JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Pseudoknot Preorganization of the preQ$_1$ Class I Riboswitch

Tobias Santner,[‡] Ulrike Rieder,[‡,†] Christoph Kreutz,* and Ronald Micura*

Institute of Organic Chemistry and Center for Molecular Biosciences, University of Innsbruck, Austria

Supporting Information

ABSTRACT: To explore folding and ligand recognition of metabolite-responsive RNAs is of major importance to comprehend gene regulation by riboswitches. Here, we demonstrate, using NMR spectroscopy, that the free aptamer of a preQ$_1$ class I riboswitch preorganizes into a pseudoknot fold in a temperature- and Mg^{2+}-dependent manner. The preformed pseudoknot represents a structure that is close to the ligand-bound state and that likely represents the conformation selected by the ligand. Importantly, a defined base pair mutation within the pseudoknot interaction stipulates whether, in the absence of ligand, dimer formation of the aptamer competes with intramolecular pseudoknot formation. This study pinpoints how RNA preorganization is a crucial determinant for the adaptive recognition process of RNA and ligand.

As much as 4 % of bacterial genes are controlled by riboswitches that are metabolite-responsive mRNA domains located in 5′-untranslated leader sequences.[1] A typical riboswitch consists of a high-affinity aptamer that, upon binding of a specific metabolite, induces a conformational change in the adjoining expression platform, thereby signaling ON or OFF for gene expression.[2−6] Tremendous progress has been made in the structure determination of ligand-bound aptamers,[7,8] while the characterization of unbound riboswitches and the ligand-induced folding pathway that is directly associated with the conformational change in the expression platform, and hence the regulatory outcome, remains elusive.[9,10]

Here we shed light on the conformational diversity of the free form of the preQ$_1$ class I riboswitch in solution. This riboswitch is the smallest riboswitch known to date with an aptamer that comprises only 34 nucleotides and that specifically recognizes 7-aminomethyl-7-deazaguanine (preQ$_1$).[11] The minimal sequence and structure consensus refer to a hairpin comprising a 5 base-pair stem (P1) and a loop of 11−13 nucleotides (L1) together with a 3′ single-stranded nucleoside overhang, as originally proposed by Breaker and et al.[11] Several structural studies including our own revealed that preQ$_1$ binds with concurrent pseudoknot formation of the aptamer.[12−20] Very recently, Wedekind and et al. succeeded in crystallizing a ligand-free aptamer of the preQ$_1$ riboswitch from *Thermoanaerobacter tengcongensis*.[10,21] Comparison of the preQ$_1$-bound and ligand-free states showed an all-atom rmsd of 1.7 Å, with the most significant rearrangement being for an adenosine in the ligand-binding pocket (A14). This nucleobase flanks the pocket in the bound state, where it pairs with a guanosine (G11). In the free state, A14 moves 7.5 Å away from G11 into the location that was occupied by preQ$_1$ in the bound state. Because crystal structures are subject to packing interactions, the compactness and degree of folding of the *T. tengcongensis* aptamer were assessed in solution using small-angle X-ray scattering (SAXS).[21] Strikingly, from the SAXS data it appeared that the preQ$_1$ aptamer of the thermophilic organism was equally compact in the ligand-bound and -free states. However, when control SAXS measurements were conducted on the mesophilic counterpart (from *Fusobacterium nucleatum*), the two states were drastically different (preQ$_1$-bound state, R_G = 19.3 ± 0.03 Å and r_{Max} 64 ± 6 Å; ligand-free state 31.0 ± 0.07 and 107 ± 11 Å).[21] This surprising difference prompted us to explore the molecular basis of the different solution behavior of the *F. nucleatum* riboswitch.

To disclose the conformational performance of the *F. nucleatum* preQ$_1$ aptamer in solution, we focused on NMR spectroscopic methods. Based on a mutagenesis study in our laboratory, we started with a representative preQ$_1$ RNA, encompassing 34 nucleotides that provided an excellent dispersion of the imino proton resonances in the ^{1}H NMR sprectrum when bound to its dedicated ligand (Figure 1A). The ^{1}H,^{1}H-nuclear Overhauser enhancement spectroscopy (NOESY) spectrum together with the ^{15}N-filtered NOESY of a site-specifically labeled variant (C17−15N(4)) provided unambiguous NH signal assignments of the two stem regions P1 and P2 of the pseudoknot (P) fold (Figure 1B).[15] For the present work, we added a 5-fluorine label to the uridine in position 32 (U32), that was expected to be a sensitive probe for Watson−Crick (WC) base pairing, in order to persue pseudoknot formation (Figure 1C). At 298 K, and in the absence of ligand, the four signals observed in the ^{1}H NMR imino proton region were consistent with a highly developed stem P1 of the free RNA (Figure 1D, left). However, several low-intensity signals in the same chemical shift area indicated conformational heterogeneity. When the temperature was lowered to 283 K, a rather well-dispersed imino proton spectrum was obtained for the free RNA reflecting characteristic patterns that resembled, to some extent, the imino proton spectrum of the same RNA bound to the ligand (Figure 1B,D). We interpreted this vague resemblance as a hint that pseudoknot formation can markedly occur at lower temperatures for the free RNA. This interpretation was in accordance with the ^{19}F NMR spectra of the same RNA sample. At 298 K, we observed a major signal at −164.5 ppm, typical for an unpaired 5-F uridine[22] and in accordance with the single-stranded 3′-RNA overhang as proposed in the secondary structure model (Figure 1A,D). A very minor signal group (of

Received: May 23, 2012
Published: July 9, 2012

dx.doi.org/10.1021/ja3049964 | J. Am. Chem. Soc. 2012, 134, 11928−11931

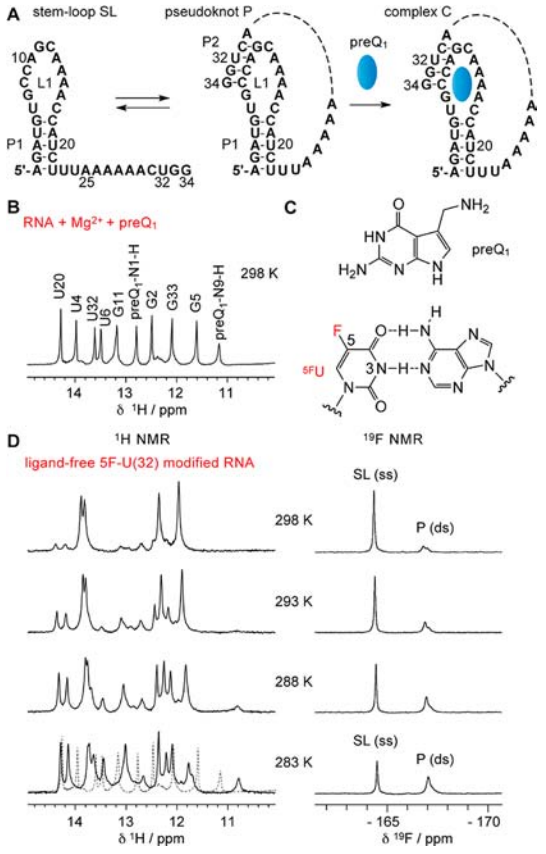

Figure 1. Initial investigations on a typical preQ$_1$ riboswitch. (A) Proposed structure equilibrium of SL and P folds. (B) ^{1}H NMR imino proton spectrum of the unmodified 34 nt preQ$_1$ RNA in complex with preQ$_1$. (C) Chemical structure of preQ$_1$ (top) and a 5F-U:A base pair (bottom). D) Temperature-dependent series of ^{1}H NMR (left) and ^{19}F NMR (right) spectra of ligand-free 5F-U labeled RNA; overlay with spectrum in B (dashed line). Conditions for D: c_{RNA} = 0.6 mM, 25 mM Na_2HAsO_4, pH 6.5, H_2O/D_2O 9/1, temperatures as indicated and for B: c_{RNA} = 0.9 mM, $c_{Mg^{2+}}$ = 4.0 mM, c_{preQ1} = 1.8 mM.

two overlapping resonances) was present at −167.2 ppm, a shift value typical for WC base-paired 5-F uridines.[22] Importantly, at 288 K, the ratio of the two signals approached 1:1, with a significantly broader (and merged) signal for the putative WC base pair in the pseudoknot region. The line shape perturbation of the pseudoknot fluorine resonance points toward an additional localized dynamic phenomenon in the micro/millisecond time regime in accordance with the anticipated transient nature of the P fold of the free RNA.

Encouraged by these observations, we envisaged direct evidence for the preorganization of the P fold and aimed at an RNA sample with noninvasive, selective ^{15}N labeling patterns. We considered to equip two potential A-U base pairs, one in stem P1 (A3:U20) and the other in the pseudoknot P2 (A10:U32) with ^{15}N(3)-U and ^{15}N(1)-A modified nucleosides to enable direct detection of WC base pairs by applying J_{NN} HNN COSY experiments.[23−25] We therefore developed a straightforward synthesis of the corresponding ^{15}N(3)- and ^{15}N(1)-modified uridine and adenosine phosphoramidites (based on reports in the literature)[26−28] and incorporated them into the 34 nt RNA target (Figure 2A).[29,30] Figure 2 depicts the corresponding ^{1}H−^{15}N heteronuclear single quantum coherence (HSQC) and ^{1}H/^{15}N/^{15}N(HNN) COSY spectra of free RNA (Figure 2B), RNA with Mg^{2+} (Figure 2D),

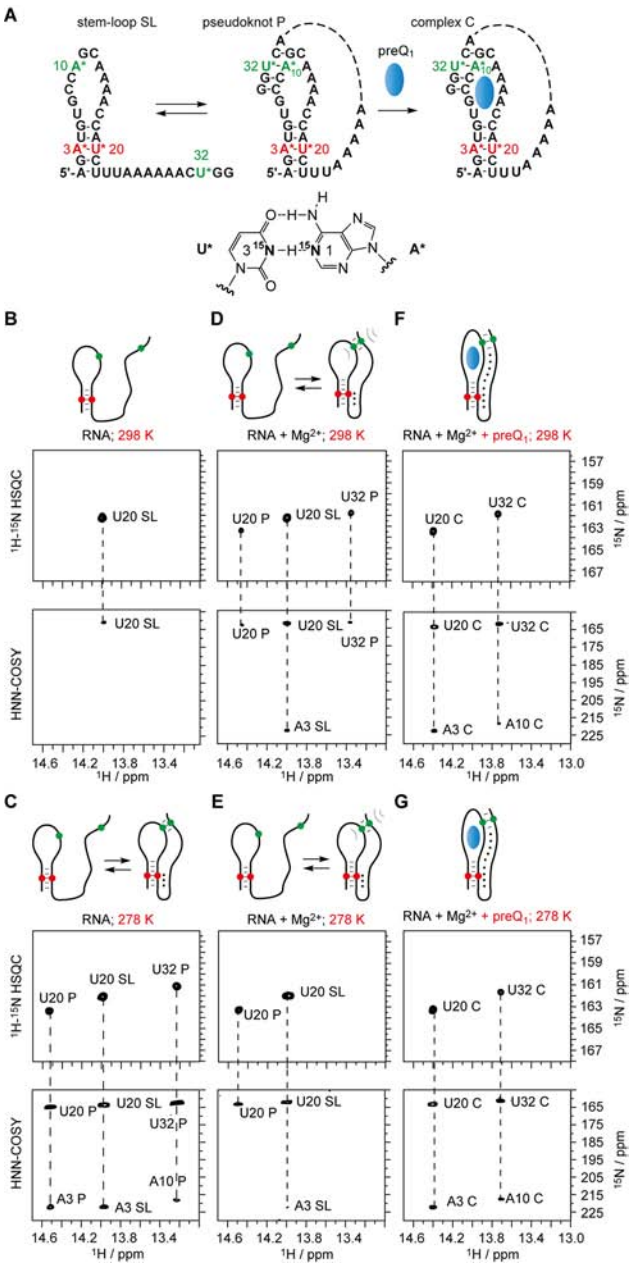

Figure 2. Direct evidence for base pair formation in the free and ligand-unbound preQ$_1$ aptamer. (A) Proposed structure model with labeling positions highlighted in red and green colors. Chemical structure of labeled nucleobases: ^{15}N(3)-U:^{15}N(1)-A base pairs. (B−G) ^{1}H−^{15}N HSQC and J_{NN} HNN COSY spectra and conformational cartoons at 298 K of the labeled RNA at 298 K (B,D,F) and 278 K (C,E,G). Conditions: c_{RNA} = 0.5 mM, 10 mM cacodylate buffer [$Na(CH_3)_2AsO_2{\cdot}3H_2O$], pH 6.5, temperatures as indicated; $c_{Mg^{2+}}$ = 2.0 mM, c_{preQ1} = 1 mM. Resonance assignment was supported by the A*10/U*32 single base pair labeled RNA (Figure 2, SI).

and RNA, Mg^{2+}, with preQ$_1$ ligand (Figure 2F) at 298 K. The same series of experiments was measured at 278 K (Figure 2C,E,G). At 298 K, the free RNA (no Mg^{2+}) provided a single signal in the HSQC spectrum that can be assigned to U20:A3 of the stem−loop fold (SL) (Figure 2B). When the temperature was lowered by 20 K, three pronounced signals in the HSQC spectrum occurred that can be assigned to U20:A3 of the SL fold and to U20:A3 and U32:A10 of the P fold, based on the corresponding HNN COSY experiment

dx.doi.org/10.1021/ja3049964 | J. Am. Chem. Soc. 2012, 134, 11928−11931

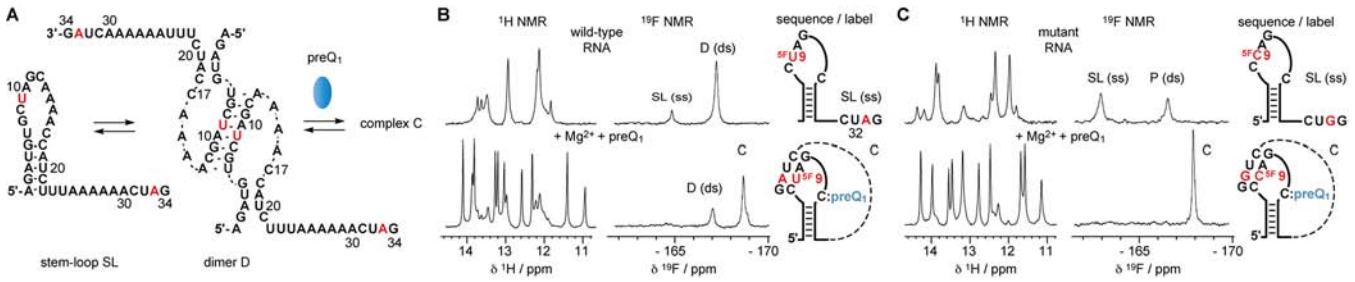

Figure 3. Comparison of wild-type (wt) and mutant *F. nucleatum* preQ$_1$ riboswitch aptamers. (A) Proposed equilibrium of SL and fold for the wt RNA. (B) ^{1}H- and ^{19}F-NMR spectra of 5F-U(9) modified, free (top), and preQ$_1$-bound (bottom) wt RNA. (C) ^{1}H- and ^{19}F-NMR spectra of 5F-C(9) modified free (top) RNA and preQ$_1$-bound (bottom) mutant RNA. For interpretation see main text. Conditions: c_{RNA} = 0.6 mM, 25 mM Na_2HAsO_4, pH 6.5, H_2O/D_2O 9/1, 298 K; $c_{Mg^{2+}}$ = 3.0 mM, c_{preQ1} = 1.2 mM. Double-stranded ds; single-stranded ss.

(Figure 2C). A potential HSQC signal for the unpaired U32 in the SL fold could not be detected because of proton exchange with the solvent. Clearly, the two folds (SL and P) were in slow conformational exchange with respect to the NMR time scale and evenly distributed. Upon addition of Mg^{2+}, the spectra of the free RNA, measured at 298 K (Figure 2D), resembled the preceding low-temperature/no Mg^{2+} spectra. According to signal intensities, the SL fold now appeared somewhat higher populated compared to the P fold. Noteworthily, in the presence of magnesium ions, the conformational exchange became expedited as reflected by signal broadening (see also Figure 1, Supporting Information, SI). Also, imino proton exchange with the solvent seemed enhanced, consistent with the lower intensities of the correlation signals in the HNN COSY experiment (Figure 2D).

When the RNA/Mg^{2+} sample was cooled down to 278 K, the spectra appeared simplified (Figure 2E) and, at first sight, implied a single species. However, the corresponding HSQC and HNN COSY resonances can be unequivocally assigned to U20:A3 of the SL fold and to U20:A3 of the P fold. Base pairing of U32:A10 in the P fold exhibited intermediate exchange dynamics that significantly broadened the corresponding resonances (below threshold set in Figure 2E). The expected HSQC resonance of U32 had a low S/N ratio, but it was unambiguously observed, as depicted for the analogous set of NMR experiments on the same RNA with a sole ^{15}N-labeled base pair of U32:A10 in Figure 2, SI. We furthermore note that the nearly identical chemical shift values of H−N(1) U20 of the P fold, observed for the three different conditions (no ligands, Mg^{2+}, Mg^{2+} and preQ$_1$) at two different temperatures, indicate a close conformational resemblance of the P1 containing triple helical segment of free and bound states of the aptamer.

The above set of experiments unequivocally corroborates the existence of a preorganized pseudoknot structure in solution for the preQ$_1$ riboswitch aptamer from *F. nucleatum*, that at low temperatures or in the presence of Mg^{2+} can become a significant population. It is therefore surprising, at first sight, that the SAXS data reported in the literature accentuate a difference in compactness and degree of folding by a factor of ∼2 (between free and bound forms), which is a discrepancy with respect to the model of a preorganized fold.[21] We therefore invested efforts to understand this issue. As mentioned previously, we relied on a preQ$_1$ riboswitch sequence that was obtained from a broad mutagenesis study.[15] The selected sequence contained a single base pair replacement within the pseudoknot stem P2, represented by the double mutation A33G/U9C of the wild-type (wt) sequence (Figure 3A; Figure 3, SI). From preliminary isothermal titration calorimetry (ITC) experiments, the mutant indicated a 2-fold increase in ligand affinity. Furthermore, we had indications from previously performed gel-shift assays that, for the wt sequence, dimerization of the aptamer could occur at high RNA concentrations.[16] We therefore synthesized the wt RNA with a 5-F uridine label at loop site 9, because this position was assumed to be sensitive for dimer formation (Figure 3B; Figure 3E,F, SI) as opposed to position 32 which can only sense intramolecular pseudoknot formation (Figure 3G,H, SI). Two ^{19}F resonances from 5F-U(9) in a 1:9 ratio were indeed observed with the major signal at −167.2 ppm, assigned to the hypothetic dimer that can form because of a palindromic, fully WC paired interface (Figure 3A,B; Figure 3E, SI). In complementary line of evidence, the 5F-U(32) label provided a sole signal at −164.4 ppm characteristic for an unpaired 5F-U and therefore excludes the possibility of an intramolecular P fold (Figure 3G, SI). Strong support for these conformational assignments further came from a truncated 22 nt SL RNA without the single-stranded nucleoside overhang (Figure 3C,D, SI), therefore lacking the possibility for intramolecular pseudoknot formation, that showed the same signal patterns for imino proton and fluorine resonances as the full-size aptamer (Figure 3E, SI). For the 22 nt RNA, these NH signals can only originate from intermolecular WC base pairing. In addition, UV melting profiles of the shorter 22 nt RNA showed a clear dependence of the T_m values from concentration, typical for a bimolecular interaction (Figure 4, SI). We also mention that the dimer of the free RNA presumably contains an open stem P1, as suggested by analysis of the ^{1}H NMR NH signal ratios of monomers and dimers for the various labeled and unlabeled wt preQ$_1$ RNAs (Figure 3B–E,G, SI).

Importantly, when a ligand was added in one-fold excess to the 5F-U(9)-labeled wt aptamer, the complex was significantly formed, although a minor population of the dimer was still competing (Figure 3B, bottom). We additionally investigated the A33G/U9C mutant with a 5-F reporter at C9 and observed two ^{19}F signals in 6:4 ratio in favor of the SL fold (Figure 3C). The signal at −162.8 ppm is characteristic for a G:^{5F}C base pair[22] and was assigned to the monomolecular P fold, consistent with the set of HSQC and HNN NMR experiments (Figure 2). Transformation into a single conformer (complex C) was detected for the mutant aptamer when a ligand was added in one-fold excess (Figure 3C, bottom).

The different behavior of wt compared to mutant *F. nucleatum* aptamer with respect to dimerization can be rationalized by the disruption of the two central A:U base pairs (wt) in the palindromic 8 nt WC interface (Figure 3A). In

dx.doi.org/10.1021/ja3049964 | J. Am. Chem. Soc. 2012, 134, 11928−11931

a hypothetic dimer of the mutant, two A/C mismatches in the center of the palindrom make this possibility rather unlikely. Additionally, in the mutant, the pseudoknot double helix P2 is thermodynamically stabilized because of the higher G:C content and certainly supports pseudoknot preorganization. We believe that the dimerization of the wt riboswitch has most likely no biological function but rather represents an artifact that has to be handled when high concentrations are involved, as required for many biophysical investigations, such as for NMR spectroscopy, X-ray crystallography, or SAXS analysis. It is therefore reasonable to explain the different SAXS parameters reported for free and bound forms of the *F. nucleatum* riboswitch[21] by dimer formation of the free RNA, while monomers exist in the bound form. In this context, we mention that 2AP-fold[9] experiments were performed at low concentrations of wt RNA and were in accordance with stabilization of the pseudoknot through Mg^{2+} (Figure 5, SI).

In summary, we provide clear evidence that the smallest known riboswitch becomes preorganized from a SL into a P fold that is structurally close to the ligand-bound complex. These open and closed conformations interchange slowly and exist in dynamic equilibrium. The preformed pseudoknot likely represents the conformation that is selected by the ligand to make initial contacts with the nucleobases of the binding pocket. Although more sophisticated NMR experiments would be required for a direct proof that pseudoknot formation indeed facilitates recognition of $preQ_1$, these features are characteristics for the 'conformational capture/selection' mechanism that postulates a population shift of pre-existing conformational states upon ligand interaction.[31] The following optimization of the binding pocket by smaller conformational rearrangements of nucleobases, ribose units and phosphate backbone refers to 'induced fit' recognition.[32,33] Both concepts play a significant role during riboswitch–ligand recognition.[9]

Our study particularly contributes to understanding the folding behavior of pseudoknot-type riboswitches and how specific mutations impact on their conformational heterogeneity. This knowledge will be of use to manipulate wt riboswitch sequences to engineer well-behaved two-state systems for applications in biotechnology.[34−37]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Experimental details and characterization data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

ronald.micura@uibk.ac.at; christoph.kreutz@uibk.ac.at

Present Address

[†]Institute of Organic Chemistry, University of Zurich.

Author Contributions

[‡]These authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Funding by the Austrian Science Fund FWF (P21641, I317, I844) and the EU (ITN RNPnet) is acknowledged. We thank M. Aigner for the synthesis of 5-F cytidine, C. Reymond for ITC measurements, and C. Höbartner and M. F. Soulière for critical reading of the manuscript.

■ REFERENCES

(1) Breaker, R. R. *Mol. Cell* **2011**, *43*, 867.
(2) Garst, A. D.; Edwards, A. L.; Batey, R. T. *Cold Spring Harbor Perspect. Biol.* **2011**, *3*, doi:10.1101/cshperspect.a003533.
(3) Deigan, K. E.; Ferré-D'Amaré, A. R. *Acc. Chem. Res.* **2011**, *44*, 1329.
(4) Serganov, A.; Patel, D. J. *Curr. Opin. Struct. Biol.* **2012**, *22*, 1.
(5) Blouin, S.; Mulhbacher, J.; Penedo, J. C.; Lafontaine, D. A. *ChemBioChem* **2009**, *10*, 400.
(6) Nudler, E.; Mironov, A. S. *Trends Biochem. Sci.* **2004**, *29*, 11.
(7) Schwalbe, H.; Buck, J.; Furtig, B.; Noeske, J.; Wöhnert, J. *Angew. Chem., Int. Ed.* **2007**, *46*, 1212.
(8) Serganov, A. *RNA Biol.* **2010**, *7*, 98.
(9) Haller, A.; Soulière, M. F.; Micura, R. *Acc. Chem. Res.* **2011**, *44*, 1339.
(10) Liberman, J. A.; Wedekind, J. E. *WIREs RNA* **2011**, *3*, 369.
(11) Roth, A.; Winkler, W. C.; Regulski, E. E.; Lee, B. W. K.; Lim, J.; Jona, I.; Barrick, J. E.; Ritwik, A; Kim, J. N.; Welz, R.; Iwata-Reuyl, D.; Breaker, R. R. *Nat. Struct. Mol. Biol.* **2007**, *14*, 308.
(12) Klein, D. J.; Edwards, T. E.; Ferré-D'Amaré, A. R. *Nat. Struct. Mol. Biol.* **2009**, *16*, 343.
(13) Kang, M.; Peterson, R.; Feigon, J. *Mol. Cell* **2009**, *33*, 784.
(14) Spitale, R. C.; Torelli, A. T.; Krucinska, J.; Bandarian, V.; Wedekind, J. E. *J. Biol. Chem.* **2009**, *284*, 11012.
(15) Rieder, U.; Lang, K.; Kreutz, C.; Polacek, N.; Micura, R. *Chembiochem* **2009**, *10*, 1141.
(16) Rieder, U.; Kreutz, C.; Micura, R. *Proc. Nat. Acad. Sci. U.S.A.* **2010**, *107*, 10804.
(17) Feng, J.; Walter, N. G.; Brooks, C. L., III *J. Am. Chem. Soc.* **2011**, *133*, 4196.
(18) Zhang, Q.; Kang, M.; Peterson, R. D.; Feigon, J. *J. Am. Chem. Soc.* **2011**, *133*, 5190.
(19) Petrone, P. M.; Dewhurst, J.; Tommasi, R.; Whitehead, L.; Pomerantz, A. K. *J. Mol. Graphics Modell.* **2011**, *30*, 179.
(20) Eichhorn, C. D.; Feng, J.; Suddala, K. C.; Walter, N. G.; Brooks, C. L.; Al-Hashimi, H. M. *Nucleic Acids Res.* **2012**, *40*, 1345.
(21) Jenkins, J. L.; Krucinska, J.; McCarty, R. M.; Bandarian, V.; Wedekind, J. E. *J. Biol. Chem.* **2011**, *286*, 24626.
(22) Puffer, B.; Kreutz, C.; Rieder, U.; Ebert, M. O.; Konrat, R.; Micura, R. *Nucleic Acids Res.* **2009**, *37*, 7728.
(23) Fürtig, B.; Richter, C.; Wöhnert, J.; Schwalbe, H. *ChemBioChem* **2003**, *4*, 936.
(24) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293.
(25) Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14147.
(26) Wenter, P.; Pitsch, S. *Helv. Chim. Acta* **2003**, *86*, 3955.
(27) Baral, B.; Kumar, P.; Anderson, B. A.; Østergaard, M. E.; Sharma, P. K.; Hrdlicka, P. J. *Tetrahedron Lett.* **2009**, *50*, 5850.
(28) Ariza, X.; Bou, V.; Vilarrasa, J. *J. Am. Chem. Soc.* **1995**, *117*, 3665.
(29) Micura, R. *Angew. Chem., Int. Ed.* **2002**, *41*, 2265.
(30) Wachowius, F.; Höbartner, C. *ChemBioChem* **2010**, *11*, 469.
(31) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789.
(32) Leulliot, N.; Varani, G. *Biochemistry* **2001**, *40*, 7947.
(33) Hermann, T.; Patel, D. J. *Science* **2000**, *287*, 820.
(34) Wieland, M.; Benz, A.; Klauser, B.; Hartig, J. S. *Angew. Chem., Int. Ed.* **2009**, *48*, 2715.
(35) Weigand, J. E.; Schmidtke, S. R.; Will, T. J.; Duchardt-Ferner, E.; Hammann, C.; Wöhnert, J.; Suess, B. *Nucleic Acids Res.* **2011**, *39*, 3363.
(36) Chang, A. L.; Wolf, J. J.; Smolke, C. D. *Curr. Opin. Biotechnol.* **2012**, *23*, doi 10.1016/j.copbio.2012.01.005.
(37) Quarta, G.; Sin, K.; Schlick, T. *PLoS Comp. Biol.* **2012**, *8*, e1002368.