RNA Structure Elucidation

DOI: 10.1002/anie.201101743

Rapid Identification of Noncanonical RNA Structure Elements by Direct Detection of OH···O=P, NH···O=P, and NH_2···O=P Hydrogen Bonds in Solution NMR Spectroscopy**

*Elke Duchardt-Ferner, Jan Ferner, and Jens Wöhnert**

Dedicated to Professor Heinz Rüterjans on the occasion of his 75th birthday

Hydrogen bonds are of fundamental importance for the formation of functional nucleic acid structures. In double-helical DNA hydrogen bonding occurs according to the classical Watson–Crick base-pairing scheme. In contrast, hydrogen bonds in RNA are much more diverse with respect to their donor and acceptor groups. Most importantly, non-canonical hydrogen bonds deviating from the Watson–Crick scheme are the basis for the intricate, globular proteinlike folds observed in functional RNA molecules such as tRNAs, self-splicing introns, ribozymes, riboswitches, and ribosomal RNAs.

These hydrogen bonds also involve donor and acceptor groups on the Hoogsteen and sugar edges of the nucleobases as well as the 2′-hydroxy groups of the ribose and the oxygen atoms of the phosphate backbone. Hydrogen bonds including the negatively charged oxygen atoms of the phosphate backbone as acceptor groups are particularly strong.[1] They are very important as an alternative to cation binding in neutralizing the negative charge of backbone phosphate groups and thereby reducing the electrostatic repulsion between closely approaching backbone segments in complex RNA structures. Accordingly, many of the recurring three-dimensional folding motifs of functional RNAs are stabilized by conserved "signature" hydrogen bonds involving backbone phosphate groups as hydrogen-bond acceptors. Classical examples are the "U-turn" motif initially identified in the anticodon and T-loops of tRNAs, where a hydrogen bond between the U imino proton and the phosphate of the (N+3) nucleotide stabilizes a sharp turn in the RNA backbone; the ultrastable GNRA tetraloop with a bifurcated hydrogen bond between the G amino and imino groups and the phosphate backbone; or the sarcin/ricin internal loop motif.[2]

The analysis of RNA structures has been revolutionized by the realization that hydrogen bonds in RNA and DNA can be directly detected using NMR spectroscopy owing to the presence of through-hydrogen-bond scalar couplings between nuclei of the hydrogen-bond donor and acceptor groups as a result of their partially covalent nature.[3] NMR spectroscopy experiments linking individual donor and acceptor groups were initially developed for NH···N-type hydrogen bonds in Watson–Crick and noncanonical base pairs and later extended to the characterization of NH···O=C and OH···N-type hydrogen bonds in RNA.[3,4] In many cases the information obtained from such experiments alone is sufficient to derive the secondary structure and gain insight into the tertiary folding of RNAs.[5]

For proteins binding to cofactors containing phosphate groups, such as guanosine triphosphate and flavin mononucleotide, it has been shown that hydrogen bonds between backbone amide or amino acid side chain hydroxy groups and the phosphate group can also be detected by NMR spectroscopy.[6] Similar observations have been reported for small organic phosphate receptor molecules.[7] Surprisingly, for RNA no NMR spectroscopy experiments have been reported to date for the direct identification of hydrogen bonds involving phosphate groups as the hydrogen-bond acceptor, despite their importance in the stabilization of RNA non-canonical structure elements.

Herein we demonstrate that OH···O=P, NH···O=P, and NH_2···O=P hydrogen bonds typical for different noncanonical structure elements in RNA can be readily detected by NMR spectroscopy in solution, as they give rise to sizeable scalar $^{2h}J_{HP}$ couplings between the donor-group protons and the ^{31}P nuclei of the phosphate backbone. We used two model systems for our experiments—a 27 nt synthetic neomycin sensing riboswitch in complex with ribostamycin, and a 14 nt RNA containing a stable UUCG-tetraloop (nt = nucleotide). High-resolution 3D structures have been reported for both systems.[8]

The NOE-based structure of the riboswitch RNA revealed two intertwined structural elements known to

[*] Dr. E. Duchardt-Ferner, Prof. Dr. J. Wöhnert
Institute of Molecular Biosciences
Center for Biomolecular Magnetic Resonance (BMRZ)
Johann-Wolfgang-Goethe-Universität Frankfurt
Max-von-Laue-Strasse 9, 60438 Frankfurt (Germany)
E-mail: woehnert@bio.uni-frankfurt.de
Dr. J. Ferner
Institute of Organic Chemistry and Chemical Biology
Center for Biomolecular Magnetic Resonance (BMRZ)
Johann-Wolfgang-Goethe-Universität Frankfurt
Max-von-Laue-Strasse 7, 60438 Frankfurt (Germany)

[**] This work was supported by an Aventis Foundation endowed professorship in Chemical Biology to J.W., the Deutsche Forschungsgemeinschaft (DFG) (grant WO 901/1-2 to J.W.), the Center for Biomolecular Magnetic Resonance (BMRZ) and the Center of Excellence "Macromolecular complexes" of the Johann-Wolfgang-Goethe-Universität Frankfurt. We are grateful to J. Weigand, B. Suess, S. Häfner, M. Görlach, and H. Schwalbe for providing the RNA samples and to H. Schwalbe and J. P. Wurm for valuable discussions.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201101743.

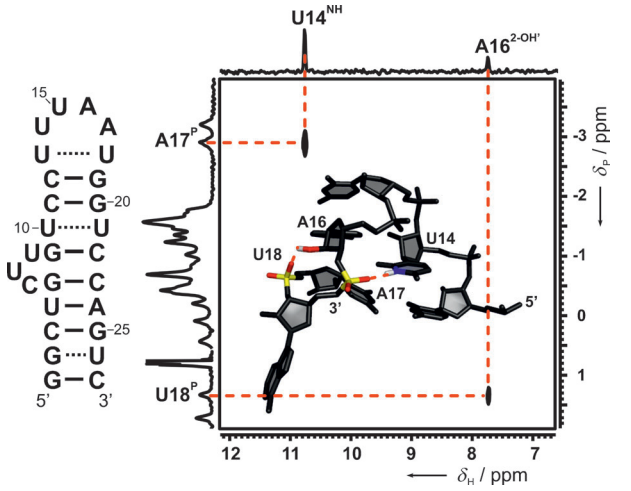

Figure 1. Left: Secondary structure and numbering scheme of the 27 nt neomycin riboswitch when bound to ribostamycin. Right: Long-range WATERGATE ^{1}H,^{31}P HSQC spectrum recorded on the RNA–ribostamycin complex at 5 °C using ^{2}H-labeled RNA. At the left side, the full 1D ^{31}P spectrum of the RNA is shown, and the relevant ^{31}P assignments are indicated.[10] The two observed correlations correspond to the presence of hydrogen bonds between the A16-2′-OH and the U18 phosphate group and the U14-NH and the of A17 phosphate group. The section of the NOE-based 3D structure, including the bulged-out base motif (A16-A17-U18) and the U-turn motif (U14 to A17), is shown as an inset. Functional groups predicted to be involved in hydrogen bonds to backbone phosphate groups are colored according to atom type (H white, N blue, O red, P yellow), and the proposed hydrogen bonds are indicated by dashed red lines.

contain hydrogen bonds with backbone phosphate groups as acceptors: a bulged-out base motif (Figure 1, A16-A17-U18) often stabilized by a hydrogen bond between the 2′-OH group of nucleotide N and the 5′-phosphate group of nucleotide N+2 and a "U-turn" motif (Figure 1, U14-U15-A16-A17) with a hydrogen bond between the U imino group and the phosphate group of nucleotide N+3.[2a,9]

For both motifs, the NOE-based structure already suggested the presence of the respective hydrogen bond on the basis of distance and bond-angle criteria. Furthermore, both the A16-2′-OH and the U14-NH proton can be detected in standard NMR spectra, even at room temperature (Figures S1 and S4 in the Supporting Information). Their protection from fast exchange with the bulk solvent as well as their downfield chemical shifts (A16-2′-OH $\delta = 7.7$ ppm, U14-NH $\delta =$ 10.8 ppm) also agree with their involvement in stable hydrogen bonds.

We unambiguously verified the presence of hydrogen bonds between these two protons and phosphate backbone groups by detecting long-range hydrogen-bond-mediated scalar couplings between the protons of the hydrogen-bond donor groups and the ^{31}P nuclei of the acceptor groups in a simple 2D WATERGATE ^{1}H,^{31}P HSQC experiment with an extended INEPT delay (35–70 ms) for magnetization transfer between ^{1}H and ^{31}P recorded on a deuterated but otherwise unlabeled sample (see the Experimental Section) at 5 or 20 °C. The resulting spectrum clearly showed two correlation signals connecting the resonance of the A16-2′-OH proton with the U18 phosphate group and the U14-NH proton resonance with the A17 phosphate group (Figure 1; for ^{31}P assignments, see Figure S2 in the Supporting Information). Besides the scalar couplings across the hydrogen bonds, cross-correlated relaxation between the P,H dipolar interaction and the ^{31}P chemical shift anisotropy might also contribute to the magnetization transfer observed in this experiment. However, since this additional mechanism of magnetization transfer also requires close spatial proximity between the proton and the phosphate group tantamount to hydrogen-bond formation, it even contributes favorably to the signal-to-noise ratio in this experiment.

Alternatively, we recorded a non-refocused Carr–Purcell–Meiboom–Gill (CPMG) echo ^{1}H,^{31}P HSQC experiment where ^{1}H,^{31}P antiphase magnetization is recorded during t_2 (Figure S3 in the Supporting Information). This experiment is conceptually very similar to an experiment proposed by Hennig and co-workers for detecting 2′-OH···N-hydrogen bonds.[4g] It yielded the same correlations as the standard HSQC experiment (Figure S3 in the Supporting Information).

To quantify the $^{2h}J_{HP}$ couplings, we compared signal intensities in 1D ^{31}P{^{1}H} spin-echo experiments (Figure 2a) with and without proton decoupling using a variable spin-echo delay. For these experiments, a deuterated sample was advantageous, as unwanted magnetization transfer between ^{31}P nuclei and ribose protons across covalent bonds as well as ^{31}P relaxation by dipolar interactions with protons are minimized, thus allowing longer dephasing periods. Special care was taken to suppress scalar couplings to the residual ribose RNA protons by selective ^{1}H decoupling (Figure 2, Figure S5 in the Supporting Information) as well as contributions of cross-correlated relaxation between the P,H dipolar interaction and the ^{31}P chemical shift anisotropy to the measured scalar coupling. The $^{2h}J_{HP}$ scalar couplings across the A16-2′-

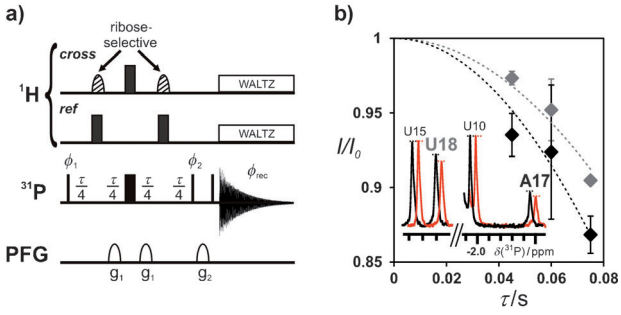

Figure 2. Quantification of scalar $^{2h}J_{HP}$ couplings for OH···O=P and NH···O=P hydrogen bonds in the neomycin riboswitch. a) Pulse sequences for cross- and reference spectra of 1D ^{31}P{^{1}H} spin-echo experiments. Narrow and wide black bars denote hard 90° and 180° pulses, respectively, open semiellipses pulse field gradient pulses. The shaded semiellipses in the cross experiment are selective ^{1}H 180° pulses targeted at the ribose protons. The phase cycling was $\phi_1 = x$, $-x$; $\phi_2 = x, x, -x, -x$; $\phi_{rec} = x, -x, -x, x$. Sine-shaped gradients of 1 ms duration with amplitudes of 11 G cm^{-1} (g_1) and 22 G cm^{-1} (g_2) were used. b) Signal intensity data for the A17 (black) and the U18 (gray) phosphorous resonance with variable spin-echo delays. The data were fit using the equation $I/I_0 = \cos(\pi J \tau)$ to obtain the $^{2h}J_{HP}$ values given in the text.

OH···U18 phosphate and the U14-NH···A17 phosphate hydrogen bonds are (1.8 ± 0.06) and (2.2 ± 0.07) Hz, respectively. Thus, their size is similar to the $^{2h}J_{HP}$ coupling constants found between protein functional groups and the phosphate groups of tightly bound cofactors.[6] The measured scalar couplings in this experiment might actually be smaller than the true coupling, since chemical exchange between the proton in the hydrogen bond and the water protons in principle leads to self-decoupling. However, an ^{1}H-detected spin-echo experiment, which is not affected by self-decoupling, yielded only a slightly larger value of (2.5 ± 0.09) Hz for the $^{2h}J_{HP}$ coupling constant between the U14-NH and the A17 phosphate group (Figure S6 in the Supporting Information).

Base NH_2 groups are the third type of hydrogen-bond donors occurring in RNA. An NH_2···O=P hydrogen bond is predicted to stabilize the structure of YNCG tetraloops based on NOE-derived NMR and X-ray structures and rationalizes the requirement for the C at position 3 of the consensus sequence for these loops.[8a–c,11] To test if hydrogen bonds with amino groups as the donor are also directly detectable by scalar couplings across the hydrogen bond, we recorded a long-range ^{1}H,^{31}P HSQC experiment on a ^{15}N-labeled 14 nt RNA containing a UUCG tetraloop (Figure 3). Besides numerous correlations between ribose protons and ^{31}P nuclei mediated by through-bond scalar couplings, the resulting spectrum shows a correlation between the ^{31}P nucleus of the U7 phosphate group and a proton with a chemical shift of $\delta = 6.8$ ppm. This proton is nitrogen-bound, since it appears as a doublet with a splitting of about 90 Hz in the absence of ^{15}N decoupling and has been assigned previously to the C8 amino group. Thus, hydrogen bonds with NH_2 groups as donor and backbone phosphate groups as acceptor also give rise to measurable scalar couplings across the hydrogen bond and can be identified directly by NMR spectroscopy.

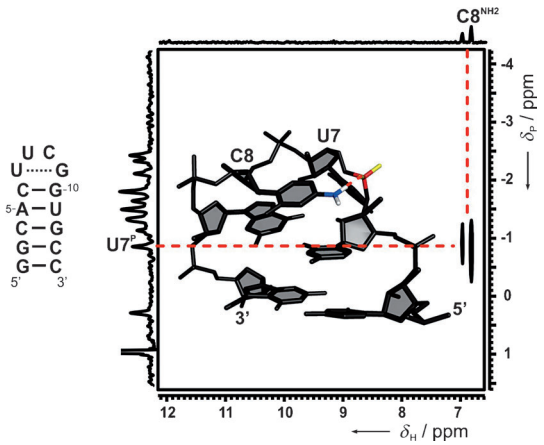

Figure 3. Left: Secondary structure and nucleotide numbering of the 14 nt RNA containing the stable UUCG tetraloop. Right: Long-range WATERGATE ^{1}H,^{31}P HSQC spectrum recorded on the ^{15}N-labeled RNA at 5 °C. At the left side, the full 1D ^{31}P spectrum of the RNA is shown. The ^{31}P assignments were reported previously.[12] The observed correlation directly reveals the presence of the hydrogen bond between the amino group of C8 and the U7 phosphate group highlighted in the 3D structure of the UUCG tetraloop shown in the inset.

Taken together, we show herein that all three types of hydrogen bonds with phosphate groups as the hydrogen-bond acceptor occurring in RNA result in measurable ^{2h}J scalar couplings across the hydrogen bond between the proton of the donor group (OH, NH, NH_2) and the ^{31}P nucleus of the acceptor. These hydrogen bonds are surprisingly easy to detect even in unlabeled RNA samples by standard ^{1}H,^{31}P correlation experiments requiring only moderate amounts of measurement time, even when only standard room-temperature NMR probes are available (Figure S7 in the Supporting Information). Hydrogen bonds including backbone phosphate groups do not occur in regular secondary structure elements such as A-form helices but only in long range noncanonical interactions. This situation should in many cases result in a good dispersion of the relevant ^{31}P signals. Therefore, donor and acceptor groups might often be identifiable in a straightforward manner. Furthermore, many recurring three-dimensional folding modules in RNA are stabilized by unique types of hydrogen bonds to phosphate groups. Thus, we anticipate that the direct identification of OH···O=P, NH···O=P, and NH_2···O=P hydrogen bonds by the experiments described above will greatly facilitate the rapid characterization of RNA tertiary structure by NMR spectroscopy even in instances where only limited assignments are available.

Experimental Section

The 27 nt neomycin riboswitch RNA was synthesized by in vitro transcription using ^{2}H-labeled nucleotides (Silantes GmbH, Munich) and purified as described.[8d] A 0.9 mM sample in NMR buffer containing 25 mM potassium phosphate (pH 6.2) and 50 mM KCl in 10% (v/v) D_2O was used for all NMR experiments. The 14 nt tetraloop RNA was purchased from Silantes GmbH (Munich) in ^{13}C,^{15}N-labeled form. The sample concentration was 0.7 mM dissolved in 20 mM potassium phosphate (pH 6.4) and 0.4 mM EDTA in 10% (v/v) D_2O.

All 2D H,P correlation spectra were recorded on a Bruker Avance 600 MHz spectrometer equipped with a 5 mm cryogenic HCP z-gradient probe. Long range ^{1}H,^{31}P HSQC spectra were recorded with an optimized INEPT transfer delay of 34 ms at 278 K. The temperature was optimized for signal intensity of both the HN and the 2′-OH resonance. A spectral width of 10 ppm in the indirect ^{31}P dimension was recorded with 24 complex points and 256 transients per increments in around 7 h. For the NH_2 group in the 14 nt RNA, the ^{1}H,^{31}P HSQC spectrum was recorded with a transfer delay of 30 ms, 20 complex t_1 increments, and 2560 transients per increment in about 60 h. For the determination of the $^{2h}J_{HP}$ coupling constants for the 27 nt RNA, quantitative spin-echo ^{31}P 1D spectra were obtained on a Bruker DRX300 spectrometer equipped with a 5 mm room-temperature BBO z-gradient probe. Coupling evolution delays of 45, 60, and 75 ms were recorded in duplicate measurements with 45056, 67584, and 81920 transients, respectively. Band-selective 180° Q3 Gaussian cascade pulses with a band width of 860 Hz and duration of 2.8 ms were used for selective decoupling of ribose 2′, 3′, 4′, and 5′ protons in the cross experiments and applied at 4.1 ppm.[13] Spectra processing and peak deconvolution was performed with Topspin2.1 software (Bruker Biospin). Curve fitting of deconvoluted signal intensities was carried out using Origin8 (OriginLab).

Received: March 10, 2011
Revised: May 31, 2011
Published online: July 21, 2011

Keywords: hydrogen bonds · NMR spectroscopy · phosphate groups · RNA · tertiary structures

[1] C. L. Zirbel, J. E. Sponer, J. Sponer, J. Stombaugh, N. B. Leontis, *Nucleic Acids Res.* **2009**, *37*, 4898 – 4918.

[2] a) G. J. Quigley, A. Rich, *Science* **1976**, *194*, 796 – 806; b) A. A. Szewczak, P. B. Moore, Y. L. Chan, I. G. Wool, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9581 – 9585; c) F. M. Jucker, H. A. Heus, P. F. Yip, E. H. Moors, A. Pardi, *J. Mol. Biol.* **1996**, *264*, 968 – 980.

[3] a) A. J. Dingley, S. Grzesiek, *J. Am. Chem. Soc.* **1998**, *120*, 8293 – 8297; b) K. Pervushin, A. Ono, C. Fernandez, T. Szyperski, M. Kainosho, K. Wuthrich, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14147 – 14151.

[4] a) A. J. Dingley, J. E. Masse, J. Feigon, S. Grzesiek, *J. Biomol. NMR* **2000**, *16*, 279 – 289; b) A. Majumdar, A. Kettani, E. Skripkin, D. J. Patel, *J. Biomol. NMR* **1999**, *15*, 207 – 211; c) J. Wöhnert, A. J. Dingley, M. Stoldt, M. Görlach, S. Grzesiek, L. R. Brown, *Nucleic Acids Res.* **1999**, *27*, 3104 – 3110; d) A. Majumdar, A. Kettani, E. Skripkin, *J. Biomol. NMR* **1999**, *14*, 67 – 70; e) A. Z. Liu, A. Majumdar, W. D. Hu, A. Kettani, E. Skripkin, D. J. Patel, *J. Am. Chem. Soc.* **2000**, *122*, 3206 – 3210; f) D. P. Giedroc, P. V. Cornish, M. Hennig, *J. Am. Chem. Soc.* **2003**, *125*, 4676 – 4677.

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

[6] a) F. Löhr, S. G. Mayhew, H. Ruterjans, *J. Am. Chem. Soc.* **2000**, *122*, 9289 – 9295; b) M. Mishima, M. Hatanaka, S. Yokoyama, T. Ikegami, M. Walchli, Y. Ito, M. Shirakawa, *J. Am. Chem. Soc.* **2000**, *122*, 5883 – 5884.

[7] R. M. Gschwind, M. Armbruster, I. Z. Zubrzycki, *J. Am. Chem. Soc.* **2004**, *126*, 10228 – 10229.

[8] a) F. H. T. Allain, G. Varani, *J. Mol. Biol.* **1995**, *250*, 333 – 353; b) E. Ennifar, A. Nikulin, S. Tishchenko, A. Serganov, N. Nevskaya, M. Garber, B. Ehresmann, C. Ehresmann, S. Nikonov, P. Dumas, *J. Mol. Biol.* **2000**, *304*, 35 – 42; c) S. Nozinovic, B. Fürtig, H. R. A. Jonker, C. Richter, H. Schwalbe, *Nucleic Acids Res.* **2010**, *38*, 683 – 694; d) E. Duchardt-Ferner, J. E. Weigand, O. Ohlenschlager, S. R. Schmidtke, B. Suess, J. Wöhnert, *Angew. Chem.* **2010**, *122*, 6352 – 6355; *Angew. Chem. Int. Ed.* **2010**, *49*, 6216 – 6219.

[9] N. B. Ulyanov, T. L. James, *New J. Chem.* **2010**, *34*, 910 – 917.

[10] S. R. Schmidtke, E. Duchardt-Ferner, J. E. Weigand, B. Suess, J. Wöhnert, *Biomol. NMR Assignments* **2010**, *4*, 115 – 118.

[11] O. Ohlenschläger, J. Wöhnert, E. Bucci, S. Seitz, S. Häfner, R. Ramachandran, R. Zell, M. Görlach, *Structure* **2004**, *12*, 237 – 248.

[12] B. Fürtig, C. Richter, W. Bermel, H. Schwalbe, *J. Biomol. NMR* **2004**, *28*, 69 – 79.

[13] L. Emsley, G. Bodenhausen, *J. Magn. Reson.* **1992**, *97*, 135 – 148.

