Supporting online material:

A procedure to validate and correct the ^{13}C chemical shift calibration of RNA datasets

Thomas Aeschbacher, Mario Schubert, Frédéric H.-T. Allain

Supplementary Methods

Preparation of RNA stem-loops by in vitro transcription

Six RNAs were synthesized by run-off transcription using synthetic DNA templates (Price et al. 1998): two stem-loops of the small non-coding RNA RSMZ from *Pseudomonas fluorescens*, the 22 nt stem-loop FZL2 with the sequence 5′-GGGCCAUCAAGGACGAUGGUCC-3′, the 21 nt stem-loop FZL4 with the sequence 5′-GGGUCAUCAGGACGAUGACCC-3′, the 20 nt stem-loop harboring the Shine-Dalgarno sequence from the hcnA mRNA of *Pseudomonas fluorescens* with the sequence 5′-GGGCUUCACGGAUGAAGCCC-3′, the 22 nt artificial stem-loop TASL1 with the sequence 5′-GGGAUUCAUUUCGAUGAAUCCC-3′, the 26 nt artificial stem-loop TASL2 with sequence 5′-GGGAUUCAUGCUUCGGCAUGAAUCCC-3′ and the 30 nt artificial stem-loop TASL3 with sequence 5′-GGGAUUCAUGCACUUCGGAGCAUGAAUCCC-3′. Secondary structures of these RNAs are shown in Supplementary Fig. 1.

The DNA templates were obtained from Microsynth AG (Balgach, Switzerland). We engineered GGG at the 5′-end, since purine-rich sequences result in high efficiency of transcription. RNA transcription was accomplished using recombinantly expressed T7 RNA polymerase, 1M $MgCl_2$, NTP mixture with a concentration of 80 mM for each NTP, 20x TB-Buffer (800mM TrisHCl pH = 8.1, 20 mM spermidine, 0.2 % Triton X-100, 100 mM DTT) and, in some cases, pyrophosphatase and GMP following published protocols (Price et al. 1998). To find the optimal transcription conditions for all the constructs, 40 µl test reactions were performed with varying $MgCl_2$, DNA template, NTP and T7 polymerase concentration as well as in the presence and absence of GMP and pyrophosphatase. Transcription was performed at 37°C for 4 hours, and stopped with 0.5 M EDTA in a slight excess over the $MgCl_2$ concentration. The best conditions were scaled up to 20 ml for large-scale transcriptions. After stopping the reaction with EDTA, the RNA sample was passed through a 0.22 µm Millex GP filter. The RNA was purified from the reaction mixture using a Dionex DNAPAC PA-100 anion exchange column 22*250 mm combined with a Dionex DNAPAC PA-100 anion exchange column 22*50 mm at 84°C under denaturing conditions. The running buffer was composed of 12.5 mM Tris-HCl (pH 7.4) and 6M urea, the elution buffer contained 6M urea, 12.5 mM Tris-HCl (pH 7.4) and 500mM $NaClO_4$. After injecting, a flow of 20 ml/min pure running buffer was used for six minutes, followed by a mixture of 90 % washing buffer and 10 % elution buffer (which results in a $NaClO_4$ concentration of 50mM) was applied for 5 minutes. For elution, a ramp with increasing concentration of $NaClO_4$ (up to 400mM) was applied over 17 minutes. The RNAs of 20–30 nts elute typically after 11–13 minutes.

The RNA was concentrated and purified by Butanol extraction (Sawadogo and Van Dyke 1991). The RNA pellet was then dissolved with 1 ml D_2O, heated to 95°C for five minutes, frozen in liquid nitrogen, and lyophilized. The RNAs were dissolved in either D_2O or H_2O, with 5 % D_2O for NMR measurements. In addition salt concentrations and pH values were varied for stem-loop TASL2 (Supplementary Table 1).

The origin of the 2.66 ppm offset

Note that the default referencing on Bruker spectrometers is to TMS whereas biomolecules should be referenced via the absolute ^{1}H frequency of DSS multiplied by the ratio 0.251449530 yielding the absolute ^{13}C frequency of DSS that is set to 0 ppm (Markley et al. 1998). The difference of the ^{13}C frequency in the internal Bruker nucleus table (Topspin or Xwinnmr: ../exp/stan/nmr/lists/nuclei.all) and the ^{13}C DSS frequency calculated from 0.251449530 × the ^{1}H frequency from the Bruker nucleus table results in a 2.66 ppm difference in the ^{13}C chemical shifts. The nucleus table contains 100.000000 MHz for ^{1}H and 25.145020 MHz for ^{13}C. The ^{13}C DSS frequency is 25.1449530 (calculated as 0.251449530 × 100.000000 MHz) resulting in a difference to the value in the nucleus table of 67 Hz. This corresponds to 2.66 ppm (67 Hz/25.1449530 MHz). In conclusion the 2.66 ppm offset occurs when the default Bruker referencing for ^{13}C is applied.

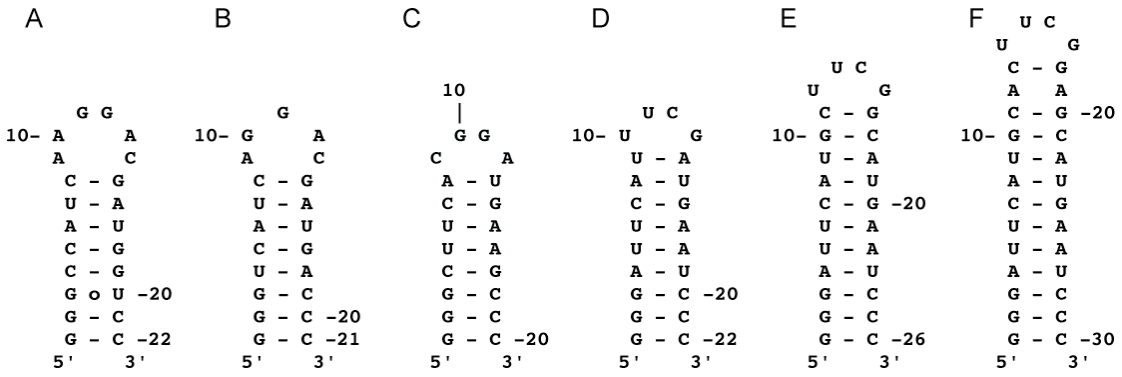
Step-by-step procedure for indirect ^{13}C referencing with the Bruker software Topspin

A small amount of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) is added to the sample resulting in a concentration of 0.2–2 mM and a ^{1}H-1D spectrum is recorded in order to determine the absolute frequency of the DSS methyl resonance. If addition of DSS is undesirable for whatever reason, DSS signal of the Bruker standard sample 2 mM Sucrose/0.5 mM DSS in H_2O/D_2O can be used as external reference. This is a reasonable alternative since RNA samples are typically measured without salt or at very low buffer concentrations so that the conditions are comparable to the sucrose sample. The 1D spectrum is referenced in Topspin by the command 'cal', selecting 'manual calibration' and setting the DSS signal to 0 ppm. Topspin stores then the absolute frequency of the DSS signal under the parameter SF displayed with the command 'edp'. This value should be the sum of the basis frequency BF1 and the spectrum reference frequency SR. The absolute standard frequencies for the indirect dimensions are calculated from the SF (^{1}H) by multiplying with the recommended factors for indirect referencing (Markley et al. 1998; Wishart et al. 1995). SF (^{13}C) is then SF (^{1}H) multiplied by 0.251449530. Any spectrum recorded under the same conditions can now be referenced by the command '1s sf' which displays the SF values of all dimensions that can now be edited. This way the spectrum reference frequency SR is set automatically. Note that it is important that the external reference must be measured at the same temperature and locked to the same solvent as the spectrum that is being calibrated.

Step-by-step procedure for indirect ^{13}C referencing with Varian software

A ^{1}H-1D spectrum containing a DSS methyl signal (see in the above Bruker section) is used to extract the absolute ^{1}H frequency of 0 ppm, $\omega_0(^{1}H)$. The cursor is set on the DSS signal and the transmitter offset 'tof' is temporarily set to this position using the command 'movetof'. The command 'spcfrq' gives then the absolute ^{1}H frequency corresponding to 0 ppm. In addition the absolute frequencies of the decoupler channels are given corresponding to the decoupler offsets dof, dof2, dof3. The absolute ^{13}C frequency corresponding to 0 ppm is calculated by the recommended factor for indirect referencing ($\omega_0(^{13}C) = \omega_0(^{1}H) \cdot 0.251449530$) (Markley et al. 1998; Wishart et al. 1995). The correct chemical shift of the center line of the ^{13}C axis is then calculated by $(\omega_{dof} - \omega_0)/ \omega_{dof}$. Agilent users who have installed the free optional BioPack software package can automatically perform these calculations by placing the

graphical user interface cursor on the ^{1}H DSS methyl resonance and selecting the 'Cursor on DSS' button in the 'Setup' panel and 'Globals and Probefile' subpanel, or using the macro 'BPfindfrqs' directly from the VNMRJ command line.



Supplementary Figure 1: RNA stem-loops that were used to derive additional ^{13}C chemical shifts. A) stem-loop 2 of the small non-coding RNA RsmZ from *Pseudomonas fluorescens* named FZL2. B) stem-loop 4 of the small non-coding RNA RsmZ from *Pseudomonas fluorescens* named FZL4. C) Shine-Dalgarno sequence from the hcnA mRNA of *Pseudomonas fluorescens* named RP1. D) 22 nt artificial stem-loop TASL1. E) 26 nt artificial stem-loop TASL2. F) 30 nt artificial stem-loop TASL3 containing an A-A mismatch.

Supplementary Table 1. Reference ^{13}C values of the 26 nt stem-loop TASL2 under different sample conditions. Spectra were recorded at different temperatures ranging from 10 ° to 40 °C, at different pH values ranging from 5 to 8, at different NaCl concentrations ranging from 0 to 200 mM and in the presence of potassium and phosphate. Values in red indicate the largest deviations.

TASL2	Spectrometer	Temperature [°C]	C8 ^{13}C chemical shift of the 5′G [ppm]	C8 ^{13}C chemical shift of the 5′GG [ppm]	C5 ^{13}C chemical shift of the 3′C [ppm]	C1′ ^{13}C chemical shift of the 3′C [ppm]	C3′ ^{13}C chemical shift of the 3′C [ppm]
low salt 1.3mM pH8.0	600	10	139.2	136.8	98.0	92.9	69.6
low salt 1.3mM pH8.0	600	20	139.2	136.9	98.1	92.9	69.8
low salt 1.3mM pH8.0	600	30	139.1	136.9	98.1	92.9	69.9
low salt 2.7mM pH8.0	900	30	139.2	136.9	98.1	92.9	69.8
low salt 1.3mM pH8.0	600	40	139.2	137.0	98.2	92.9	70.0
low salt 1.3mM pH7.0	600	30	139.2	136.9	98.1	92.9	69.9
low salt 1.3mM pH6.0	600	30	139.1	136.9	98.1	92.9	69.9
low salt 1.3mM pH5.0	600	30	139.1	136.9	98.1	92.9	69.8
50mM NaCl 1.3mM pH6.5	600	30	139.2	136.9	98.1	92.9	69.8
100mM NaCl 1.3mM pH6.5	600	30	139.2	136.8	98.1	92.9	69.8
200mM NaCl 1.3mM pH6.5	600	30	139.3	136.8	98.1	93.0	69.8
50mM KH2PO4 1.3mM pH 6.0	600	30	139.1	136.9	98.1	92.9	69.8

Supplementary Table 2. List of all available ^{13}C chemical shift datasets: 58 from BMRB entries (excluding complexes and two pseudoknots), 6 from RNAs obtained from shifts published only in the literature and 6 RNAs of the Allain lab (unpublished). The chemical shifts are color coded: the value was in the expected range (green), shifted by 2.7 ppm (blue) or outside those two ranges (red).

BMRB or PDB code /internal name	C8 ^{13}C chemical shift of the 5′G [ppm]	C8 ^{13}C chemical shift of the 5′GG [ppm]	C5 ^{13}C chemical shift of the 3′C [ppm]	C1′ ^{13}C chemical shift of the 3′C [ppm]	C3′ ^{13}C chemical shift of the 3′C [ppm]	Temperature [°C]	nts	PDB code of structures used to analyze the sugar conformation	Reference
4120	part of tertiary structure (142.1)	part of tertiary structure (141.4)	3′A not 3′C	3′A not 3′C (89.5)	3′A not 3′C (72.0)	30	44		(Kolk et al. 1998)
4226	139.4	not 5′GG	97.8	93.1	69.9	25	30	1LDZ	(Legault et al. 1998)
4253	ssRNA 5′G (139.4)	not 5′GG but GC	ssRNA 3′C (94.35)	ssRNA 3′C (89.39)	ssRNA 3′C (66.96)	25	36		(Holland et al. 1999)
4346	139.4	137.6	98.7	93.1	70.2	15	33		(Mao et al. 1999)
4780[a]	107.8	136.7	98.2	92.9	69.7	25	19	1ESY	(Amarasinghe et al. 2000)
4816	not 5′G/ missing	not 5′GG/ missing	3′U not 3′C / missing	3′U not 3′C / (89.8)	3′U not 3′C / missing	5	31		(Bae et al. 2001)
5007[b]	136.3	before GU wobble (134.3)	missing	91.3	missing	1	30		(Flinders and Dieckmann 2001)
5170[c]	139.3	137.4	99.5	93.3	73.0	7	28		(DeJong et al. 2002)
5256 list 1	139.1	136.6	97.9	92.9	69.8	25	17	1KKA	(Cabello-Villegas et al. 2002)
5259 list 1	missing	136.8	98.0	93.0	69.6	25	17		(Cabello-Villegas et al. 2002)
5278	ssRNA 5′G / missing	ssRNA 5′GA	ssRNA 3′A	ssRNA 3′A / missing	ssRNA 3′A / missing	10	33		(Nixon et al. 2002)
5371	138.7	missing	97.6	missing	missing	30	24	1LC6	(Huppler et al. 2002)
5553	not 5′G	not 5′GG	missing	missing	missing	21	31		(Lee et al. 2003)
5559	ssRNA 5′G / missing	not 5′GG but GA/ missing	3′G	3′G (90.9)	3′G (71.0)	30	20		(Comolli et al. 2002)
5632[d]	136.4	134.2	ssRNA 3′C (95.3)	ssRNA 3′C (90.28)	ssRNA 3′C (73.2)	5	30		(Theimer et al. 2003b)
5655	139.2	137.1	98.3	93.1	70.2	30	24	1NC0	(Sashital et al. 2003)
5703[e]	138.1	136.3	97.5	92.0	missing	20	24		(Reiter et al. 2003)
5705	139.0	137.0	98.4	93.0	69.7	25	14	2KOC	(Furtig et al. 2004)
5773[f]	136.9	135.1	92.1	86.7	59.8	35	36		(Lawrence et al. 2003)
5834[g]	139.0	137.0	98.1	92.7	missing	10	22	1PJY	(Staple and Butcher 2003)
5852	138.8	not 5′GG	97.8	92.7	70.0	25	23	1OW9	(Hoffmann et al. 2003)
5919[h]	139.4	137.0	98.4	92.4	missing	25	42		(Vallurupalli and Moore 2003)
5932[i]	139.1	136.8	98.3	90.6	67.1	10	15		(Theimer et al. 2003a)
5962	139.1	137.0	97.5	missing	missing	30	36	1R2P	(Sigel et al. 2004)
6062[j]	139.2	136.9	98.0	92.9	99.6	25	23	1S34	(Cabello-Villegas et al. 2004)
6076 list 1	139.2	137.0	98.3	93.2	70.1	25	34	1R7W	(Du et al. 2004)
6077 list 1	139.2	137.1	98.1	missing	69.9	25	34	1R7Z	(Du et al. 2004)
6094	ssRNA 5′G (138.1)	ssRNA 5′GG (136.5)	ssRNA 3′U /	ssRNA 3′U / (89.7)	ssRNA 3′U / (68.8)	35	101		(D'Souza et al. 2004)
6239[k]	133.8	before GU wobble (131.5)	94.5	89.9	67.6	20	22		(Flinders and Dieckmann 2004)
6320[l]	missing	not 5′GG	missing	90.4	68.0	10	32		(Sashital et al. 2004)
6477	part of tertiary structure (139.0)	part of tertiary structure (136.4)	part of tertiary structure 3′A not 3′C	part of tertiary structure 3′A not 3′C (90.6)	part of tertiary structure 3′A not 3′C (71.2)	10/20	47		(Theimer et al. 2005)
6485[m]	135.3	133.0	96.5	92.6	69.6	30	27		(Stefl and Allain 2005)
Corrected 6485	139.3	137.0	98.2	93.0	70.0	30	27	1YSV	(Stefl and Allain 2005)

6509	ssRNA 5′A not 5′G (139.9)	ssRNA not 5′GG	ssRNA 3′A not 3′C	ssRNA 3′A not 3′C	ssRNA 3′A not 3′C	10	31		(Cornish et al. 2005)
6543	139.3	136.79	97.9	missing	missing	30	45	1Z2J	(Staple and Butcher 2005)
6562 list 1	before GU wobble (138.0)	part of GU wobble (135.7)	after GU wobble (97.4)	after GU wobble (90.1)	after GU wobble (76.5)	25	18		(Ihle et al. 2005)
6562 list 2	before GU wobble missing	part of GU wobble missing	after GU wobble (97.3)	after GU wobble missing	after GU wobble (76.1)	10	18		(Ihle et al. 2005)
6633[n]	137.1	135.0	95.9	90.7	67.4	10	41		(Gaudin et al. 2005)
6652 (dimer)	GU wobble/ missing	after GU wobble (135.9)	U not C at 3′ end/ (102.8)	U not C at 3′ end/ (91.3)	U not C at 3′ end/ missing	30	43		(Davis et al. 2005)
6756[o]	136.7	134.7	97.3	89.0	73.8	30	27		(Erat et al. 2007)
7090[p]	missing	not 5′GG	missing	92.6	73.5	25	18		(Joli et al. 2006)
7098	139.3	not 5′GG	98.1	92.7	70.0	30	35	2GM0	(Ulyanov et al. 2006)
7403	136.7	134.5	95.5	90.4	67.1	10	24	2QH2	(Theimer et al. 2007)
7404	136.5	134.4	missing	90.2	missing	10	23	2QH3	(Theimer et al. 2007)
7405	136.7	134.5	95.8	90.5	67.1	10	18	2QH4	(Theimer et al. 2007)
15080	139.2	137.2	98.7	93.0	70.0	30	20	2O33	(Sashital et al. 2007)
15081	Six modified bases, also at 5′-end								(Sashital et al. 2007)
15417	139.4	137.3	missing	92.9	69.7	4	34	2JTP	(Marcheschi et al. 2007)
15538[q]	139.2	not 5′GG	95.8	92.2	70.2	25	29		(Reiter et al. 2008)
15571	138.8	not 5′GG	after G (97.4)	92.7	69.7	25	10+10	2XJQ	(Popenda et al. 2008)
15572	138.8	not 5′GG	after G (97.5)	92.8	69.5	25	11+10	2JXS	(Popenda et al. 2008)
15656 list2[r]	139.0	136.8	99.96	92.86	69.8	15	29	2K5Z	(Ampt et al. 2008)
15745	before GU wobble (137.6)	part of GU wobble (135.7)	after GU wobble (97.3)	after GU wobble (89.9)	after GU wobble (76.3)	20	22		(Schwalbe et al. 2008)
15780	missing	not 5′GG	97.7	92.7	69.9	25	9+8	2K3Z	(Popenda et al. 2009)
15781	139.0	not 5′GG	97.7	92.7	69.8	25	9+8	2K41	(Popenda et al. 2009)
15786 list1[s]	missing	136.2	97.9	missing	missing	15	37		(van der Werf et al. 2008)
15786 list2	missing	missing	missing	92.8	70.0	25	37		(van der Werf et al. 2008)
15856	136.2	137.9	95.3	90.0	missing	20	29		unpublished
15858	before GU wobble (137.41)	part of GU wobble (134.17)	after GU wobble (94.7)	after GU wobble (90.2)	missing	15	7+7		unpublished
15859	136.31	136.72	95.3	90.4	missing	20	22		unpublished
15869	136.5	134.3	missing	90.4	67.1	25	30		(Carlomagno et al. 2008)
17RA	139.3	136.93	98.3	93.2	69.8	25	21	17RA	(Smith and Nikonowicz 1998)
1AFX[t]	143.4	145.3	99.1	92.8	70	20	12	1AFX	(Butcher et al. 1997)
1RNG	before GU wobble (140.6)	part of GU wobble (138.6)	after GU wobble (98.0)	after GU wobble (92.2)	after GU wobble (70.1)	20	12		(Jucker and Pardi 1995)
1SCL	136.4	134.2	95.5	90.3	missing	30	29	1SCL	(Szewczak and Moore 1995)
1UUU[u]	139.0	136.9	98.1	92.8	69.6	28	19	1UUU	(Sich et al. 1997)
1YFV[v]	138.2	135.2	97.2	91.9	68.7	35	8+8		(SantaLucia and Turner 1993)
20 nt hcnA	138.9	136.8	97.9	92.8	69.7	30	20		unpublished
TASL1	139.2	136.9	98.0	92.8	69.8	30	22		unpublished
TASL2	139.2	136.9	98.1	92.9	69.8	30	26		unpublished
TASL3	139.1	136.9	98.0	92.8	69.9	30	30		unpublished
FZL2	139.2	before GU wobble (137.1)	98.2	92.8	69.8	30	22		unpublished
FZL4	139.1	136.8	98.1	92.8	69.8	30	21		unpublished

a) One reference value is shifted by more than 30 ppm. No other errors were detected.

b) C8 and C1′ chemical shifts are systematically shifted.

c) C5 chemical shifts systematically too high. C3′ could be wrongly assigned or it is not the 3′end.

d) The C8 reference values are shifted by 2.7 ppm. C5 values seem to be less shifted and C1′ values are in the expected ranges.

e) Recalibration by 0.7 ppm, shifts all reference values in the expected regions.

f) C8 shifts seem to be ~ -2ppm off, C3′ by ~ -10 ppm, C1′ and C5′ by ~ 6 ppm.

g) C6 shifts of uridines and C8 shifts of adenosines are shifted by around 2.7 ppm whereas C8 chemical shifts of guanosines or C6 chemical shifts of uridines are in the expected regions. Since two labeled samples were used for recording the NMR data, an AU and a GC 13C/15N-labeled sample, we suspect that the spectrum of the AU labeled RNA was not calibrated properly. It is likely that as well the C2 shifts of adenosines are not properly referenced. For our database we removed these shifts.

h) C1′ reference value only slightly out of the expected range. Other C1′ values are as well only slightly low.

i) C1′ and C3′ chemical shifts are off by ~ 2.7 ppm.

j) One reference value is shifted by more than 30 ppm. No other errors were detected.

k) C6 and C8 are ~ -5ppm off, sugars ~ -2 ppm, C5 3 ppm.

l) Sugars reference values are off by 2–3 ppm. Also the other values seem to be off by 2–3 ppm.

m) By reprocessing and analyzing the original data it turned out that the ^{13}C assignments of the bases (including C5) were derived from a ^{13}C-HSQC spectrum optimized for the aromatic region (offset 145 ppm, spectral width 30 ppm) which was mis-calibrated by 4.0 ppm. The ribose ^{13}C assignments originated from a 3D ^{13}C-edited NOESY which was mis-calibrated by 0.4 ppm. The ^{13}C C5 assignments were mis-calibrated by 1.7 ppm. Interestingly after calibrating the chemical shifts to DSS, all the reference values lie in the very narrow ranges that we received by measuring the RNA stem-loop TASL2 under different conditions.

n) Chemical shifts shifted by ~ 2.2 ppm.

o) C8 resonances are shifted by ~ 2.7 ppm. No clear systematic offsets detected for the other atom types.

p) Origin off the unexpected C3′ reference value not clear. C5, C6 and C8 resonances are not assigned. Therefore we cannot validate the reliability of the data.

q) All C5 shifts are systematically shifted by around 2.7 ppm. Here we assume improper calibration of the ^{13}C spectrum of the C5 chemical shifts. The C1′ reference value is only 0.3 ppm out of the expected range and other C1′ seem as well to have only a slight tendency towards too small values.

r) All C5 shifts are shifted by around 2.0 ppm (Fig. 5).

s) One reference value slightly too low.

t) Sugar shifts seem to be correct. The origin of the inconsistency of the other reference values is unclear.

u) ^{13}C chemical shifts were originally referenced to external 1 % dioxane at 67.8 ppm. Since 1 % dioxane was reported to resonate at 69.3 ppm in regard to DSS as reference, we shifted the values by the difference of 1.5 ppm.

v) Resonances shifted by ~ 1 ppm.

References

Amarasinghe GK, De Guzman RN, Turner RB and Summers MF (2000) NMR structure of stem-loop SL2 of the HIV-1 psi RNA packaging signal reveals a novel A-U-A base-triple platform. J Mol Biol 299: 145-156.

Ampt KA, Ottink OM, Girard FC, Nelissen F, Tessari M and Wijmenga SS (2008) 1H, 13C and 15N NMR assignments of Duck HBV apical stem loop of the epsilon encapsidation signal. Biomol NMR Assign 2: 159-162.

Bae SH, Cheong HK, Lee JH, Cheong C, Kainosho M and Choi BS (2001) Structural features of an influenza virus promoter and their implications for viral RNA synthesis. Proc Natl Acad Sci U S A 98: 10602-10607.

Butcher SE, Dieckmann T and Feigon J (1997) Solution structure of the conserved 16 S-like ribosomal RNA UGAA tetraloop. J Mol Biol 268: 348-358.

Cabello-Villegas J, Giles KE, Soto AM, Yu P, Mougin A, Beemon KL and Wang YX (2004) Solution structure of the pseudo-5' splice site of a retroviral splicing suppressor. RNA 10: 1388-1398.

Cabello-Villegas J, Winkler ME and Nikonowicz EP (2002) Solution conformations of unmodified and A(37)N(6)-dimethylallyl modified anticodon stem-loops of Escherichia coli tRNA(Phe). J Mol Biol 319: 1015-1034.

Carlomagno T, Amata I, Williamson JR and Hennig M (2008) NMR assignments of HIV-2 TAR RNA. Biomol NMR Assign 2: 167-169.

Comolli LR, Ulyanov NB, Soto AM, Marky LA, James TL and Gmeiner WH (2002) NMR structure of the 3' stem-loop from human U4 snRNA. Nucleic Acids Res 30: 4371-4379.

Cornish PV, Hennig M and Giedroc DP (2005) A loop 2 cytidine-stem 1 minor groove interaction as a positive determinant for pseudoknot-stimulated -1 ribosomal frameshifting. Proc Natl Acad Sci U S A 102: 12694-12699.

D'Souza V, Dey A, Habib D and Summers MF (2004) NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus. J Mol Biol 337: 427-442.

Davis JH, Tonelli M, Scott LG, Jaeger L, Williamson JR and Butcher SE (2005) RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex. J Mol Biol 351: 371-382.

DeJong ES, Marzluff WF and Nikonowicz EP (2002) NMR structure and dynamics of the RNA-binding site for the histone mRNA stem-loop binding protein. RNA 8: 83-96.

Du Z, Ulyanov NB, Yu J, Andino R and James TL (2004) NMR structures of loop B RNAs from the stem-loop IV domain of the enterovirus internal ribosome entry site: a single C to U substitution drastically changes the shape and flexibility of RNA. Biochemistry 43: 5757-5771.

Erat MC, Zerbe O, Fox T and Sigel RK (2007) Solution structure of domain 6 from a self-splicing group II intron ribozyme: a Mg(2+) binding site is located close to the stacked branch adenosine. ChemBioChem 8: 306-314.

Flinders J and Dieckmann T (2001) A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA. J Mol Biol 308: 665-679.

Flinders J and Dieckmann T (2004) The solution structure of the VS ribozyme active site loop reveals a dynamic "hot-spot". J Mol Biol 341: 935-949.

Furtig B, Richter C, Bermel W and Schwalbe H (2004) New NMR experiments for RNA nucleobase resonance assignment and chemical shift analysis of an RNA UUCG tetraloop. J Biomol NMR 28: 69-79.

Gaudin C, Mazauric MH, Traikia M, Guittet E, Yoshizawa S and Fourmy D (2005) Structure of the RNA signal essential for translational frameshifting in HIV-1. J Mol Biol 349: 1024-1035.

Hoffmann B, Mitchell GT, Gendron P, Major F, Andersen AA, Collins RA and Legault P (2003) NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. Proc Natl Acad Sci U S A 100: 7003-7008.

Holland JA, Hansen MR, Du Z and Hoffman DW (1999) An examination of coaxial stacking of helical stems in a pseudoknot motif: the gene 32 messenger RNA pseudoknot of bacteriophage T2. RNA 5: 257-271.

Huppler A, Nikstad LJ, Allmann AM, Brow DA and Butcher SE (2002) Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. Nat Struct Biol 9: 431-435.

Ihle Y, Ohlenschlager O, Hafner S, Duchardt E, Zacharias M, Seitz S, Zell R, Ramachandran R and Gorlach M (2005) A novel cGUUAg tetraloop structure with a conserved yYNMGg-type backbone conformation from cloverleaf 1 of bovine enterovirus 1 RNA. Nucleic Acids Res 33: 2003-2011.

Joli F, Bouchemal N, Laigle A, Hartmann B and Hantz E (2006) Solution structure of a purine rich hexaloop hairpin belonging to PGY/MDR1 mRNA and targeted by antisense oligonucleotides. Nucleic Acids Res 34: 5740-5751.

Jucker FM and Pardi A (1995) Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif. Biochemistry 34: 14416-14427.

Kolk MH, van der Graaf M, Wijmenga SS, Pleij CW, Heus HA and Hilbers CW (1998) NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA. Science 280: 434-438.

Lawrence DC, Stover CC, Noznitsky J, Wu Z and Summers MF (2003) Structure of the intact stem and bulge of HIV-1 Psi-RNA stem-loop SL1. J Mol Biol 326: 529-542.

Lee MK, Bae SH, Park CJ, Cheong HK, Cheong C and Choi BS (2003) A single-nucleotide natural variation (U4 to C4) in an influenza A virus promoter exhibits a large structural change: implications for differential viral RNA synthesis by RNA-dependent RNA polymerase. Nucleic Acids Res 31: 1216-1223.

Legault P, Hoogstraten CG, Metlitzky E and Pardi A (1998) Order, dynamics and metal-binding in the lead-dependent ribozyme. J Mol Biol 284: 325-335.

Mao H, White SA and Williamson JR (1999) A novel loop-loop recognition motif in the yeast ribosomal protein L30 autoregulatory RNA complex. Nat Struct Biol 6: 1139-1147.

Marcheschi RJ, Staple DW and Butcher SE (2007) Programmed ribosomal frameshifting in SIV is induced by a highly structured RNA stem-loop. J Mol Biol 373: 652-663.

Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE and Wuthrich K (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy. J Biomol NMR 12: 1-23.

Nixon PL, Rangan A, Kim YG, Rich A, Hoffman DW, Hennig M and Giedroc DP (2002) Solution structure of a luteoviral P1-P2 frameshifting mRNA pseudoknot. J Mol Biol 322: 621-633.

Popenda L, Adamiak RW and Gdaniec Z (2008) Bulged adenosine influence on the RNA duplex conformation in solution. Biochemistry 47: 5059-5067.

Popenda L, Bielecki L, Gdaniec Z and Adamiak RW (2009) Structure and dynamics of adenosine bulged RNA duplex reveals formation of the dinucleotide platform in the C:G-A triple. Arkivoc130-144.

Price RP, Oubridge C, Varani G and Nagai K. (1998) Preparation of RNA:protein complexes for X-ray cristallography and NMR. In Smith, CWJ (ed.), *RNA-Protein interactions: A practical approach*. Oxford University Press, pp. 37-74.

Reiter NJ, Maher LJ, 3rd and Butcher SE (2008) DNA mimicry by a high-affinity anti-NF-kappaB RNA aptamer. Nucleic Acids Res 36: 1227-1236.

Reiter NJ, Nikstad LJ, Allmann AM, Johnson RJ and Butcher SE (2003) Structure of the U6 RNA intramolecular stem-loop harboring an S(P)-phosphorothioate modification. RNA 9: 533-542.

SantaLucia J, Jr. and Turner DH (1993) Structure of (rGGCGAGCC)2 in solution from NMR and restrained molecular dynamics. Biochemistry 32: 12612-12623.

Sashital DG, Allmann AM, Van Doren SR and Butcher SE (2003) Structural basis for a lethal mutation in U6 RNA. Biochemistry 42: 1470-1477.

Sashital DG, Cornilescu G, McManus CJ, Brow DA and Butcher SE (2004) U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. Nat Struct Mol Biol 11: 1237-1242.

Sashital DG, Venditti V, Angers CG, Cornilescu G and Butcher SE (2007) Structure and thermodynamics of a conserved U2 snRNA domain from yeast and human. RNA 13: 328-338.

Sawadogo M and Van Dyke MW (1991) A rapid method for the purification of deprotected oligodeoxynucleotides. Nucleic Acids Res 19: 674.

Schwalbe M, Ohlenschlager O, Marchanka A, Ramachandran R, Hafner S, Heise T and Gorlach M (2008) Solution structure of stem-loop alpha of the hepatitis B virus post-transcriptional regulatory element. Nucleic Acids Res 36: 1681-1689.

Sich C, Ohlenschlager O, Ramachandran R, Gorlach M and Brown LR (1997) Structure of an RNA hairpin loop with a 5'-CGUUUCG-3' loop motif by heteronuclear NMR spectroscopy and distance geometry. Biochemistry 36: 13989-14002.

Sigel RK, Sashital DG, Abramovitz DL, Palmer AG, Butcher SE and Pyle AM (2004) Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif. Nat Struct Mol Biol 11: 187-192.

Smith JS and Nikonowicz EP (1998) NMR structure and dynamics of an RNA motif common to the spliceosome branch-point helix and the RNA-binding site for phage GA coat protein. Biochemistry 37: 13486-13498.

Staple DW and Butcher SE (2003) Solution structure of the HIV-1 frameshift inducing stem-loop RNA. Nucleic Acids Res 31: 4326-4331.

Staple DW and Butcher SE (2005) Solution structure and thermodynamic investigation of the HIV-1 frameshift inducing element. J Mol Biol 349: 1011-1023.

Stefl R and Allain FH (2005) A novel RNA pentaloop fold involved in targeting ADAR2. RNA 11: 592-597.

Szewczak AA and Moore PB (1995) The sarcin/ricin loop, a modular RNA. J Mol Biol 247: 81-98.

Theimer CA, Blois CA and Feigon J (2005) Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. Mol Cell 17: 671-682.

Theimer CA, Finger LD and Feigon J (2003a) YNMG tetraloop formation by a dyskeratosis congenita mutation in human telomerase RNA. RNA 9: 1446-1455.

Theimer CA, Finger LD, Trantirek L and Feigon J (2003b) Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. Proc Natl Acad Sci U S A 100: 449-454.

Theimer CA, Jady BE, Chim N, Richard P, Breece KE, Kiss T and Feigon J (2007) Structural and functional characterization of human telomerase RNA processing and cajal body localization signals. Mol Cell 27: 869-881.

Ulyanov NB, Mujeeb A, Du Z, Tonelli M, Parslow TG and James TL (2006) NMR structure of the full-length linear dimer of stem-loop-1 RNA in the HIV-1 dimer initiation site. J Biol Chem 281: 16168-16177.

Vallurupalli P and Moore PB (2003) The solution structure of the loop E region of the 5S rRNA from spinach chloroplasts. J Mol Biol 325: 843-856.

van der Werf RM, Girard FC, Nelissen F, Tessari M and Wijmenga SS (2008) 1H, 13C and 15N NMR assignments of Duck HBV primer loop of the encapsidation signal epsilon. Biomol NMR Assign 2: 143-145.

Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL and Sykes BD (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J Biomol NMR 6: 135-140.