Journal of Biomolecular NMR **28:** 69–79, 2004.
KLUWER/ESCOM

New NMR experiments for RNA nucleobase resonance assignment and chemical shift analysis of an RNA UUCG tetraloop

Boris Fürtig[a], Christian Richter[a], Wolfgang Bermel[b] & Harald Schwalbe[a,*]
[a]*Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance, Johann Wolfgang Goethe-University, Marie-Curie-Strasse 11, D-60439 Frankfurt/M., Germany*
[b]*Bruker Biospin GmbH, Silberstreifen 4, D-76287 Rheinstetten, Germany*

Received 22 May 2003; Accepted 22 July 2003

Key words: chemical shift analysis, NMR resonance assignment, NMR spectroscopy, RNA, UUCG tetraloop

Introduction

Hairpins are among the most important secondary structure elements found in RNA. They are involved in a variety of RNA functions ranging from nucleation sites for RNA folding to mediating intermolecular interactions with other RNAs or proteins. Among all hairpin loops, tetraloops are the most abundantly found in RNA. Common stable tetraloop motifs are e.g. UNCG, GNRA, and CUUG loop sequences (N representing a pyrimidine nucleotide). Structure determination of RNA tetraloops has provided significant insight into the structural basis of their stability.

In particular, the structures of the UNCG tetraloops have been solved both by NMR (Cheong et al., 1990; Varani et al., 1991; Allain & Varani, 1995a, b) and by X-ray methods (Nissen et al., 2000; Wimberly et al., 2000; Ennifar et al., 2000). The 14mer RNA UUCG tetraloop discussed here is very stable and a wealth of structural data is available. In addition, the RNA can be purchased from commercial sources.

This paper provides the near to complete resonance assignment of a 14mer hairpin containing a cUUCGg-tetraloop including the resonances of nuclei not directly bonded to hydrogens. To this end, new experiments were developed for the assignment of quaternary carbons in the nucleobases. In addition, the information of chemical shifts regarding conformation is evaluated.

While the newly available chemical shifts of carbon atoms C2 in cytosines and C4 and C5 in purines of the nucleobases differ only marginally along the sequence and from the chemical shifts found in mononucleotides, ^{13}C and ^{31}P chemical shifts in the phosphodiester backbone are sensitive measures of RNA conformation.

In the first part of this paper we report the resonance assignment of ^{1}H, ^{15}N, ^{13}C, and ^{31}P atoms. In the second part of this report, the information content of the chemical shifts of C1′, C2′, C3′, C4′, and C5′ to derive the sugar pucker mode and the exocyclic torsion angle γ, the chemical shifts of ^{31}P to detect non-canonical backbone conformation and the combined chemical shifts of H5′/H5″ and C5′ to determine the spectroscopic assignment at the diastereotopic protons H5′/H5″ are analyzed.

Methods and experiments

The uniformly ^{13}C,^{15}N labeled RNA tetraloop sample with the sequence 5′-PO_3^--PO_2^--PO_2^--GGC-AC(UUCG)GUGCC-OH-3′ has been purchased from Silantes GmbH (Munich, Germany). Samples for NMR-spectroscopy contained ∼0.7 mM RNA in 20 mM $KHPO_4$, pH 6.4, 0.4 mM EDTA and 10% $^{2}H_2O$. In NMR spectra essentially no duplex of the RNA could be detected. ^{1}H chemical shifts are referenced directly to 3-[2,2,3,3-$^{2}H_4$]-trimethylsilylpropionate (TSP) as an external reference. ^{13}C and ^{15}N chemical shifts are referenced indirectly to external TSP (Wishart et al., 1995). The shifts of ^{31}P are referenced according to the recommendations of IUPAC (Harris et al., 2001).

Spectra were acquired at 25 °C, 10 °C and at 5 °C on Bruker DRX600, AV700, and AV800 spectrometers equipped with z-axis gradient ^{1}H{^{13}C,^{15}N} or

*To whom correspondence should be addressed. E-mail: schwalbe@nmr.uni-frankfurt.de

Table 1. ^{1}H-, ^{13}C-, and ^{31}P-chemical shift assignments of the ribose moiety and the phosphodiester backbone. Chemical shifts were referenced to TSP. Superscripts indicate the experiments used for resonance assignment a: NOESY, b: fwd. HCC-TOCSY-CCH-E.COSY, c: HCP, d: stereospecific assignment by use of selective C5′H5′-HSQC and fwd. HCC-TOCSY-CCH-E.COSY. Signals assigned in different experiments, are annotated with α: HCCH-COSY, β: ^{13}C edited NOESY and ^{13}C HSQC, γ: 2D-HH-NOESY

	C1′[a] (ppm)	C2′[b] (ppm)	C3′[b,c] (ppm)	C4′[b,c] (ppm)	C5′[b,c] (ppm)	H1′[a] (ppm)	H2′[b] (ppm)	H3′[b,c] (ppm)	H4′[b,c] (ppm)	H5′[b,c,d] (ppm)	H5″[b,c,d] (ppm)	p[c] (ppm)
G1	90.9	75.0$^{\alpha}$	74.2$^{\beta}$	83.7	67.7	5.86	4.98$^{\alpha}$	4.46$^{\beta}$	4.60	4.43	4.33	$P_{\gamma}-14$ $P_{\beta}-25.6$ $P_{\alpha}-12.9$
G2	93.0	75.5	73.4$^{\beta}$	82.7	66.5	5.95	4.6	4.61$^{\beta}$	4.60	4.54	4.36	−3.55
C3	93.8	75.3	72.1$^{\beta}$	81.8	64.2	5.57	4.71	4.48$^{\beta}$	4.49	4.55	4.08	−4.18
A4	93.0	75.8	72.9	81.9	65.2	6.01	4.65	4.67	4.54	4.60	4.19	−3.88
C5	93.8	75.6	72.1	81.9	65.0	5.38	4.39	4.17	4.43	4.47	4.06	−4.05
U6	94.5	76.0	73.1	82.4	64.6	5.67	3.80 OH2′6.4$^{\gamma}$	4.56	4.39	4.62	4.17	−4.35
U7	89.3	74.7	77.7	87.0	67.8	6.14	4.04	4.69	4.51	4.27	4.07	−3.41
C8	89.1	77.7	80.3	84.5	67.3	5.99	4.69	4.51	3.83	2.73	3.65	−5.03
G9	94.5	77.4	75.9	83.2	69.0	6.00	4.85	5.67	4.45	4.23	4.44	−4.92
G10	93.3	75.1	74.6	83.1	69.8	4.47	4.47	4.29	4.42	4.53	4.23	−2.4
U11	93.8	75.4	72.2	81.9	64.0	5.59	4.60	4.60	4.43	4.51	4.11	−4.35
G12	92.9	75.5	72.9	81.9	65.5	5.87	4.6	4.62	4.60	4.56	4.17	−3.95
C13	94.1	75.7	72.1	81.9$^{\beta}$	64.5	5.51	4.28	4.45	4.17$^{\beta}$	4.59	4.09	−4.4
C14	93.0	77.8$^{\alpha}$	69.7	83.4	65.1	5.79	4.06$^{\alpha}$	4.21	4.19	4.51	4.06	−4.1

Table 2. ^{1}H-, ^{13}C-chemical shift assignments of the nucleobases. Entries given in italics were measured at 10 °C, all other chemical shift values were measured at 25 °C. Chemical shifts were referenced to TSP. Superscripts indicate the experiments which were used for chemical shift assignment: a: HNCO, b: HCNC, c: HNCC, d: ^{13}C-HSQC, e: HCN, f: NOESY, g: ^{13}C edited NOESY, h: CPMG-NOESY and i: TROSY relayed HCCH-COSY

	C2 (ppm)	C4 (ppm)	C5 (ppm)	C6 (ppm)	C8 (ppm)	H1 (ppm)	H2 (ppm)	H3 (ppm)	H5 (ppm)	H6 (ppm)	H8 (ppm)	NH2 (ppm)
G1	*156.3*[a]	*153.1*[b]	*118.8*[c]	*157.1*[a]	139.0[de]	*13.01*[f]					8.17[ge]	
G2	*157*[a]	*152*[b]	*118.9*[c]	*161.7*[a]	137.0[de]	*13.41*[f]					7.70[ge]	
C3	158.7[b]	*168.8*[a]	97.6[d]	140.8[de]					5.34[g]	7.75[ge]		*8.67/7.02*[h]
A4	153.1[d]	147.3[i]	109.8[i]	159.6[i]	139.4[d]		7.46[gd]				8.09[ge]	*8.33/6.46*[h]
C5	158.3[b]	*168.5*[a]	97.4[d]	140.3[de]					5.23	7.31[ge]		*8.59/7.16*[h]
U6	*153.1*[a]	*168.0*[a]	105.1[d]	140.6[de]				*11.83*[f]	5.69[g]	7.77[ge]		
U7	*155.1*	*169.1*[a]	105.5[d]	144.7[de]				*11.26*[f]	5.89[g]	8.05[ge]		
C8	159.9[b]	*168.0*[a]	98.6[d]	142.8[de]					6.14[g]	7.72[ge]		*7.18/6.4*[h]
G9	*155.6*[a]	153.4[b]	*118.7*[c]	*161.6*[a]	142.9[de]	*9.96*[f]					7.90[ge]	
G10	*162.5*[a]	152.3[b]	*119.3*[c]	*157.0*[a]	139.0[de]	*13.52*[f]					8.35[ge]	*8.78/6.63*[h]
U11	*152.7*[a]	*169.4*[a]	102.8[d]	141.6[de]				*13.80*[f]	5.19[g]	7.80[ge]		
G12	*156.4*[a]	151.8[b]	*119.0*[c]	*161.5*[a]	136.2[de]	*12.66*[f]					7.79[ge]	*8.15/6.02*[h]
C13	156.1[b]	*168.4*[a]	97.2[d]	141.0[de]					5.30[g]	7.69[ge]		*8.67/7.06*[h]
C14	156.1[b]	*169.1*[a]	98.4[d]	141.7[de]					5.57[g]	7.69[ge]		*8.48/7.13*[h]

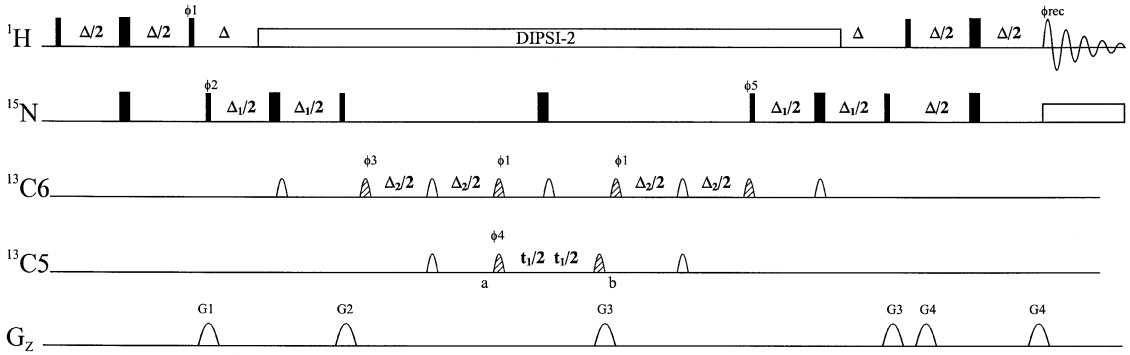

Figure 1. Pulse sequence of the HNC6C5 experiment with phase sensitive detection using States-TPPI (Marion et al. 1989). Q3 and Q5 pulses (Emsley, L. and Bodenhausen, G., 1992) were used as selective 180°- and 90°-pulses (where hatched symbols indicate 90° and empty symbols 180° pulses) with a pulse length of 500 μs and 750 μs at 700 MHz, respectively. The delays for the INEPT steps were set to $\Delta = 1/2J(H,N)$, $\Delta 1 = 1/2J(N,C6)$ and $\Delta 2 = 1/2J(C6,C5)$. The DIPSI-2 sequence (Shaka et al., 1988) was applied for decoupling of ^{1}H, and the GARP sequence (Shaka et al., 1985) for decoupling of ^{15}N during acquisition. Sine-shaped gradients were applied with gradient strength of G1 = 60%, G2 = 40%, G3 = −10%, and G4 = 80%, where 100% is approx. 55 G/cm. Phase cycle: $\phi 1 = y$; $\phi 2 = x_8,-x_8$; $\phi 3 = x,-x$; $\phi 4 = x_4,-x_4$; $\phi 5 = x_2,-x_2$; $\phi rec = (x,-x,-x,x)(-x,x,x,-x)_2(x,-x,-x,x)$.

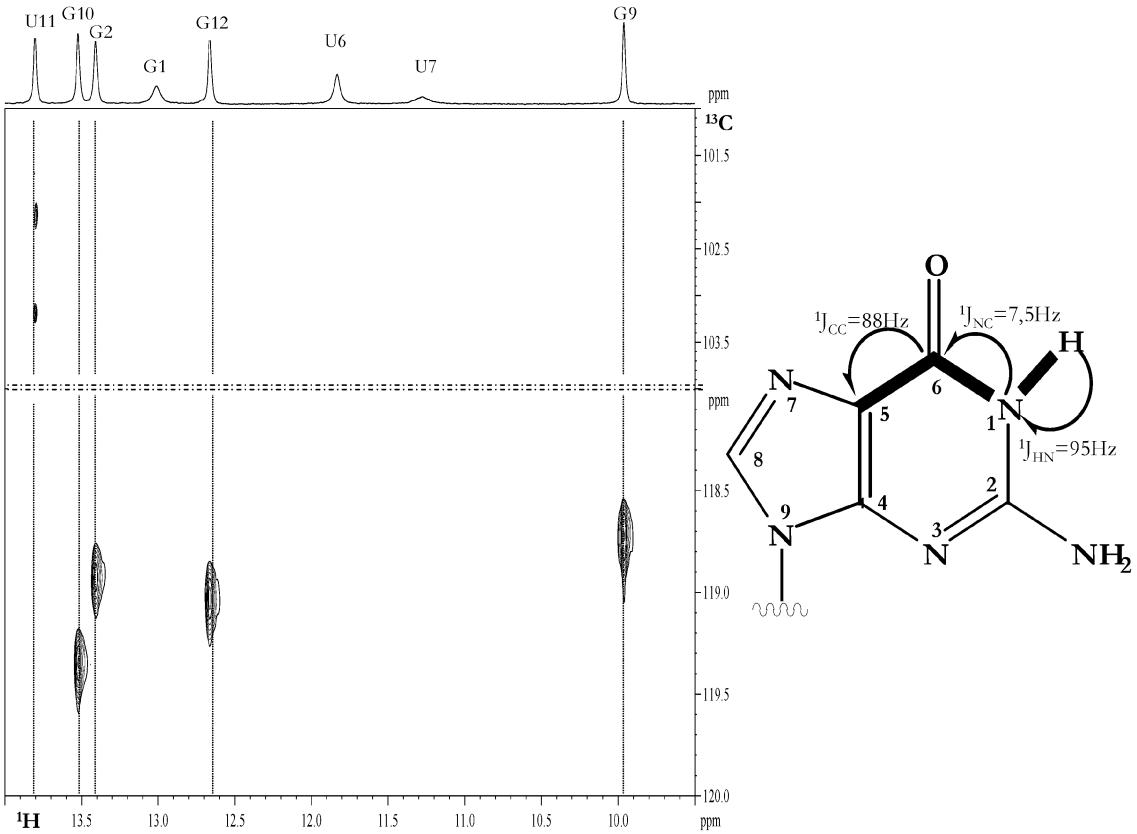

Figure 2. Spectrum of HNC6C5-experiment at 700 MHz and 10 °C. The carrier for ^{15}N was set to 154 ppm. Hard ^{1}H pulses were applied on-resonance at the water frequency. For hard ^{1}H and ^{15}N pulses, field strengths of 27.5 kHz and 7.4 kHz were used, respectively. The GARP-decoupling (Shaka et al., 1985) during acquisition was applied with a field strength of 1.3 kHz. The 2D experiment was recorded for 8.5h with t_{max} of 28 ms (100pts, complex), 63 ms (1024 pts, complex) for t_1 and t_2. The relaxation delay was set to 1.0 s. The dashed lines separate the spectrum in two parts. The upper part shows the region of ^{13}C chemical shifts between 101 and 104 ppm, whereas the lower one shows the region between 118 and 120 ppm. The pointed line connects the cross peaks with respective imino-proton resonances.

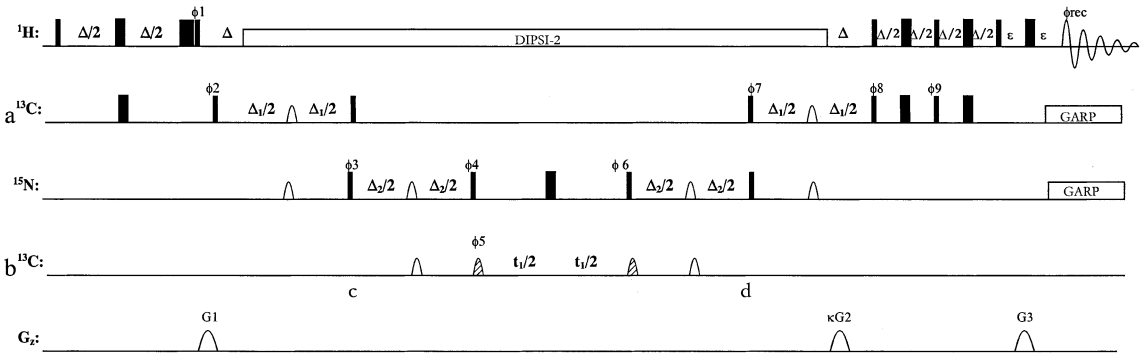

Figure 3. Pulse sequence of the HCNC-experiments with phase sensitive detection according to echo/antiecho modulation (Cavanagh et al., 1991; Palmer et al., 1991). The upper trace for the ^{13}C-channel (a) shows pulses applied at the frequency of C1′ in the H1′-C1′-N9(N1)-C4(C2)-experiment or at the frequency of the aromatic C8(C6) in the H8(H6)-C8(C6)-N9(N1)-C4(C2)-experiment. The lower trace for the ^{13}C-channel (b) represents in both cases the frequency of C4(C2). As selective ^{13}C-pulses Q3 (180°, τ_P = 2900 μs, empty symbols) and Q5 (90°, τ_P = 2500 μs, hatched symbols) (Emsley and Bodenhausen, 1992) were applied; as selective ^{15}N-pulses IBURP2 pulses (180° τ_P = 1300 μs) (Geen and Freeman, 1991) were applied. The delays for the inept steps were set to Δ = 1/2J(HC^a), $\Delta1$ = 1/2J(C^{a}N) and $\Delta2$ = 1/2J(NC^b). The default phase was x. Phase cycle: $\phi1$ = y $\phi2$ = x,−x; $\phi5$ = $(x)_{16}$,$(-x)_{16}$; $\phi7$ = x; $\phi8$ = x,x,−x,−x; $\phi9$ = y,y,−y,−y; $\phi4$ = $(y)_8$, $(-y)_{-8}$; $\phi3$ = $(x)_4$,$(-x)_4$; ϕrec = $((x,-x,-x,x,-x,x,x,-x)(-x,x,x,-x,-x,x,x,-x))_2$. Sine shaped gradients were applied with gradient strength of G1 = 50%, G2 = 80%, and G3 = 20.1%, where 100% is approx. 55 G/cm. With the change of the sign of κ, the phase $\phi9$ is shifted by 180° in every second experiment, stored separately and processed according to sensitive-enhanced echo/antiecho sign discrimination.

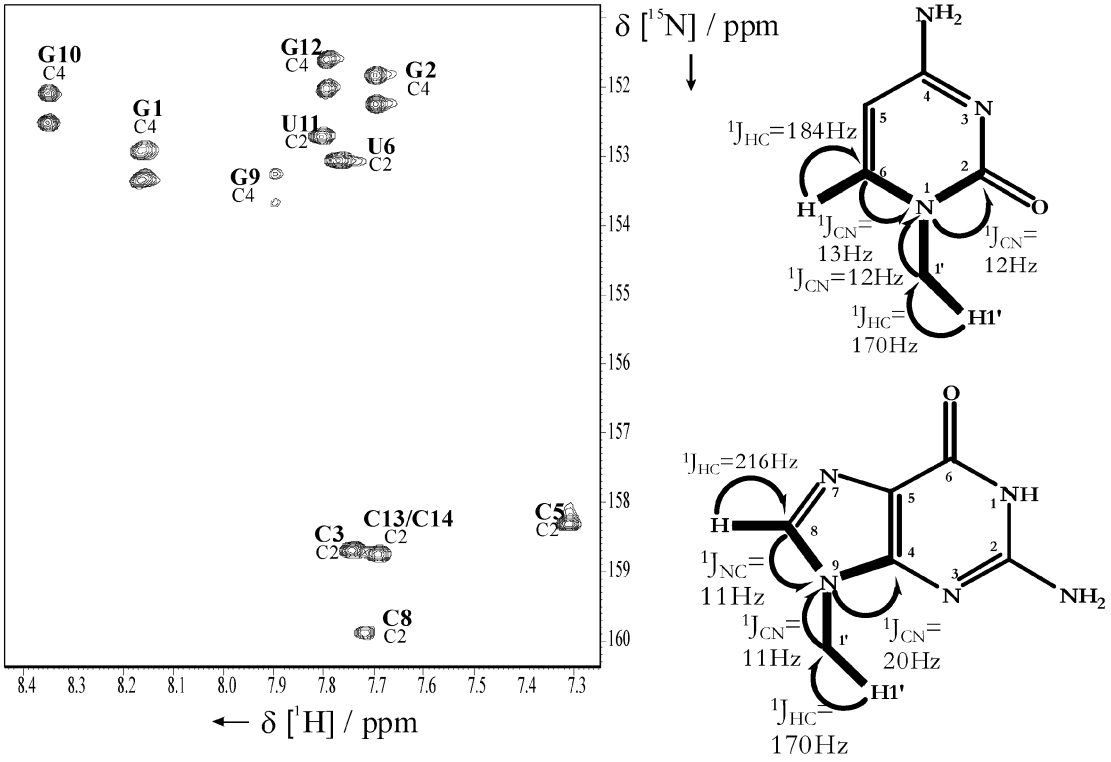

Figure 4. H8(H6)-C8(C6)-N9(N1)-C4(C2) experiment at 600 MHz and 25 °C. The carrier for ^{15}N was set to 158 ppm. The carrier for ^{13}C is shifted during the experiments from 138 ppm to 152 ppm and back to 138 ppm (at points c and d annotated in the Figure 3). All ^{1}H pulses were applied on-resonance at the resonance frequency of water. For hard ^{1}H, ^{13}C and ^{15}N pulses, field strengths of 27 kHz, 22.5 kHz and 7.5 kHz were applied, respectively. The GARP-decoupling sequence (Shaka et al., 1985) was applied during acquisition with field strength of 0.8 kHz to ^{15}N and with 1.8 kHz to ^{13}C. DIPSI-2 decoupling of ^{1}H (Shaka et al., 1988) was applied with a field strength of 3.1 kHz. The 2D experiment was recorded in 16h with t_{max} of 72 ms (64 pts, complex), 73 ms (512 pts, complex) for t_1 and t_2. The relaxation delay was set to 1.5 s.

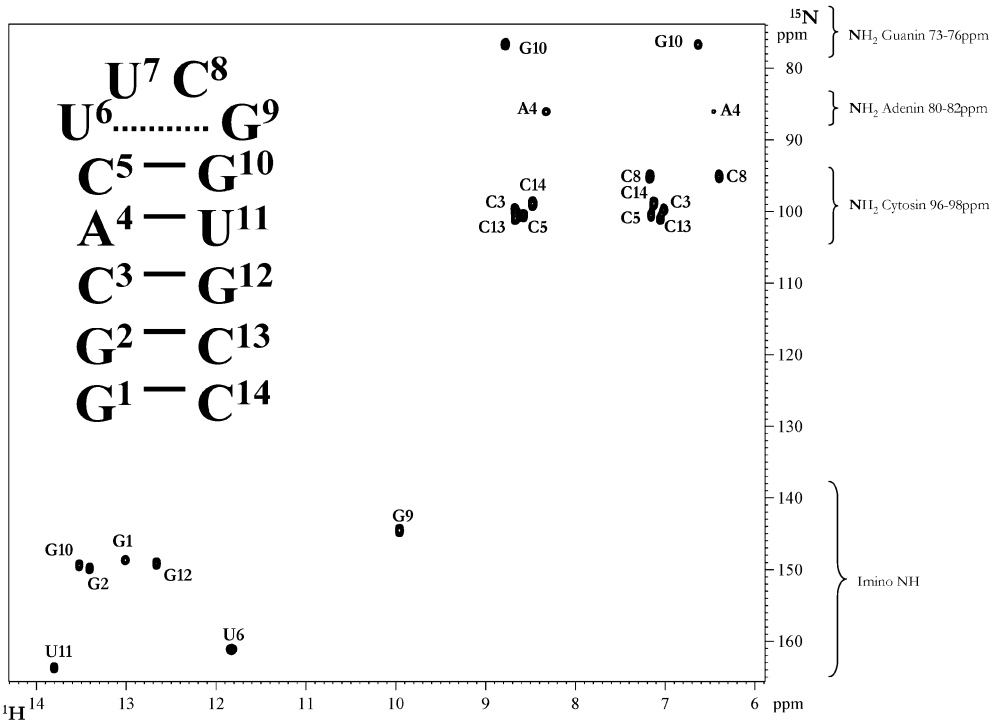

Figure 5. $^{1}H,^{15}N$-HSQC at 700 MHz with phase sensitive detection using States-TPPI (Marion et al., 1989) recorded at 10 °C with the assignments indicated. In the left part of the spectrum, the secondary structure of the tetraloop is shown. The carrier for ^{15}N was set to 110ppm. Hard pulses were applied with field strengths of 27 kHz and 7.5 kHz for ^{1}H and ^{15}N, respectively. During acquisition, GARP decoupling (Shaka et al., 1985) was applied with a field strength of 1.3 kHz. The 2D experiment was recorded in 1.5 h with t_{max} of 7.6 ms (64 complex pts, complex), 61 ms (512 complex pts) for t_1 and t_2. The relaxation delay was set to 1.3 s.

z-axis gradient $^{1}H\{^{13}C,^{31}P\}$ triple resonance probes. Spectra were processed with XwinNMR 3.5 (Bruker) and analyzed with felix2000 (MSI). Resonance assignments were obtained from triple resonance 2D and 3D NMR experiments such as 2D $^{1}H,^{13}C$ CT-HSQC, 2D $^{1}H,^{15}N$ ^{1}J-FHSQC, 2D $^{1}H,^{15}N$ ^{2}J-FHSQC (Mori et al., 1995), 3D HCP (Marino et al., 1994), 3D HCN (Sklenar et al., 1993), 3D HCCH-TOCSY (Kay et al., 1993), 3D forward directed HCC-TOCSY-CCH-COSY (Schwalbe et al., 1995), 2D H(N)CO (Kay et al., 1994), 3D^{15}N-HSQC-NOESY (Sklenar et al., 1993), $^{1}H,^{15}N$-CPMG-NOESY (Mueller et al., 1995) as well as 2D TROSY relayed HCCH-COSY (Simon et al., 2001). Experiments used to assign the nuclei are indicated in the Tables 1–3 by superscripts.

New NMR pulse sequences

The following new experiments are described: 2D HNC6C5, 2D H8(H6)-C8(C6)-N9(N1)-C4(C2) and 2D H1′-C1′-N9(N1)-C4(C2). The names of these pulse sequences are derived from the correlated atoms, names given in parenthesis are the atoms in pyrimidine nucleotides, names given without parenthesis are atoms in purine nucleotides.

The 2D HNC6C5 experiment

For the resonance assignment of the quaternary carbon C5 in guanine, a triple-resonance 2D experiment similar to the HNCOCA sequence (Muhandiram et al., 1994) has been developed (Figure 1). In an out-and-back-manner, magnetization is transferred from the imino proton via the imino nitrogen and the carboxyl carbon atom to the C5 ($^{1}J(N,H^{N})$ ∼95Hz; $^{1}J(N,C6)$ ∼7.5 Hz; $^{1}J(C6,C5)$ ∼88 Hz), whose chemical shift evolves in t_1. After backtransfer, the imino resonances are detected during t_2. Coherences of carbons C5 (115–120 ppm) and C6 (160–165 ppm) have been differentiated using shaped pulses. The use of these selective pulses also prevents the magnetization transfer from nitrogen to carbon C2.

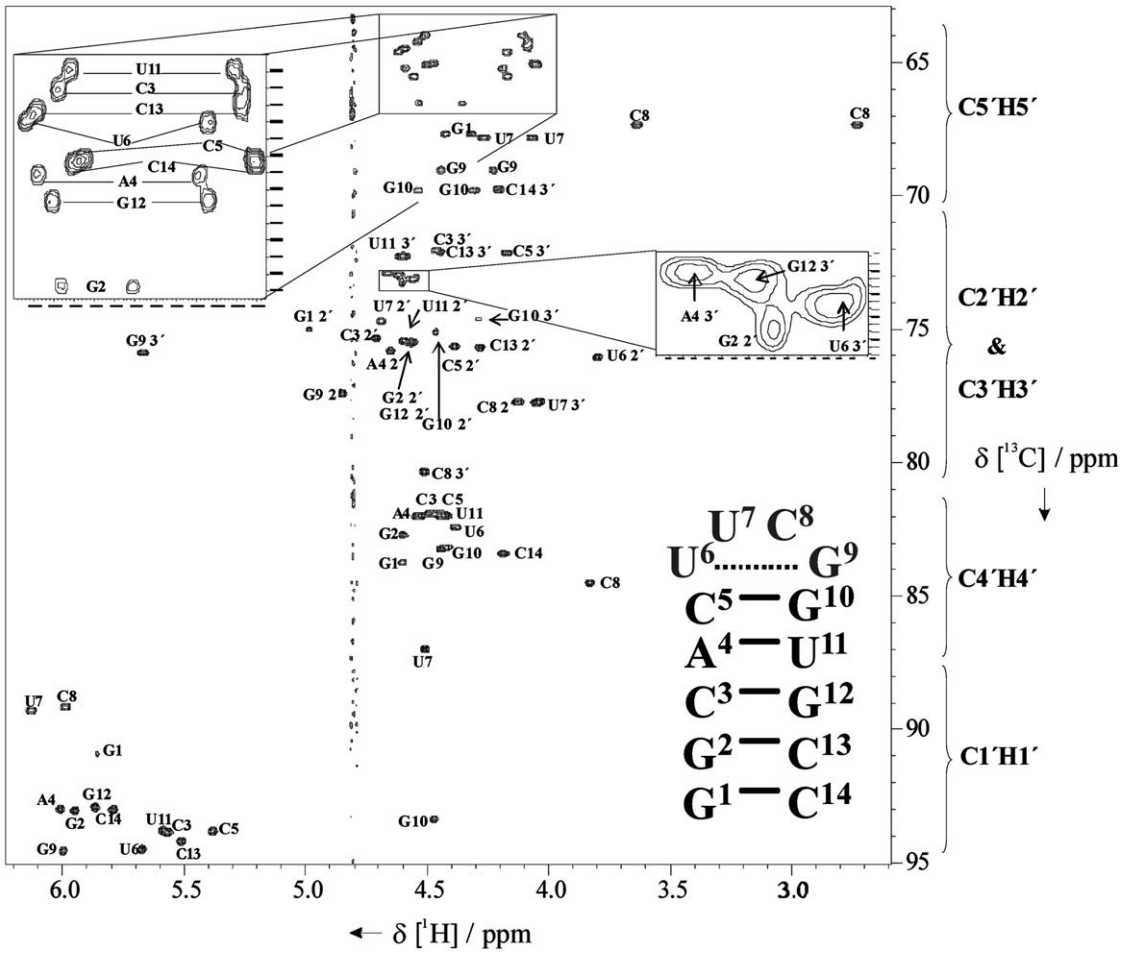

Figure 6. $^{1}H,^{13}C$-CT-HSQC at 600 MHz with the assignments of the sugar region indicated. The carrier for ^{13}C was set to 106 ppm. All proton pulses were applied on-resonance at the water frequency. Hard pulses were applied with a field strength of 29 and 19 kHz for proton and carbon, respectively. During acquisition, GARP decoupling was applied with a field strength of 3.8kHz. The 2D experiment was recorded in 2 h with t_{max} of 17.6 ms (256 complex pts), 142 ms (1024 complex) for t_1 and t_2. The relaxation delay was set to 1.5 s.

In the resulting spectrum (Figure 2), the resonances of the quaternary carbons C5 (as singlet due to selective decoupling of the neighboring carbon atoms) of nucleotides G2, G9, G10, and G12 and, in addition, carbon C6 of the only based paired U nucleotide U11 (as a doublet due to $^{1}J(C5,C6)$) can be observed. Four of the five quaternary C5 carbons could therefore be assigned, the imino resonance of the terminal G1 is too broad to be observed. Nearly uniform chemical shift values are observed for the carbons C5 along the sequence. The average value is similar to the chemical shift of carbon C5 in a GTP mononucleotide dissolved in the same buffer ($\delta C5^{GTP}$:119 ppm).

The 2D H8(H6)-C8(C6)-N9(N1)-C4(C2) and 2D H1′-C1′-N9(N1)-C4(C2) experiment

For the assignment of the quaternary carbons C2 in pyrimidines and C4 in purines, two new pulse sequences have been developed that either correlate the resonances of the sugar H1′ protons or the resonances of the aromatic H8/H6 protons with C4 and C2 in purines and pyrimidines, respectively. The experiment is of the out-and-back-type (Figure 3). The flow of magnetization is briefly discussed for the experiment correlating H1′ with N9 and C4 in purines. It consists of three subsequent INEPT-steps exploiting the scalar coupling constants $^{1}J(H1',C1')$ ~170 Hz, $^{1}J(C1',N9)$ ~10–12 Hz, and $^{1}J(N9,C4)$ ~20 Hz. For the transfer from H8 or H6 to the directly bound carbon, delays

Table 3. ^{15}N-chemical shift assignments of the nucleobases. Entries given in italics were measured at 10 °C, all other chemical shift values were measured at 25 °C. Chemical shifts were referenced to TSP. Superscripts indicate the experiments which were used for chemical shift assignment: a: ^{15}N-HSQC, b:HCN, c: ^{2}J-^{15}N-HSQC, d: HNN-COSY, e: HCCN

	N1 (ppm)	N2 (ppm)	N3 (ppm)	N4 (ppm)	N6 (ppm)	N7 (ppm)	N9 (ppm)
G1	*148.6*[a]		*161.7*[d]			233.1[c]	169.5[bc]
G2	*149.7*[a]		*162.6*[d]			234.3[c]	170.5[bc]
C3	151.9[b]		198.1[d]	*99.7*[a]			
A4	214.6[c]		223.6[c]		*86.1*[a]	231.6[c]	172.2[bc]
C5	152.4[b]		197.4[d]	*100.9*[a]			
U6	148.1[b]		*161.0*[a]				
U7	145.1[b]		*159.5*[a]				
C8	151.9[b]		204.3[e]	*95.1*[a]			
G9	*144.4*[a]		*171.7*[d]			232.9[c]	172.9[bc]
G10	*149.3*[a]	*76.7*[a]	*165.5*[d]			234.5[c]	171.7[bc]
U11	147.6[b]		*163.6*[a]				
G12	*149.1*[a]		*162.9*[d]			235.6[c]	170.8[bc]
C13	152.4[b]		199.1[d]	*100.6*[a]			
C14	153.7[b]		197.8[d]	*98.9*[a]			

Table 4. Canonical coordinates *can*1 and *can*2 (ppm) as defined in the text

Nuc.	*can1* (ppm)	*can2* (ppm)
G1	−6.5	−17.3
G2	−5.9	−17.1
C3	−5.4	−16.9
A4	−5.6	−17.0
C5	−5.5	−16.9
U6	−5.4	−17.0
U7	−7.6	−17.6
C8	−7.0	−17.6
G9	−5.8	−17.5
G10	−6.1	−17.3
U11	−5.4	−16.8
G12	−5.7	−17.0
C13	−5.4	−16.9
C14	−5.9	−16.8

were tuned to ^{1}J(H8,C8) ∼220 Hz and ^{1}J(H6,C6) ∼185 Hz, respectively, while for transferring magnetization from N1 to C2 in cytosine and uracile, the delay was optimized based on a scalar coupling ^{1}J(N1,C2) ∼12 Hz. The transfer between glycosidic nitrogen N1 or N9 and the quaternary carbons C2 and C4 was achieved using selective carbon pulses. During the first evolution period, no decoupling for carbons was applied.

In the spectrum of such a type of experiment (Figure 4), correlation peaks are observed between H8/H6 and the corresponding C4 and C2 resonances as annotated in the figure. The assignment of the quaternary carbons is possible by relating the spectrum to the assigned regions H6C6 or H8C8 (or H1′C1′ for the other experiment) of a constant-time ^{1}H,^{13}C-HSQC. C4 resonances of guanines are observed as doublets due to modulation with ^{1}J(C4,C5).

The S/N ratio is better for the experiment starting at the aromatic hydrogens. Nevertheless, both pulse sequences could also be applied to an RNA 30mer (Ohlenschläger et al., 2003) with ^{13}C and ^{15}N labeled guanine and cytosine residues; for the 30mer, resonance assignment for 14 out of 16 possible resonances could be achieved (data not shown). According to the analysis using only two different data sets, the chemical shifts of the carbon C4 in guanines seem to vary depending on the extent of base stacking. Although the maximum difference of chemical shift values is just 1.7 ppm, resonances for guanines involved in base stacking are shifted upfield compared to non-base-stacked nucleotides and to mononucleotides ($\delta C4^{GTP}$:154.6 ppm).

Extent of NMR resonance assignment

For the 14 residues of the UUCG tetra loop, essentially complete assignments could be obtained (Figures 5 and 6). Of all ^{1}H (except 0H-resonances), ^{13}C, ^{15}N, and ^{31}P resonances, 324 resonances could be assigned, corresponding to 97% of all resonances. Not assigned are the resonances of the NH_2-groups of guanines G1, G2, and G9 which are broadened by exchange beyond detection.

The assigned ^{1}H, ^{15}N, ^{13}C, ^{31}P chemical shifts and a number of scalar coupling constants of the cUUCGg tetraloop RNA have been deposited in the BioMagResBank under accession number BMRB-5705. The resonance assignment found for nucleotides C5 to G10 forming the apical tetraloop are similar to the resonance assignment by the Varani group (Allain et al., 1995b) for an cUUCGg tetraloop with different stem residues.

Analysis of chemical shift data

For RNA molecules, chemical shift analysis is espe-

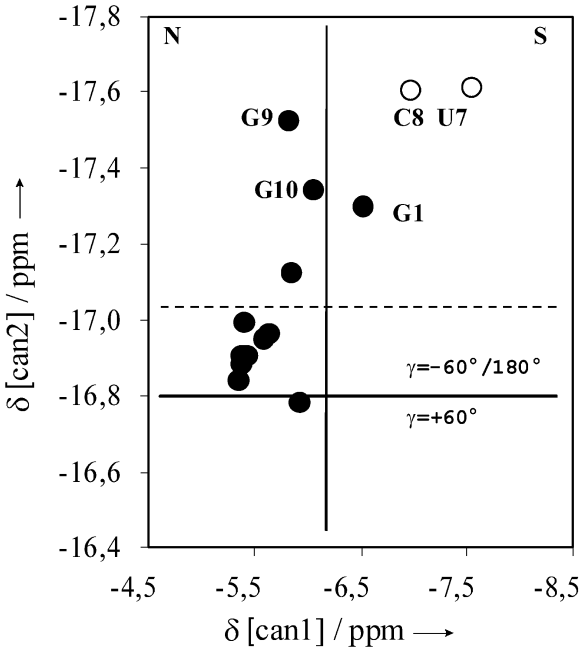

Figure 7. Plot of canonical coordinate *can*2 versus *can*1. Canonical coordinates are defined in equation 1 in the text following Ebrahimi et al. (2001). N and S annotate the North and the South conformation of the ribose moiety respectively, and gg and gt give the conformation around the exocyclic torsion angles γ (gg stands for $\gamma = 60°$; gt stands for $\gamma = -60°$).

cially valid, since deviations from average chemical shifts are predominantly observed for residues deviating from canonical A-form helical conformation such as the loop region of the 14mer RNA (Cheong et al., 1990; Varani et al., 1991; Allain & Varani, 1995b). For example, the resonance frequencies of C2′H2′ of U6, of C4′H4′ of U7, of C5′H5′/C5′H5″ of C8 and of N1H1 of guanine 9 differ from those of the atoms in canonical regions. The observed chemical shift deviations can in part be explained by inspection of the structure: The chemical shifts of C2′H2′ of U6 and N1H1 of G9 can be due to the unusual hydrogen bonds between U6(O2) and G9($N2H_2$) and between U6(O2′H2′) and G9(O6). This unusual hydrogen bonding pattern results also in the ability to assign the chemical shift of the hydrogen in the 2′-hydroxy group of U6 in the NOESY spectrum (δ2′O$\mathbf{H}^{U6}$:6.8 ppm). The unusual chemical shifts of C4′H4′ of U7 and C5′H5′/C5′H5″ of C8 can be the result of a strong variation in the backbone conformation of the loop. The downfield shift of resonances C3′H3′ and C2′H2′ of G9 can be caused by base stacking effects.

The dependence of ^{1}H chemical shifts on secondary structure has been analyzed in detail for a number of different RNAs including the UUCG tetraloop (Cromsigt et al., 2001). Here, we focus on the analysis of the ^{13}C- and ^{31}P-chemical shift data. ^{13}C-chemical shift data yielding the sugar pucker modes could be obtained following the analysis of solid-state NMR data by Ebrahimi et al. (Ebrahimi et al. 2001), which is based on the measurement and analysis of ^{13}C-chemical shifts. For this analysis, canonical coordinates were calculated using the chemical shift data δ.

$$\begin{aligned} \text{can1} &= 0.179\delta_{C1'} - 0.225\delta_{C4'} - 0.0585\delta_{C5'} \\ \text{can2} &= -0.0605(\delta_{C2'} + \delta_{C3'}) - 0.0556\delta_{C4'} - 0.0524\delta_{C5'} \end{aligned} \quad (1)$$

The first canonical coordinate *can*1 describes the conformation of the sugar. For *can*1 > -6.25 ppm, the sugar is in a North-conformation, for *can*1 < -6.25 ppm, the sugar adopts a South-conformation. The second coordinate *can2* determines the conformation of the exocyclic torsion angle γ. If the torsion angle γ is in a gauche-gauche conformation ($\gamma = +60°$), *can2* > -16.8 ppm. Population of either of the two gauche-trans conformations ($\gamma = -60°$,$180°$) results in a coordinate *can*2 larger than -16.8 ppm. Application of these rules to the chemical shift data of the UUCG tetraloop allows determination of the sugar pucker mode of all residues (Table 1, Figure 7). According to this analysis, residues cytosine 8 and uracile 7 adopt a C2′-endo conformation; while all other residues are in a C3′-endo conformation. The coordinate analysis fails for residue guanine 1, may be due to a higher conformational flexibility at the 5′-end of the stem –that can be detected by the analysis of the ^{3}J(H1′H2′) and ^{3}J(H3′H4′) coupling constants (to be published elsewhere)- or due to the additional charge of the 5′-terminal phosphate group. These results are in good agreement with the published structures and with the analysis of scalar ^{3}J(H,H)-coupling constants and cross-correlated relaxation rates in the ribose ring (to be published elsewhere).

Discrimination between gg- and gt-conformations on the basis of *can*2 is less convincing. Most of the values of *can*2 are clustered in the region between -16.6 ppm and -17.0 ppm. Only the nucleotides guanine 1, uracile 7, cytosine 8, guanine 9 and guanine 10 are clearly in the gt-conformation. The uncertainty of the analysis can be due to the simultaneous use of the chemical shift of C2′ and C3′ in the calculation. For the 14mer, the analysis would predict residues

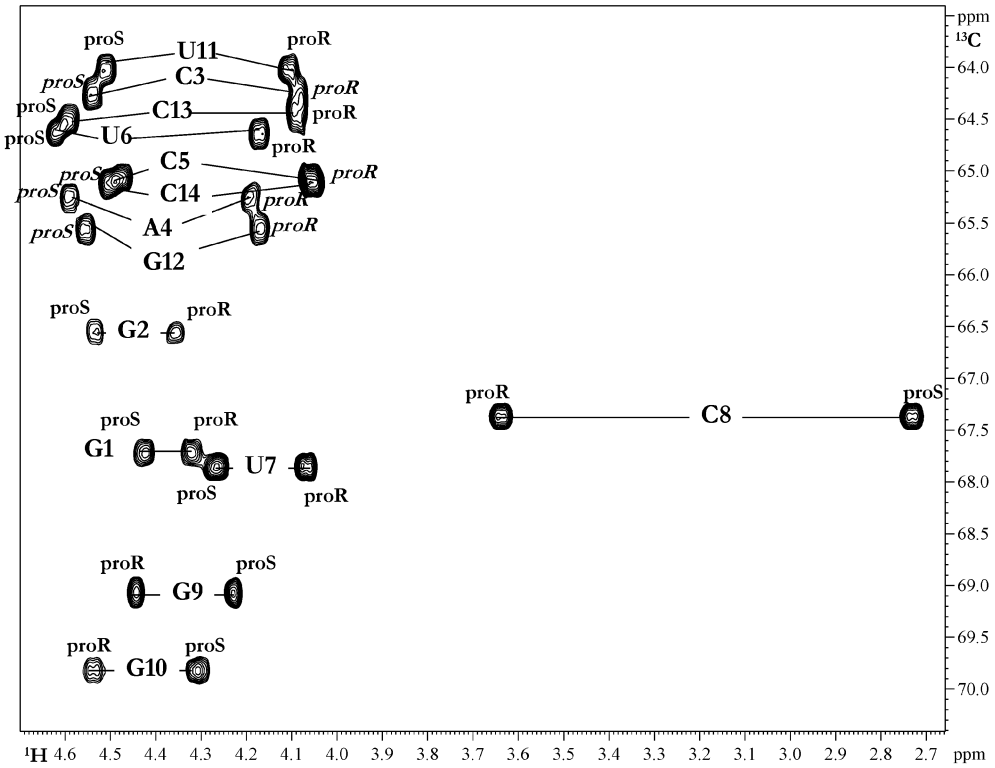

Figure 8. Expanded plot of the C5′H5′ region of a ^{1}H,^{13}C-CT-HSQC performed on the UUCG-tetraloop with the assignments annotated. The stereochemical assignment is given in italics if derived from chemical shift arguments alone; all other assignments have been verified by analysis of ^{2}J(C4′,H5′/H5″)- and ^{3}J(H4′,H5′/H5″)-coupling constants.

with can2 > -17 ppm to be in a gg-conformation around the exocyclic torsion angles γ.

By analyzing the chemical shift data of the C5′H5′/C5′H5″ resonances, a stereo-specific assignment of the prochiral H5′/H5″ protons could be obtained. It is observed that in helical RNA structures the resonance of the H5′(proS) proton is up-field shifted with respect to the resonance of the H5′(proR) proton (Remin et al., 1972). Unfortunately, this rule is not applicable to non-canonical regions of RNA structures. However, a stereo-specific assignment is possible by correlation of the difference in the proton chemical shift of the proS and proR protons with the respective carbon chemical shift $\Delta\delta$[H5(proS)-H5(proR)] (δ^{13}C) (Marino et al., 1996). If one depicts the differences of the proton chemical shifts versus the carbon chemical shift, the stereochemical assignment is revealed. The results are in good agreement with the data obtained by the analysis of the ^{2}J(C4′,H5′/H5″)- and ^{3}J(H4′,H5′/H5″)-coupling constants (to be published elsewhere). As shown in Figure 9, a general anticorrelation could be observed for the carbon chemical shift of these resonances. An exception is the residue cytosine 8 which is probably a result of the different conformation of γ at this loop position.

^{31}P-chemical shifts and their deviations from standard chemical shifts in A-form RNA are often taken to be indicative for unusual conformations around the phosphodiester backbone. Significant variations of the phosphorous chemical shift from the mean value of twelve assigned RNA molecules deposited in the BMRB-database are observed for the loop residues uracile 7, cytosine 8, guanine 9 and the nucleotide guanine 10 that is part of the closing base pair (Figure 10). The chemical shifts of the residues in canonical A-form conformation are not significantly different from the mean value.

Summary

With the application of standard and new NMR-experiments, a complete resonance assignment of an 14mer RNA containing an UUCG-tetraloop was obtained. The analysis of chemical shift data revealed a

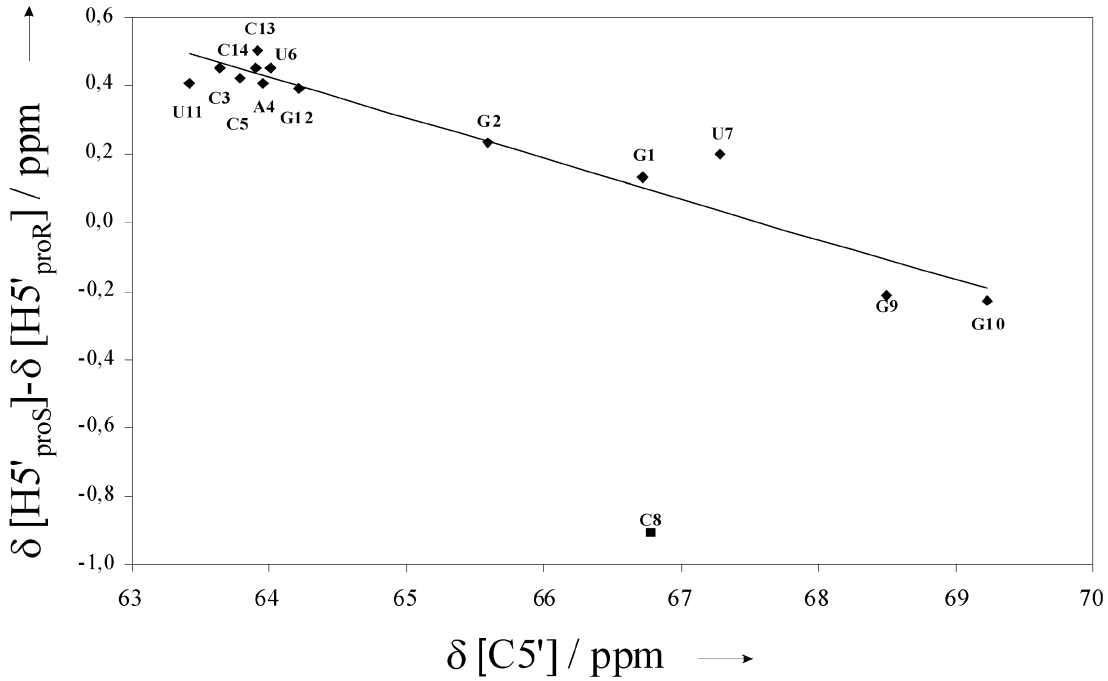

Figure 9. Plot of the difference of the proton chemical shift of the H5′(proS)proton and H5′(proR)proton versus carbon chemical shift of C5′. The fitted line has a correlation coefficient R of 0.92.

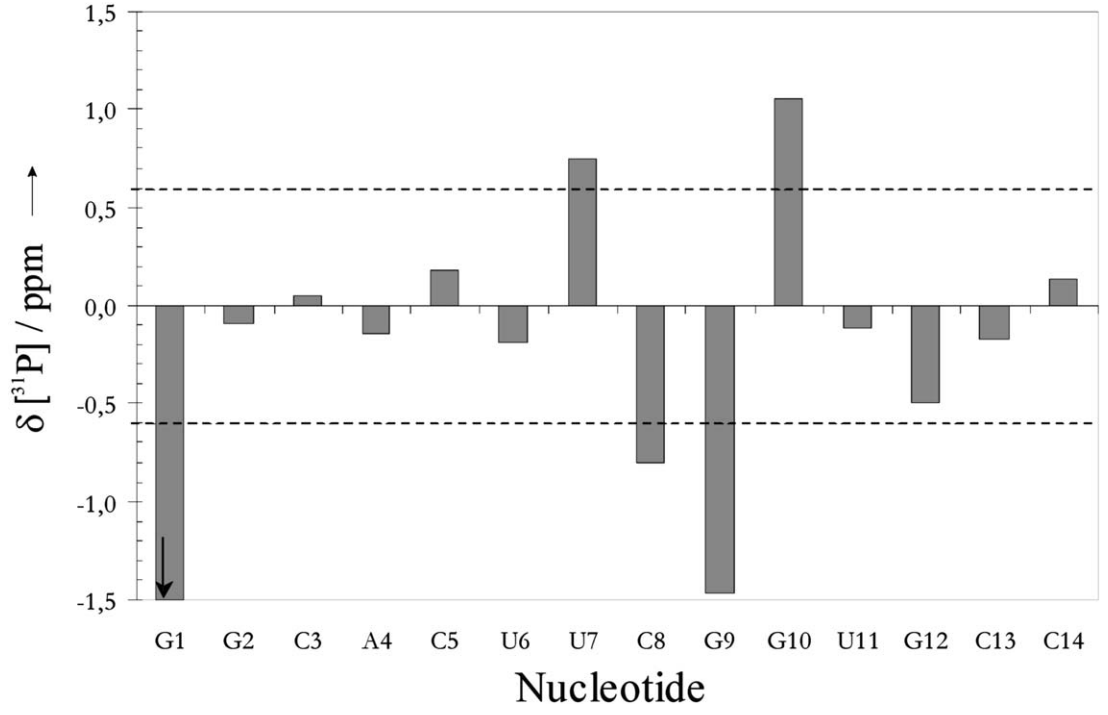

Figure 10. Deviation of the ^{31}P-chemical shift from the mean value of ^{31}P-chemical shifts from twelve different RNA-molecules deposited in the BMRB-database. The standard mean deviation is 0.6 ppm not regarding the deviation of the chemical shifts of the phosphorous atoms of the triphosphate of guanine 1.

first insight into the conformation of the tetraloop. In contrast to proteins, the number of published chemical shift assignments for RNAs is scarce. However, due to the smaller number of nucleotide building blocks and the intrinsic correlation of various torsion angles, statistical analysis of chemical shift data for secondary structure prediction is expected to be powerful.

Acknowledgements

We thank Prof Vladimír Sklenář and Dr Radek Marek for helpful discussions about the initial assignment. We are very thankful to Dr Jens Wöhnert for providing an RNA sample and for helpful discussions. This work was supported by the European Large Scale facility for Biomolecular NMR (HPRI-1999-CT-00014).

References

Allain, F.H.-T. and Varani, G. (1995a) *J. Mol. Biol.*, **250**, 333–353.
Allain, F.H.-T. and Varani, G. (1995b) *Nucl. Acid Res.*, **23**, 341–350.
Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) *Science*, **289**, 905–920.
Cavanagh, J., Palmer, A.G., Wright, P.E. and Rance, M. (1991) *J. Magn. Reson.*, **91**, 429–436.
Cheong, C., Varani, G. and Tinoco, I.J. (1990) *Nature*, **346**, 680–682.
Cromsigt, J.A.M.T.C., Hilbers, C.W. and Wijmenga, S.S. (2001) *J. Biomol. NMR*, **21**, 11–29.
Ebrahimi, M., Rossi, P., Rogers, C. and Harbison, G.S. (2001) *J. Magn. Reson.*, **150**, 1–9.
Emsley, L, Bodenhausen, G. (1992) *J. Magn. Reson.*, **97**, 135–148.
Ennifar, E., Nikulin, A., Tishchenko, S., Serganov, A., Nevskaya, N., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S. and Dumas, P. (2000) *J. Mol. Biol.*, **304**, 35–42.
Geen, H. and Freeman, R. (1991) *J. Magn. Reson.*, **93**, 93–141.
Harris, K.R., Becker, E.D., Cabral de Menezes, S.M., Goddfellow, R. and Granger, P. (2001) *Pure Appl. Chem.*, **73**, 1795–1818.
Kay, L.E., Xu, G.X. and Yamazaki, T. (1994) *J. Magn. Reson.*, **A109**, 129–133.
Kay, L.E., Xu, G.Y., Singer, A.U., Muhandiram, D.R. and Forman-Kay, J.D. (1993) *J. Magn. Reson.*, **B101**, 333–337.
Lynch, S.R., Pelton J.G. and Tinoco, I.J. (1996) *Magn. Reson. Chem.*, **34**, 11–17.
Marino, J.P., Schwalbe, H., Anklin, C., Bermel, W., Crothers, D.M. and Griesinger, C. (1994) *J. Am. Chem. Soc.*, **116**, 6472–6473.
Marino, J.P., Schwalbe, H., Glaser, S.J. and Griesinger, C. (1996) *J. Am. Chem. Soc.*, **118**, 4388–4395.
Marion, D., Ikura, R., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.
Mori, S., Abeygunawardana, C., O'Neil-Johnson, M. and van Zijl, P.C.M. (1995) *J. Magn. Reson.*, **B108**, 94–98.
Mueller, L., Legault, P. and Pardi A. (1995) *J. Am. Chem. Soc.*, **117**, 11043–11048.
Nissen, P., Hansen, J., Ban, N., Moore, P. B. and Steitz, T.A. (2000) *Science*, **289**, 920–930.
Ohlenschläger, O., Wöhnert, J., Bucci, E., Sidigi, K., Häfner, S., Zell, R. and Görlach, M. (2003) submitted.
Palmer, A.G., Cavanagh, J., Wright, P.E. and Rance, M. (1991) *J. Magn. Reson.*, **93**, 151–170.
Remin, M. and Shugar, D. (1972) *Biochem. Biophys. Res. Commun.*, **48**, 636–642.
Schwalbe, H., Marino, J.P., Glaser, S.J. and Griesinger, C. (1995) *J. Am. Chem. Soc.*, **117**, 7251–7252.
Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.
Shaka, A.J., Lee, C.J. and Pines A. (1988) *J. Magn. Reson.*, **77**, 274–293.
Simon, B., Zanier, K. and Sattler, M. (2001) *J. Biomol. NMR*, **20**, 173–176.
Sklenář, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1993) *J. Biomol. NMR*, **3**, 721–727.
Sklenář, V., Piotto, M., Leppik, R. and Saudek, V. (1993) *J. Magn. Reson.*, **A102**, 241–245.
Varani, G., Cheong, C. and Tinoco, I.J. (1991) *Biochemistry*, **30**, 3280–3289.
Wimberly, B.T., Brodersen, D.E., Clemons, Jr., W.M., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T. and Ramakrishnan, V. (2000) *Nature*, **407**, 327–339.
Wishart, D.S., Bigam, C.G., Yao J., Abildgaard F., Dyson H.J., Oldfield E., Markley J.L. and Sykes B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.