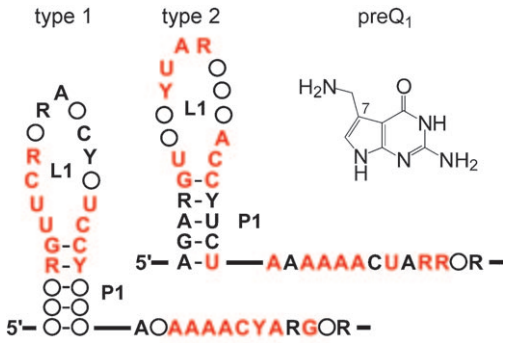

Figure 1. PreQ$_1$-responsive mRNAs are involved in the biosynthesis of the hypermodified nucleoside queuosine. Primary and secondary structure consensus models of two types of preQ$_1$ class I riboswitches (as reported by Breaker and co-workers)[2] and chemical structure of the ligand preQ$_1$. Nucleotides in red and black are more than 95 and 80% conserved, respectively. Less-conserved regions, which can vary slightly in the number of nucleotides, are represented by circles or heavy lines (R = A, G; Y = C, U). A second class of preQ$_1$ riboswitches (preQ$_1$-II) has been reported, but differs significantly in size and proposed secondary structure.[1c]

We furthermore relied on previously reported biochemical evidence that the highly conserved cytidine in position 17, which is adjacent to the loop-closing base pair G5–C18 of stem P1, most likely forms a Watson–Crick base pair with the ligand, preQ$_1$.[2] This indication stemmed from the C17U riboswitch mutant, which provided a chemical in-line probing pattern in the presence of 2,6-diaminopurine comparable to that of the wild-type sequence in the presence of preQ$_1$. This situation is moreover equivalent to that of purine riboswitches, which can change selectivity between the ligands adenine and guanine by a single mutation of their Watson–Crick binding partner within the purine aptamer scaffold.[5] To provide evidence for a possible involvement of C17 in Watson–Crick base pair formation with preQ$_1$, we synthesized an RNA that contained a site-specifically ^{15}N-labeled exocyclic amino group at residue C17. If this cytidine were involved in Watson–Crick pairing, the two hydrogen atoms of the amino group would be observed as two distinct proton resonances in a 1:1 ratio, originating from the hydrogen-bonded proton 4.1, which resonates at lower field, and the nonhydrogen-bonded proton 4.2 (Figure 2B). This was indeed observed for the preQ$_1$–aptamer complex, while only weak, broadened signals were detected for the free aptamer, due to increased proton exchange of the unpaired cytidine (Figure 2B).[6] Even more promising was the observation of the same signal behavior when we introduced a ^{15}NH$_2$-labeled cytidine at position 31 (Figure 2C). This cytidine resides within a 4 nt stretch 5′-C(31)UAG(34)-3′ that is complementary to the loop sequence portion of 5′-C(8)UAG(11)-3′, and therefore its ^{15}NH$_2$ resonances might indicate Watson–

[a] *U. Rieder, Dr. K. Lang, Dr. C. Kreutz, Prof. Dr. R. Micura*
Institute of Organic Chemistry, Center for Molecular Biosciences CMBI
6020 Innsbruck (Austria)
Fax: (+43) 507-512-2892
E-mail: ronald.micura@uibk.ac.at

[b] *Priv.-Doz. Dr. N. Polacek*
Innsbruck Biocenter, Medical University of Innsbruck
Division of Genomics and RNomics, 6020 Innsbruck (Austria)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200900155.

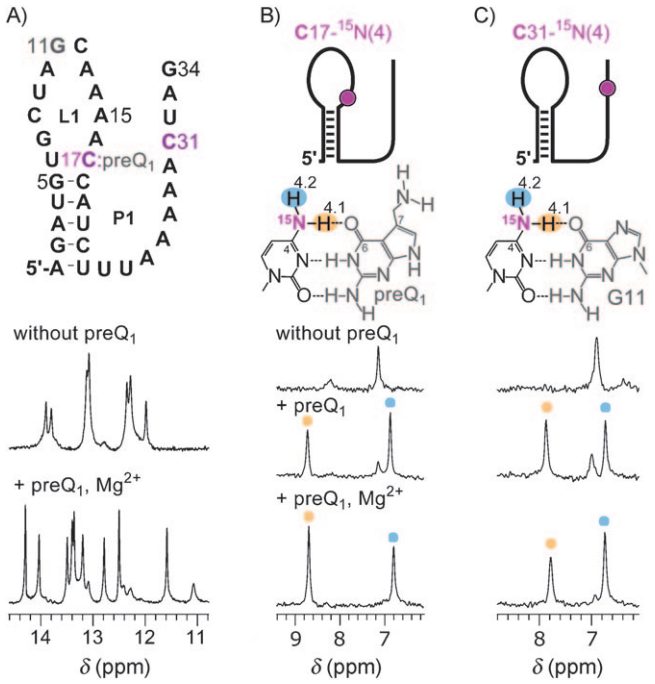

Figure 2. Minimal motif of the $preQ_1$ aptamer from *Fusobacterium nucleatum*. NMR spectra before and after addition of $preQ_1$ and Mg^{2+} ions. A) Unlabeled RNA; ^1H NMR imino proton resonances. B) Site-specifically labeled RNA with C17-$^{15}N(4)$; base pair with $preQ_1$; spectra represent the first increment of $^1H,^{15}N$ HSQC experiments. C) Site-specifically labeled RNA with C31-$^{15}N(4)$; proposed base pair with G11; first increment of $^1H,^{15}N$ HSQC. Conditions: [RNA] = 0.3 mM, [$preQ_1$] = 0.6 mM, [Mg^{2+}] = 3.0 mM, 25 mM Na_2HAsO_4, pH 6.5, H_2O/D_2O 9:1.

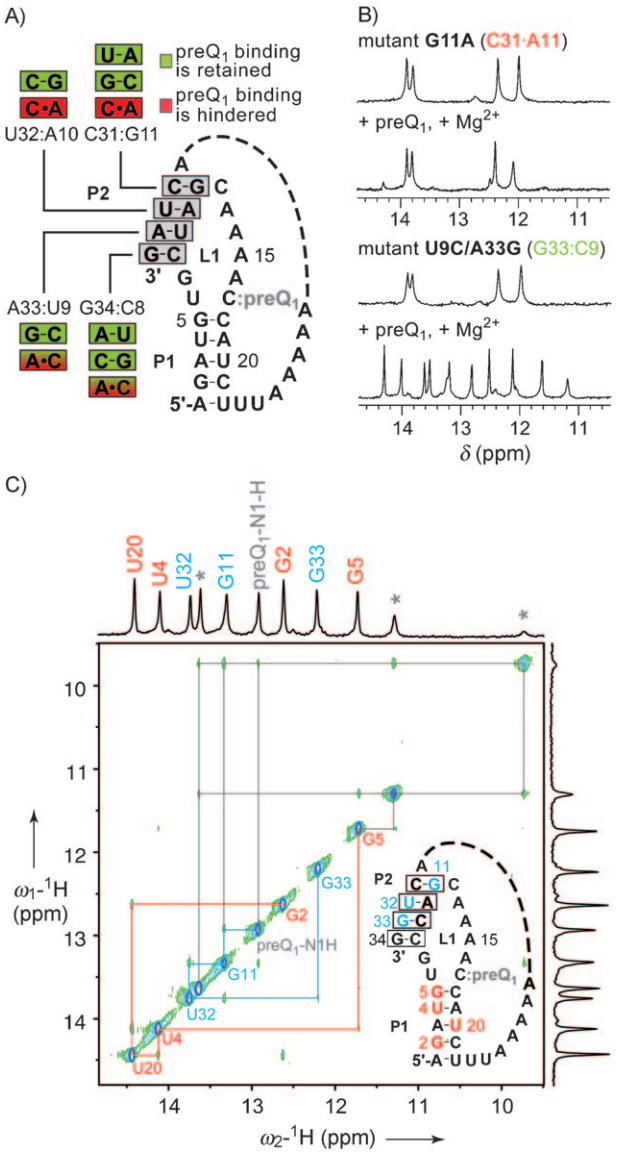

Figure 3. Mutational study of the $preQ_1$ aptamer from *Fusobacterium nucleatum*. A) Overview of mutations in pseudoknot stem P2. B) Representative ^1H NMR imino proton spectra reflecting mutant binding behavior; before and after the addition of $preQ_1$ and Mg^{2+} ions. Conditions: [RNA] = 0.2 mM, [$preQ_1$] = 0.2 (for G11A) or 0.4 mM (for U9C/A33G), [Mg^{2+}] = 2.0 mM, 25 mM Na_2HAsO_4, pH 6.5, 298K, H_2O/D_2O 9:1. C) $^1H,^1H$ NOESY spectrum of mutant aptamer U9C/A33G; preliminary assignment as indicated; asterisks indicate unassigned resonances. Conditions: [RNA] = 0.9 mM, [$preQ_1$] = 1.8 mM, [Mg^{2+}] = 4.0 mM, 25 mM Na_2HAsO_4, pH 6.5, 288K, H_2O/D_2O 9:1.

Crick base pair formation between C31 and G11, thus providing support for the hypothesis of pseudoknot formation.

Encouraged by these observations, we performed a detailed mutational analysis (Figure 3A). We systematically covaried nucleotides in each of the four proposed Watson–Crick base pairs C31–G11, U32–A10, A33–U9, and G34–C8 of the putative pseudoknot. All mutants with nucleotide covariations that retained Watson–Crick pairing capability showed binding capacity for $preQ_1$, as was deduced from their respective imino proton ^1H NMR spectra (Figure 3B and Supporting Information). This held true even for those mutants with a changed purine–pyrimidine pattern and for mutants with a double base pair covariation (Supporting Information). In contrast, single nucleobase mutations, such as G11A (Figure 3B) or U32C resulted in severely reduced binding of $preQ_1$ under the same conditions; this suggests hindrance of Watson–Crick base pair formation between loop L1 and the terminal nucleotides of the 3′ overhang. Further single-nucleobase mutations, such as U9C or G34A, also hampered binding of $preQ_1$, although to a somewhat lesser extent. Taken together, the mutational study provides strong evidence that pseudoknot formation is a requirement for a high affinity complex between $preQ_1$ and its aptamer.

One of the mutant aptamers (U9C/A33G) that replaced the respective U–A by a C–G base pair within the pseudoknot resulted in a ^1H NMR spectrum with excellent imino proton dispersion. The corresponding $^1H,^1H$ NOESY NMR spectrum of the imino proton region is depicted in Figure 3C and allowed straightforward assignment of Watson–Crick base pair NH···N resonances of stem P1 (red lines) based on the two adjacent base pairs of A3–U20 and U4–A19. Likewise, the resonances of stem P2 (pairing of 5′-C(31)UG(33)-3′ with 5′-C(9)AG(11)-3′) were assigned unequivocally (blue lines). Furthermore, the ^{15}N-edited NOESY spectrum of the U9C/A33G mutant containing a labeled C17-$^{15}N(4)$ revealed that the signal at 12.9 ppm stemmed from the imino proton N1-H of $preQ_1$ pairing with C17 (see the Supporting Information). Since this signal also showed a pronounced cross peak to the N1-H of G11, intensive

stacking of the $preQ_1$–C17 base pair onto stem P2 can be assumed. Further assignment of the resonances at 9.7, 11.3, and 13.6 ppm awaits determination in the context of a detailed NMR spectroscopy study on the interactions of $preQ_1$ and its aptamer with fully $^{13}C/^{15}N$-labeled RNAs.

From the NMR spectroscopy investigations, we also collected evidence that Mg^{2+} ions are required for complete binding of $preQ_1$. In complementary manner, this aspect was investigated by lead(II) acetate probing.[7] Figure 4 depicts the cleavage

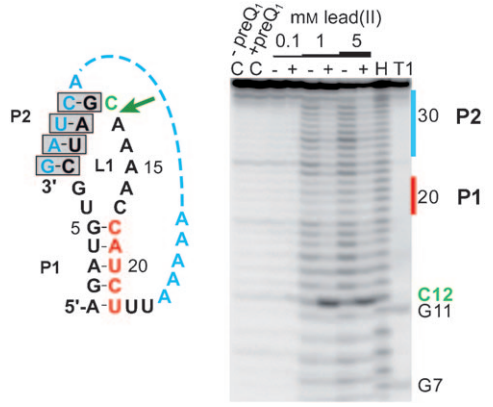

Figure 4. Chemical probing for mapping the pseudoknot structure and Mg^{2+} binding sites in vitro by lead(II) acetate in the absence (−) and presence (+) of a fivefold excess of $preQ_1$. C designates control reactions without and with $preQ_1$. T1 and H designate RNase T1 and alkaline hydrolysis ladders, respectively. Stem P1 and P2 are more resistant to Pb^{2+}-induced cleavage. A specific and prominent Pb^{2+} cleavage site is evident at C12A13 (green arrow). Conditions: [RNA] = 2.5 μM; [$preQ_1$] = 12.5 μM; buffer: 50 mM KMOPS, 100 mM KCl, 2 mM Mg^{2+}, pH 7.0, 293 K; [lead(II) acetate] = 0.1–5 mM.

pattern for the 34 nt wild-type $preQ_1$ aptamer. In the presence of $preQ_1$, significant cleavage between cytidine C12 and adenosine A13 was observed. Moreover, the tight interaction of stem P2 in the presence of $preQ_1$ is reflected by the probing pattern since this region shows reduced cleavage as observed for stem P1 (Figure 4). Furthermore, in a competition experiment with a constant concentration of Pb^{2+} ions, Pb^{2+}-induced cleavage was reduced in response to increasing Mg^{2+} ion concentrations (data not shown); this indicates a potential specific binding site for divalent metal ions close to C12A13.

In summary, our comprehensive study on mutated $preQ_1$ aptamers in combination with chemical and NMR spectroscopic structure probing[8] reveals that $preQ_1$ binds with concurrent pseudoknot formation of the aptamer. High resolution structures of several small pseudoknots have been obtained by X-ray crystallography or NMR spectroscopy. These pseudoknots include plant luteoviral RNA,[9] beet western yellow virus (BWYV) RNA,[10] sugarcane yellow leaf virus (ScYLV) RNA,[11] and mouse mammary tumor virus.[12] Their functions are usually associated with ribosomal frame-shifting. Other examples refer to RNA pseudoknots of human telomerase.[13–15] Furthermore, the structure of an in vitro selected aptamer that binds biotin was solved at high resolution and also revealed a pseudoknot structure.[16] Interestingly, biotin is bound at the interface between the two stacked helices of a 33 nt RNA forming a pseudoknot. This overall scenario seems reasonable for the $preQ_1$ bound aptamer as well, and awaits determination of its structural details and elucidation of how pseudoknot formation influences the expression platform in order to get insights into this particular gene response mechanism.

Experimental Section

All RNAs were chemically synthesized by solid-phase synthesis by using 2′-*O*-TOM protected phosphoramidites.[17] 2′-*O*-(Triisopropylsilyl)-3′-*O*-[(*N*,*N*-diisopropylamino-(β-cyanoethoxy)phosphino]-5′-*O*-(4,4′-dimethoxytrityl)-(N4-^{15}N)-acetylcytidine was chemically synthesized and used for site-specific incorporation into ^{15}N-labeled RNAs. NMR spectra were acquired by using a Varian Unity 500 MHz spectrometer. ^{1}H NMR imino proton spectra applied a selective excitation–refocusing sequence employing selective pulses shaped according to the G4 (excitation; 2.62 ms, RF amplitude 1.74 kHz) and REBURP (refocusing; 1.4 ms, RF amplitude 4.47 kHz) profile.[18] Both shaped pulses were centered at 13 ppm. HSQC experiments were recorded with 1024 complex data points. Remaining water magnetization was suppressed by a 3–9–19 WATERGATE pulse train. The proton sweep width was set to 12 kHz, and the proton carrier was centered at H_2O. The acquisition parameters for the 2D NOESY spectrum of mutant U9C/A33G were 3075 complex points in t2, 512 t1 increments and 128 transients per t1 increment in 48 h. Mixing time was 250 ms, and the relaxation delay was 2 s. The ^{15}N F2-filtered 2D NOESY was carried out with 764 complex points in t2, 128 t1 increments and 384 transients per t1 increment in 66 h. Mixing time was 300 ms (Supporting Information).

Acknowledgements

Funding from the Austrian Science Foundation FWF (P17864 to R.M., Y3150 to N.P.) and the Ministry of Science and Research (GenAU project consortium "non-coding RNAs" P0726-012-012 to R.M. and D1042-011-011 to N.P.) is acknowledged.

Keywords: pseudoknots · riboswitches · RNA · site-specific labeling · solid-phase synthesis

[1] a) R. K. Montange, R. T. Batey, *Annu. Rev. Biophys.* **2008**, *37*, 117–133; b) S. Blouin, J. Mulhbacher, J. Carlos Penedo, D. A. Lafontaine, *ChemBioChem* **2009**, *10*, 400–416; c) M. M. Meyer, A. Roth, S. M. Chervin, G. A. Garcia, R. R. Breaker, *RNA* **2008**, *14*, 685–695; d) S. D. Gilbert, R. P. Rambo, D. Van Tyne, R. T. Batey, *Nat. Struct. Mol. Biol.* **2008**, *15*, 177–182; e) H. Schwalbe, J. Buck, B. Fürtig, J. Noeske, J. Wöhnert, *Angew. Chem.* **2007**, *119*, 1232–1240; *Angew. Chem. Int. Ed.* **2007**, *46*, 1212–1219.

[2] A. Roth, W. C. Winkler, E. E. Regulski, B. W. Lee, J. Lim, I. Jona, J. E. Barrick, A. Ritwik, J. N. Kim, R. Welz, D. Iwata-Reuyl, R. R. Breaker, *Nat. Struct. Mol. Biol.* **2007**, *14*, 308–317.

[3] D. W. Staple, S. E. Butcher, *PLoS Biol.* **2005**, *3*, e213.

[4] a) F. Klepper, K. Polborn, T. Carell, *Helv. Chim. Acta* **2005**, *88*, 2610–2616; b) T. Brückl, F. Klepper, K. Gutsmiedl, T. Carell, *Org. Biomol. Chem.* **2007**, *5*, 3821–3825.

[5] A. Serganov, Y. R. Yuan, O. Pikovskaya, A. Polonskaia, L. Malinina, A. T. Phan, C. Höbartner, R. Micura, R. R. Breaker, D. J. Patel, *Chem. Biol.* **2004**, *11*, 1729–1741.

[6] a) F. Jiang, D. J. Patel, X. Zhang, H. Zhao, R. A. Jones, *J. Biomol. NMR* **1997**, *9*, 55–62; b) J. Milecki, A. Földesi, A. Fischer, R. W. Adamiak, J. Chattopadhyaya, *J. Labelled Compd. Radiopharm.* **2001**, *44*, 763–783; c) S. Rüdisser, J. G. Pelton, I. Tinoco, *J. Biomol. NMR* **1999**, *15*, 173–176;

d) B. McConnell, P. C. Seawell, *Biochemistry* **1973**, *12*, 4426–4434; e) B. McConnell, P. C. Seawell, *Biochemistry* **1972**, *11*, 4382–4392; f) J. L. Markley, A. Bax, Y. Arata, C. W. Hilbers, R. Kaptein, B. D. Sykes, P. E. Wright, K. Wüthrich, *J. Mol. Biol.* **1998**, *280*, 933–952.

[7] a) N. Polacek, A. Barta, *RNA* **1998**, *4*, 1282–1294; b) R. S. Brown, J. C. Dewan, A. Klug, *Biochemistry* **1985**, *24*, 4785–4801; c) R. S. Brown, B. E. Hingerty, J. C. Dewan, A. Klug, *Nature* **1983**, *303*, 543–546; d) E. Huntzinger, F. Winter, C. Ehresmann, P. Romby in *Handbook of RNA Biochemistry, Vol. 1* (Eds.: R. K. Hartmann, A. Bindereif, A. Schön, E. Westhof), Wiley-VCH, Weinheim, **2005**, pp. 151–171.

[8] J. M. Hart, S. D. Kennedy, D. H. Mathews, D. H. Turner, *J. Am. Chem. Soc.* **2008**, *130*, 10233–10239.

[9] P. S. Pallan, W. S. Marshall, J. Harp, F. C. Jewett, Z. Wawrzak, B. A. Brown, A. Rich, M. Egli, *Biochemistry* **2005**, *44*, 11315–11322.

[10] M. Egli, G. Minasov, L. Su, A. Rich, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4302–4307.

[11] P. V. Cornish, M. Hennig, D. P. Giedroc, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12694–12699.

[12] H. Kang, I. Tinoco, Jr., *Nucleic Acids Res.* **1997**, *25*, 1943–1949.

[13] C. A. Theimer, C. A. Blois, J. Feigon, *Mol. Cell* **2005**, *17*, 671–682.

[14] N.-K. Kim, Q. Zhang, J. Zhou, C. A. Theimer, R. D. Peterson, J. Feigon, *J. Mol. Biol.* **2008**, *384*, 1249–1261.

[15] S. Nonin-Lecomte, B. Felden, F. Dardel, *Nucleic Acids Res.* **2006**, *34*, 1847–1853.

[16] a) J. Nix, D. Sussman, C. Wilson, *J. Mol. Biol.* **2000**, *296*, 1235–1244; b) I. Harvey, P. Garneau, J. Pelletier, *RNA* **2002**, *8*, 452–463.

[17] R. Micura, *Angew. Chem.* **2002**, *114*, 2369–2373; *Angew. Chem. Int. Ed.* **2002**, *41*, 2265–2269.

[18] C. Höbartner, R. Micura, *J. Mol. Biol.* **2003**, *325*, 421–431.

Received: February 10, 2009
Published online on April 20, 2009

