J Biomol NMR (2013) 57:1–9
DOI 10.1007/s10858-013-9771-5

COMMUNICATION

4D Non-uniformly sampled C,C-NOESY experiment for sequential assignment of ^{13}C,^{15}N-labeled RNAs

Jan Stanek · Peter Podbevšek · Wiktor Koźmiński · Janez Plavec · Mirko Cevec

Received: 26 April 2013 / Accepted: 10 August 2013 / Published online: 21 August 2013

Abstract A 4D ^{13}C(aromatic),^{13}C(ribose)-edited NOESY experiment is introduced to improve sequential assignment of non-coding RNA, often hampered by a limited dispersion of ^{1}H and ^{13}C chemical shifts. The ^{13}C-labeling of RNA is fully utilized by inclusion of two ^{13}C evolution periods. These dimensions provide enhanced dispersion of resonances in the 4D spectrum. High spectral resolution is obtained using random non-uniform sampling in three indirect dimensions. The autocorrelation peaks are efficiently suppressed using band-selective pulses. Since the dynamic range of observed resonances is significantly decreased, the reconstruction of the 4D spectrum is greatly simplified. The experiment can replace two conventionally sampled 3D NOESY spectra (either ribose-^{13}C- or aromatic-^{13}C-separated), and remove most ambiguities encountered during sequential walks. The assignment strategy based on a homonuclear and 4D C,C-edited NOESY experiments is proposed and verified on a 34-nt RNA showing typical structure elements.

Keywords Multidimensional NMR · Non-uniform sampling · Isotope labeled RNA · Resonance assignment

RNA is no longer considered simply a carrier of genetic information from the DNA to the ribosomes. Recently, several new classes of non-coding RNAs were introduced which stimulated the need for a structural and mechanistic insight into their function (Mercer et al. 2009; Esteller 2011). The structure–function relationship of RNAs can be efficiently explored with NMR techniques (Varani et al. 1996; Wijmenga and van Buuren 1998; Fürtig et al. 2003). However, deriving a high resolution structure of an RNA molecule of even moderate size can be problematic. The problem of the lower proton density of RNA can be solved by isotopic labeling techniques, which have become readily available in the recent years (Lu et al. 2010). As a result, one often deals with uniformly ^{13}C and/or ^{15}N labeled RNA oligonucleotides. RNA is also notorious for its poor chemical shift dispersion, which is to a large extent caused by a small number of available building blocks. Additionally, double stranded helical regions feature little variation in chemical environment. In the case of non-coding RNAs large portions of non-base paired nucleotides are not uncommon. The lack of base stacking of such regions results in even less chemical shift dispersion. Similarly to NMR studies on intrinsically disordered proteins, novel techniques are needed to assign poorly resolved resonances.

Sequential assignment of RNA relies either on through-bond HCP type of experiments or on NOE-based

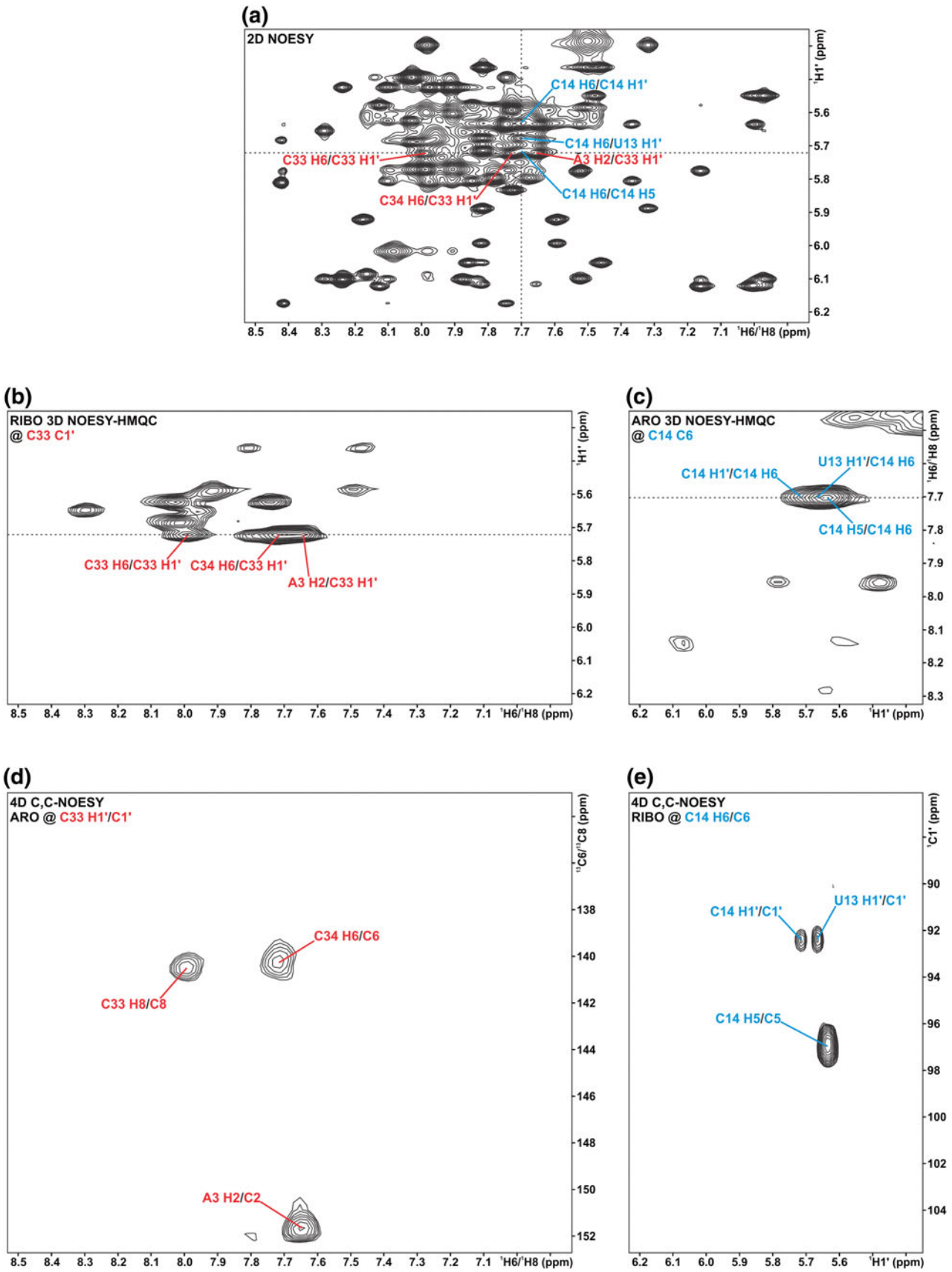
Electronic supplementary material The online version of this article (doi:10.1007/s10858-013-9771-5) contains supplementary material, which is available to authorized users.

J. Stanek · W. Koźmiński
Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw, Żwirki i Wigury 101, 02-089 Warszawa, Poland

P. Podbevšek · J. Plavec · M. Cevec (✉)
Slovenian NMR Centre, National Institute of Chemistry, Hajdrihova ulica 19, 1000 Ljubljana, Slovenia
e-mail: mirko.cevec@ki.si

J. Plavec
EN-FIST Centre of Excellence, Dunajska cesta 156, 1000 Ljubljana, Slovenia

J. Plavec
Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva cesta 5, 1000 Ljubljana, Slovenia


(a)
2D NOESY
C14 H6/C14 H1'
C14 H6/U13 H1'
C33 H6/C33 H1'
A3 H2/C33 H1'
C14 H6/C14 H5
C34 H6/C33 H1'
(b)
RIBO 3D NOESY-HMQC
@ C33 C1'
C33 H6/C33 H1'
C34 H6/C33 H1'
A3 H2/C33 H1'
(c)
ARO 3D NOESY-HMQC
@ C14 C6
U13 H1'/C14 H6
C14 H1'/C14 H6
C14 H5/C14 H6
(d)
4D C,C-NOESY
ARO @ C33 H1'/C1'
C34 H6/C6
C33 H8/C8
A3 H2/C2
(e)
4D C,C-NOESY
RIBO @ C14 H6/C6
U13 H1'/C1'
C14 H1'/C1'
C14 H5/C5

◀ **Fig. 1** Resolution enhancement provided by heteronuclear-edited spectra. Compared are homonuclear **a** 2D NOESY, **b** 3D NOESY-^{13}C(ribose)HMQC, **c** 3D NOESY-^{13}C(aromatic)HMQC, **d**, **e** 4D ^{13}C(aromatic),^{13}C(ribose)-edited HMQC-NOESY-HSQC: **d** ARO: aromatic plane, **e** RIBO: ribose plane. Addition of ribose-^{13}C dimension **b** greatly reduces spectral crowding present in 2D spectrum **a**, however, limited spectral resolution in the indirectly detected ^{1}H (aromatic) dimension results in overlap of A3 H2–C33 H1′ and C34 H8–C33 H1′ cross-peaks **b**. These peaks are resolved in the aromatic plane of the 4D spectrum **d** owing to signal dispersion in ^{13}C (aromatic) dimension. Similarly, aromatic 3D NOESY-HSQC **c** helps to resolve cross-peaks in 2D NOESY spectrum (see spectral crowding along *vertical dotted line*). Nevertheless, an essential U13 H1′–C14 H6 peak is overlapped with a very intense C14 H5–C14 H6 peak. Both peaks can be clearly identified in the 4D spectrum **e** as C1′ and C5 chemical shift regions are disjoint. Experimental parameters of the 3D and 4D spectra are given in the caption to Fig. 2 and Supplementary Fig. S2

techniques (Marino et al. 1994). The so-called "sequential walk" can be observed in the aromatic-anomeric region of 2D NOESY spectra. This region contains intra- and inter-nucleotide cross-peaks between anomeric H1′ and aromatic H6/H8 protons (see Fig. 1a). However, spectral crowding is increased by the presence of potentially useful inter- and intra-nucleotide H1′–H2 cross-peaks (to adenine H2) and less relevant intra-nucleotide H5–H6 cross-peaks of pyrimidine residues. When dealing with ^{13}C-labeled oligonucleotides 3D ^{13}C-edited NOESY experiments are used for the assignment of aromatic and anomeric ^{1}H resonances and attached ^{13}C spins (Nikonowicz and Pardi 1993). Two 3D spectra are usually acquired. The NOE transfers ^{1}H magnetization to either anomeric or aromatic protons and, as a result, the ^{13}C dimension of a 3D spectrum provides C1′ or C6/C8 chemical shifts, respectively. A single 3D spectrum, either ^{13}C(aro) or ^{13}C(ribo)-edited, allows to resolve H1′–H6/H8 peaks from *one* kind of the above mentioned extra peaks (H1′–H2 or H5–H6), which is possible due to separated ^{13}C regions of C1′ (87–94 ppm) and C5 (94–107 ppm) resonances in the case of ^{13}C(ribo)-edited 3D spectrum; similarly aromatic C6/C8 (131–145 ppm) and C2 (145–155 ppm) spins can be distinguished in 3D ^{13}C(aro)-edited NOESY. However, as illustrated in Fig. 1b, inter-nucleotide H1′–H2 and intra-nucleotide H1′–H6 peaks may still overlap and hamper the sequential assignment based on the ^{13}C(ribo)-edited 3D spectrum. Additionally, in aromatic 3D NOESY-HSQC (Fig. 1c), very intense pyrimidine H5–H6 cross-peaks often overlap with the less intense H1′–H6/H8 cross-peaks, which are essential for sequential assignment.

For this reason we have developed a 4D C,C-edited NOESY experiment, which incorporates both aromatic and ribose ^{13}C and ^{1}H evolution periods in the single pulse sequence. The advantage of higher dimensionality is demonstrated by a pair-wise comparison of two 3D spectra and corresponding regions of a 4D spectrum. The introduction of an additional ribose ^{13}C dimension (ω_3) enables us to resolve crucial H1′–H8 and H1′–H6 cross-peaks from usually more intense H5–H6 cross-peaks on the basis of separation of C5 and C1′ regions. For example, U13 H1′–C14 H6 and C14 H5–H6 peaks, which overlap in 3D spectrum, are resolved in a corresponding plane of 4D spectrum shown in Fig. 1e. Otherwise, the degeneracy of H1′ and H5 chemical shifts limits the use of aromatic 3D NOESY-HSQC for sequential assignment and requires acquisition of complementary ribose 3D NOESY-HSQC. In this communication we propose to replace this rather non-optimal solution by the acquisition of a single band-selective 4D C,C-edited NOESY experiment. The extra ^{13}C dimension reduces spectral overlap, which is especially pronounced for large RNAs or RNAs with very poor chemical shift dispersion. Additionally, reduced ambiguity of cross-peak assignment facilitates more reliable sequential assignment which is often severely hampered by the multiplicity of peaks matching a single ^{1}H chemical shift. We have tested the performance of 4D C,C-edited NOESY experiment on a fully ^{13}C,^{15}N-labeled 34-nt hairpin RNA (1.5 mM in D_2O solution) consisting of two A-RNA form stems, one adenine bulge, an asymmetric internal loop and a GAAA terminal loop (Cevec et al. 2010).

The pulse scheme for the 4D ^{13}C(aro)HMQC-NOESY-^{13}C(ribo)HMQC experiment is shown in Fig. 2. The experiment was designed with particular emphasis on the selectivity of *both* aromatic and ribose HMQC parts. Effectively, auto-correlation peaks are absent in a detected signal. This is of the particular relevance since non-uniform sampling (NUS) was employed to obtain reasonable resolution in the indirect dimensions. The elimination of intense diagonal peaks results in a considerably decreased dynamic range of signal amplitudes, which is the main difficulty when using NUS for NOESY spectra (this issue is commented below in more detail).

We wish to emphasize that the presented approach for elimination of diagonal signals does not entail any loss of sensitivity except from that induced by non-uniform inversion (refocusing) profiles of shaped pulses. This is in contrast to other approaches proposed recently that employ subtraction of exclusively diagonal spectrum (Xu et al. 2005; Wen et al. 2012) or orthogonal spin-state selection (Diercks et al. 2010; Werner-Allen et al. 2010). The solution presented here resembles the spin type selection employed in 4D C,N-NOESY (Kay et al. 1990), 3D H(aro)-NOESY-CH_3, NH (Xia et al. 2001) and 4D ^{13}C(aliph),^{13}C(aro)-edited NOESY (Stanek et al. 2013) experiments for proteins in the sense that aromatic and ribose ^{1}H (or ^{13}C) spins are treated as heteronuclei.

HMQC was employed for the frequency labeling of both aromatic (C6 or C8) and ribose (C1′) ^{13}C spins since the

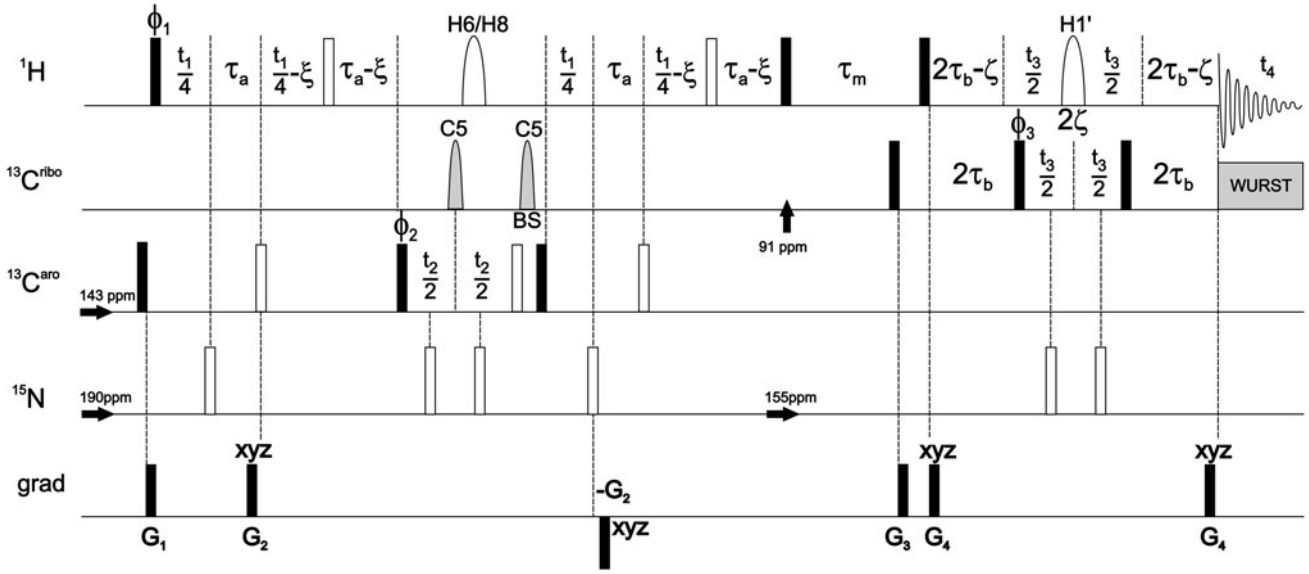


Fig. 2 Pulse scheme for the selective 4D C,C(aromatic→ribose)-edited HMQC-NOESY-HMQC experiment. *Solid and open bars* represent 'hard' 90° and 180° pulses, respectively. All pulses are applied along the *x*-axis of the rotating frame unless indicated otherwise. ^{1}H aromatic and ribose selective refocusing pulses, showed as open bell-shaped pulses (denoted H6/H8 and H1′, respectively), utilize *r-snob* profiles (Kupče et al. 1995) of the following duration, peak r.f. field and frequency modulation: 1.54 ms, 1.58 kHz, 3 ppm × $\gamma_H B_0$ (for H6/H8) and 2.2 ms, 1.12 kHz, 1 ppm × $\gamma_H B_0$ (for H1′). The grey bell-shaped ^{13}C pulse denoted as 'C5' (and a Bloch-Siegert phase compensation pulse 'BS') is introduced to refocus evolution of C5–C6 couplings in pyrimidine bases ($J_{C5-C6} \approx 68$ Hz) during frequency labeling in t_2. It employs *isnob-2* profile of length of 276 μs, peak r.f. field of 6.23 kHz and frequency modulation of −43 ppm $*\gamma_C B_0$ (effective ^{13}C offset of 100 ppm). ^{13}C adiabatic composite pulse decoupling was performed with WURST (Kupče and Freeman 1995). The ^{1}H carrier frequency is set on resonance with HDO signal. Initially, ^{13}C and ^{15}N carrier frequencies are set to 143 and 190 ppm, respectively, and shifted to 91 and 155 ppm for the ribose HQMC. NOESY mixing time of $\tau_m = 150$ ms was used. The following delays were set: $\tau_a = 1.25$ ms $\approx (4\ J_{CH}\ \text{aro})^{-1}$, $\tau_b = 1.56$ ms $\approx (4\ J_{CH}\ \text{ribo})^{-1}$, $\zeta = 1.1$ ms (1/2 duration of H1′ selective pulse). Additionally, in order to avoid sequence elongation by selective H6/H8 pulse, τ_a is decreased by $(\eta - \lambda)/4$ (if positive), where $\eta = 1.54$ ms is the duration of H6/H8 pulse and $\lambda = t_2 + 4\text{pw}_{90}(^{13}\text{C}) + 2\text{pw(C5 pulse)}$, or performed during t_2 (if $\eta - \lambda$ is negative). Note that for simplicity the H6/H8 pulse is centered with respect to entire period of transverse ^{13}C magnetization, and not w. r. t. t_2, which is allowed since no passive heteronuclear H–C coupling is active during t_2. Similarly, the duration of selective H1′ pulse (2ζ) is subtracted from coherence transfer delays ($2\tau_b$) in the ribose HMQC. This usually requires independent programming of ^{1}H and ^{13}C channels. The indirect evolution of aromatic ^{1}H (t_1) is performed in a semi-constant time fashion to utilize coherence transfer delays (2 × $2\tau_a$); assuming that $t_{1max} \geq 4\tau_a$ the contraction is $\xi = t_1/t_{1max} \times \tau_a$ (note that always $\xi \leq \tau_a$ and $\xi \leq t_1/4$). The evolution of ribose ^{13}C spins (t_3) should be started from half-dwell time (complementing delay of $(2\text{sw}_3)^{-1}$-$4\text{pw}_{90}(^{13}\text{C})$ is required) to avoid frequency-dependent phase in ω_3. Quadrature detection in t_1, t_2 and t_3 is accomplished by altering ϕ_1, ϕ_2 and ϕ_3, respectively, according to the States-TPPI procedure. Additionally, ϕ_1 is incremented by $-2\pi\Delta\Omega_H \cdot t_1$, $\Delta\Omega_H = 3.1$ ppm $*\gamma_H B_0$, to shift the effective center of ω_1 (^{1}H) dimension to ~7.8 ppm. The phase cycle employed is: $\phi_1 = x$; $\phi_2 = x, -x, x, -x$; $\phi_3 = 2(x), 2(-x)$; $\phi_{rec} = x, -x, -x, x$. Four transients were accumulated for each sampling point. Gradients were employed with the following durations and strengths: G_1 (0.5 ms, 3.55 G/cm), G_2 (0.5 ms, 5.32 G/cm), G_3 (2 ms, 16.0 G/cm), G_4 (0.5 ms, 30.2 G/cm). Gradients G_2 and G_4 are performed during delays τ_a and $2\tau_b - \zeta$, respectively. 4,100 sampling points (t_1, t_2, t_3) were randomly chosen from 30 × 32 × 32 Cartesian grid according to Gaussian probability distribution $p(t) = \exp[-(t/t_{max})^2/2\sigma^2]$; $\sigma = 0.5$. Maximum evolution times of 20 (t_{1max}), 6.4 (t_{2max}) and 7 ms (t_{3max}) were achieved in the indirectly detected dimensions. Acquisition time was set to 85 ms (t_{4max}). Spectral widths of 1.5 (ω_1), 5.0 (ω_2), 4.5 (ω_3) and 12 kHz (ω_4) were assumed. The total experiment duration was 75 h. The experiment was performed at 298 K on an Agilent 600 MHz spectrometer equipped with a room-temperature 5 mm triple-resonance probe. π/2 pulse durations were 7.9, 16.8 and 42 μs for ^{1}H, ^{13}C and ^{15}N, respectively. The interscan delay of 1.8 s for optimal recovery of ^{1}H magnetization (sensitivity per unit time) was used. The reversed experiment (ribose → aromatic) can be performed using the same scheme by swapping delays τ_a and τ_b, the selective ^{1}H pulses (denoted as H1′ and H6/H8) and reversing the order of carrier change

transverse relaxation is significantly attenuated compared to single-quantum coherences (Fiala et al. 1998, 2000; Brutscher and Simorre 2001; Marino et al. 1997). The crucial modification with respect to original double-C,C-edited NOESY (Vuister et al. 1993) is the incorporation of band-selective ^{1}H refocusing pulses. Since the H1′ and H6/H8 spins resonate in reasonably distinct spectral regions, selective pulse combined with a pair of pulsed field gradients (denoted G_2 and G_4 in Fig. 2 for aromatic and ribose HMQC, respectively) provide sufficient selectivity. An additional advantage of these ^{1}H selective pulses is the refocusing of H1′–H2′ and H6–H5 homonuclear scalar couplings. It is noteworthy that evolution of these couplings is suppressed not only in ^{13}C evolution periods (t_2 and t_3), but also in indirect ^{1}H evolution (t_1) and four coherence transfer delays (of total duration of $4\tau_a$ and $4\tau_b$) thanks to the symmetry of HMQC. It should be noted that despite significant length of the selective ^{1}H pulses the

overall duration of HMQC is not increased since they are partially performed during coherence transfer delays. However, the loss of ^{1}H resonances outside typical regions is inevitable. The deleterious effect of considerable two-bond ^{1}H–^{15}N, and one-bond ^{13}C–^{15}N couplings are eliminated in entire HMQC by the application of *two* ^{15}N inversion pulses during t_2. Additionally, band-selective off-resonance pulses were employed in the aromatic HMQC to prevent dephasing due to C5–C6 couplings in pyrimidine bases. In effect, the signal in t_1 and t_2 is only affected by relaxation and the non-uniform sampling can be applied to obtain optimal resolution. It should be also emphasized that the sensitivity is further increased owing to the utilization of two coherence transfer delays ($4\tau_a \approx (J_{CH}$ aro$)^{-1}$ in total) for the semi-constant time evolution of aromatic H6/H8 spins (t_1) as previously proposed for 4D C,N-edited NOESY for proteins (Xu et al. 2006; Stanek et al. 2012).

RNAs larger than studied here may also benefit from the use of transverse relaxation optimization (TROSY) for aromatic C–H coherences (Brutscher et al. 1998; Riek et al. 2001; Fiala et al. 2000; Brutscher and Simorre 2001). The corresponding pulse schemes are given in Supplementary Fig. S1. The aromatic-to-ribose NOESY (Fig. S1a) utilizes steady-state ^{13}C magnetization, therefore the selection of aromatic ^{13}C spins is performed using shaped ^{13}C refocusing pulse *after* ^{13}C evolution period (t_2). In contrast, the reversed (ribose-to-aromatic) experiment utilizes selective ^{13}C inversion pulse for selection of aromatic ^{13}C spins in the first INEPT (i.e. *before* ^{13}C evolution t_3) of the TROSY detection part (Fig. S1b). As before, for ribose coherences HMQC is employed as TROSY effect is not pronounced for the relaxation of C1′ (Fiala et al. 2000; Brutscher et al. 1998).

For the RNA samples in H_2O/D_2O solution residual solvent signal may heavily distort directly detected ribose H1′ resonances; in this case the reversed (ribose-to-aromatic) 4D NOESY experiment may appear beneficial (see Figs. 2 and S1b). However, the alternative experiment apparently shows approx. 35 % lower sensitivity which should be attributed to different relaxation rate and incomplete suppression of homo- and heteronuclear couplings during detection of aromatic ^{1}H resonances. Additionally, the reversed experiment with aromatic TROSY cannot utilize ^{13}C steady state magnetization which further reduces sensitivity (Fig. S1b). Another argument in favor of direct detection of ribose resonances is that the superior resolution in poorly dispersed ribose ^{1}H region is highly desired.

The principal advantage of the selective aromatic-to-ribose NOESY is the absence of auto-correlation peaks. These signals are main source of artefacts (artificial noise) in the reconstructed spectra and thus particularly deleterious in NUS NOESY experiments. The substantial reduction of dynamic range of signal amplitudes facilitates a faithful spectral reconstruction with basic or more advanced *four-dimensional* algorithms (Luan et al. 2005; Coggins and Zhou 2008; Mobli et al. 2010; Stanek et al. 2012; Hyberts et al. 2009). Here the simplest solution, namely zero-augmented Fourier Transformation was employed (Kazimierczuk et al. 2012) since thermal noise seems a limiting factor for spectral reconstruction at the specific experimental conditions. In the case of significantly higher experimental sensitivity (use of cryoprobe, stronger B_0 field, etc.), however, more sophisticated NUS reconstruction algorithms may prove beneficial. Direct FT also shows superior performance in processing of high resolution spectra. The transformation time of the 4D (non-sparse) spectrum with digital resolution of 128 in all indirect dimensions ω_1–ω_3 and 128 points in the directly detected dimension (5–6.2 ppm subregion corresponding to H1′ resonances) was only 1 min. The reasonable size of resulting 4D spectrum ($\sim$1 GB) allows for convenient analysis using common visualization tools, e.g. Sparky (Goddard and Kneller 2008). The processing software (*reconstructor4d*) can be downloaded at no charge from nmr.cent3.uw.edu.pl.

It is of a particular interest to compare the sensitivity and resolution of corresponding single ^{13}C-edited 3D and C,C-edited 4D spectra. For this purpose two uniformly sampled 3D spectra were recorded in the total experimental time corresponding to a single 4D spectrum. This is to reflect the fact that the latter basically provides more information than two 3D spectra, namely a direct correlation *of pairs* of aromatic (^{1}H, ^{13}C) and ribose (^{1}H, ^{13}C) chemical shifts. Test 3D ^{13}C(ribo)- and ^{13}C(aro)-edited NOESY experiments were performed at the same experimental conditions, with equal resolution in the indirect dimensions, and using the sequences shown in Fig. 2, where initial HMQC was replaced by a simple ^{1}H frequency labeling period (other experimental details are provided in the caption to Fig. S2). Signal-to-noise ratio (S/N) was estimated for well isolated H1′–H6/H8 cross-peaks in the 3D and 4D spectra. The comparison of peak intensities (cf. Fig. S2) shows an approx. 7 and 20 % loss of sensitivity for 4D spectrum compared to 3D ^{13}C(aro)- and ^{13}C(ribo)-edited NOESY, respectively. In the case of RNA studied, this resulted in a loss of $\sim$10 % of relevant H1′–H6/H8 cross-peaks (critical S/N = 6 for peak detection was assumed). The intrinsically lower sensitivity of the 4D spectrum is in accordance with theoretical predictions, and—similarly to the transition from 2D to 3D NOESY experiments—is a price for resolution enhancement. It is noteworthy that only 5 and 4 (of 62 present) H1′–H6/H8 cross-peaks can be unambiguously assigned in the test 3D spectra, whereas 4D spectrum provides unique assignment for 16 (of 55 detectable) cross-peaks in the studied 34-nt

Table 1 The impact of resolution on the ambiguity of cross-peak assignment 3D ^{13}C(ribo)-, ^{13}C(aro)- and 4D C,C-edited NOESY spectra

Spectrum	Average ambiguity[a]			Number of unique assignments		
	aro	ribo	Effective	aro	ribo	Effective[c]
4D C,C-edited	1.5	1.3[b]	2.1	28	33[b]	16
3D C(ribo)-edited	2.3	1.3[b]	3.2	10	32[b]	4
3D C(aro)-edited	1.3[b]	4.2	5.1	43[b]	8	5

[a] The ambiguity was defined as the number of peaks (in ribose or aromatic HSQC) that match ^{1}H chemical shift (or a pair of ^{1}H and ^{13}C chem. shifts) within the limits determined by obtained spectral resolution. The following half-widths at half height were assumed: 0.042 ppm (indirect ^{1}H), 0.52 ppm (aromatic ^{13}C), 0.47 ppm (ribose ^{13}C), 0.017 ppm (ribose ^{1}H directly detected) and 0.021 ppm (aromatic ^{1}H directly detected). 'aro' and 'ribo' refer to the multiplicity of independent assignment of aromatic and ribose chemical shift(s) of a cross-peaks, respectively

[b] The specific resonances were directly detected

[c] The cross-peak is considered uniquely assigned if *both* origin (e.g. aromatic) and destination (e.g. ribose) proton can be unambiguously identified

RNA. The average multiplicity of possible assignments of aromatic spins is reduced by a factor of 1.5 when 3D ^{13}C(ribo) and 4D C,C-edited NOESY are compared (see Table 1). For larger RNAs the complexity of 3D spectra is expected to render them even less useful for sequential assignment compared to a 4D experiment. We wish also to emphasize that sensitivity losses with increased size of RNA can be minimized by transverse relaxation optimizations (e.g. MQ-coherences or TROSY), while increased spectral overlap can only be alleviated by substantial resolution enhancement.

In the following discussion we will focus on the strategy of assignment of aromatic and anomeric protons. In order to perform sequential assignment effectively, two aromatic planes and two ribose planes of the 4D C,C-edited NOESY spectrum are needed. Additionally, one displays the aromatic 2D ^{13}C-HSQC spectrum, the anomeric region of a ribose 2D ^{13}C-HSQC spectrum and a 2D NOESY spectrum, preferably recorded in D_2O. To start the assignment, well-resolved H1′/C1′ cross-peaks in the ribose 2D ^{13}C-HSQC spectrum need to be selected (Figs. 3a, e and S3f).

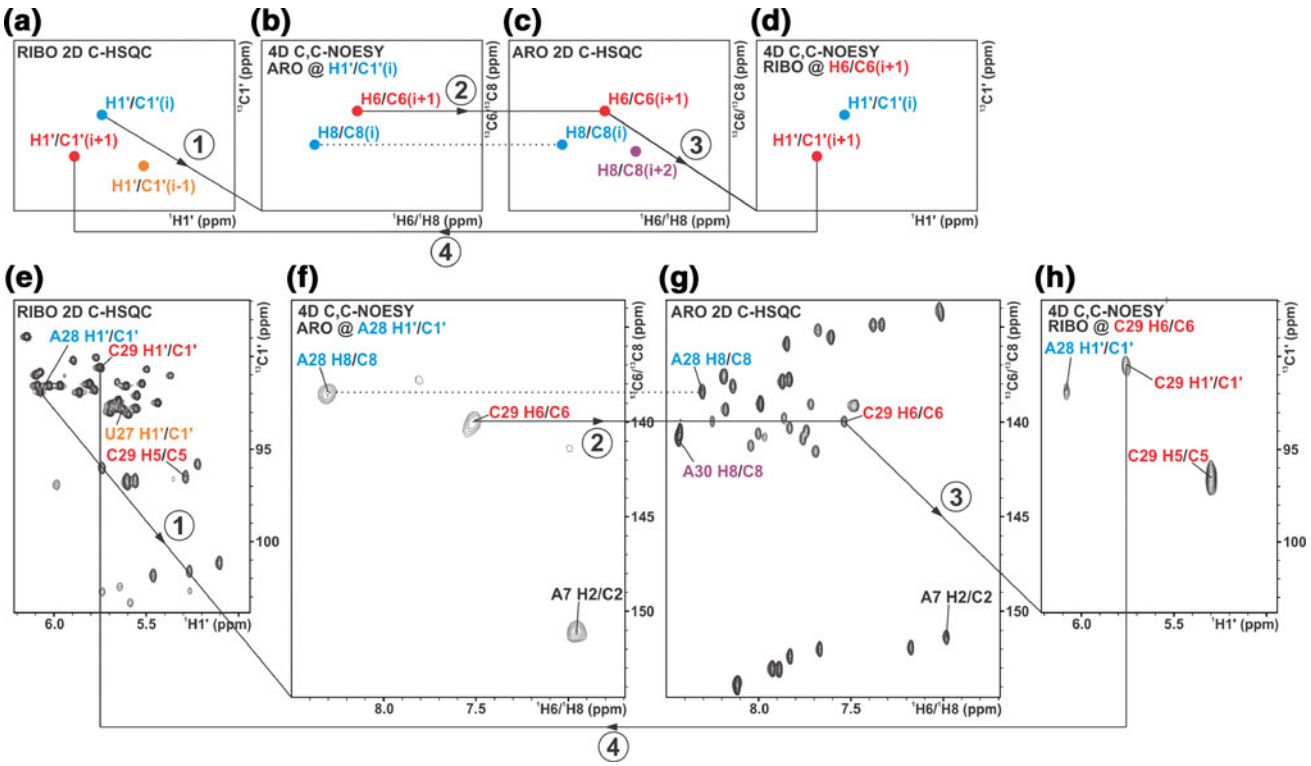


Fig. 3 Strategy of the cross-peak assignment employing 4D C,C-edited NOESY spectrum: a schematic representation (**a**–**d**) and the assignment of the A28 and C29 residues (**e**–**h**). *ARO* aromatic plane, *RIBO* ribose plane. ^{13}C-HSQC spectra (**a**, **c**) are used for convenient navigation in a 4D spectrum (**b**, **d**) and to recognize possible resonance overlap. In the steps (*1*) and (*3*) the corresponding planes of a 4D spectrum is shown for a selected peak in a HSQC spectrum (**a**, **c**). In the steps (**2**) and (**4**) the inter-residual ($i + 1$) peaks identified in a 4D spectrum are used for the assignment of peaks in the HSQC spectra. The ribose plane for H6/H8 (i), similar to (**d**), is not shown as it corresponds to the sequential walk in the *opposite* direction (it contains H1′/C1′ (i) and H1′/C1′ ($i - 1$) cross-peaks)

The direction of the sequential walk in the 2D NOESY and 4D C,C-edited NOESY spectra is unknown, since it is not clear which peak is the intra-nucleotide or the sequential one. Once a H1′/C1′ cross-peak is selected, then the aromatic plane of the 4D C,C-edited NOESY spectrum is examined at this H1′/C1′ chemical shift (Fig. 3b, f). Two cross-peaks should be observed and both of them are verified in the aromatic 2D ^{13}C-HSQC spectrum (Fig. 3c, g) and in the ribose plane of the 4D C,C-edited NOESY spectrum at the H6/C6 or H8/C8 chemical shifts (Fig. 3d, h). The common cross-peak is the starting H1′/C1′ (i) while the other cross-peaks present in two ribose planes of a 4D spectrum correspond to nucleotides $(i + 1)$ and $(i - 1)$. If more than two cross-peaks are present in the 2D planes of the 4D C,C-edited NOESY spectrum, the most intensive peaks are checked and simultaneously verified in the 2D NOESY spectrum. The sequential walk is performed as long as peaks can be unambiguously assigned in the aromatic plane of the 4D C,C-edited NOESY spectrum.

The assignment strategy was verified on 34-nt RNA using no prior information. The results are illustrated in Fig. 4. It is convenient to start the assignment of the characteristic cross-peaks, however, one should also aim at collecting as many sequential walks as possible starting from randomly chosen (non-specific) remaining ones. One characteristic peak is for the first residue in the sequence, G1, which has typical anomeric and aromatic ^{1}H chemical shifts (δ ~5.9 and ~8.2 ppm, respectively). The G1 aromatic proton (G1 H8) shows correlation only to its own anomeric proton, thus only one cross-peak is observable in the 2D NOESY at the G1 H8 proton chemical shift. Similarly, in the ribose plane of the 4D C,C-edited NOESY at the G1 H8/C8 chemical shifts only one cross-peak is observed. Employing the strategy outlined above, it was possible to perform the walk from G1 till the U9 H6/C6 cross-peak (marked in blue in Fig. 4). At this point the walk is interrupted, because of the overlap with the other cross-peaks in the ribose plane of the 4D C,C-edited NOESY spectrum at the U9 H6/C6 chemical shifts. Another characteristic H1′/C1′ cross-peak belongs to G19 (δ ~3.7 ppm), the first residue after the GAAA tetraloop. Although this cross-peak is outside the recorded anomeric proton region in the 4D C,C-edited NOESY, we could assign residues A20 and C21 (marked in blue in Fig. 4) by utilizing cross-strand H2-H1′ contacts, however, overlap in the aromatic plane of the 4D C,C-edited NOESY spectrum at the C21 H1′/C1′ chemical shift interrupts the sequential walk.

Subsequently, remaining and well-resolved H1′/C1′ cross-peaks were chosen as a starting point for other sequential walks. Four additional walks were found (marked in red in Fig. 4). In a few cases the walk was possible only to the neighbouring residue due to the high number of the cross-peaks or due to missing peaks (marked with the dashed line and a × sign in Fig. 4). Once all possible sequential walks are constructed, the obtained chains are put into the sequence. One could distinguish H8/

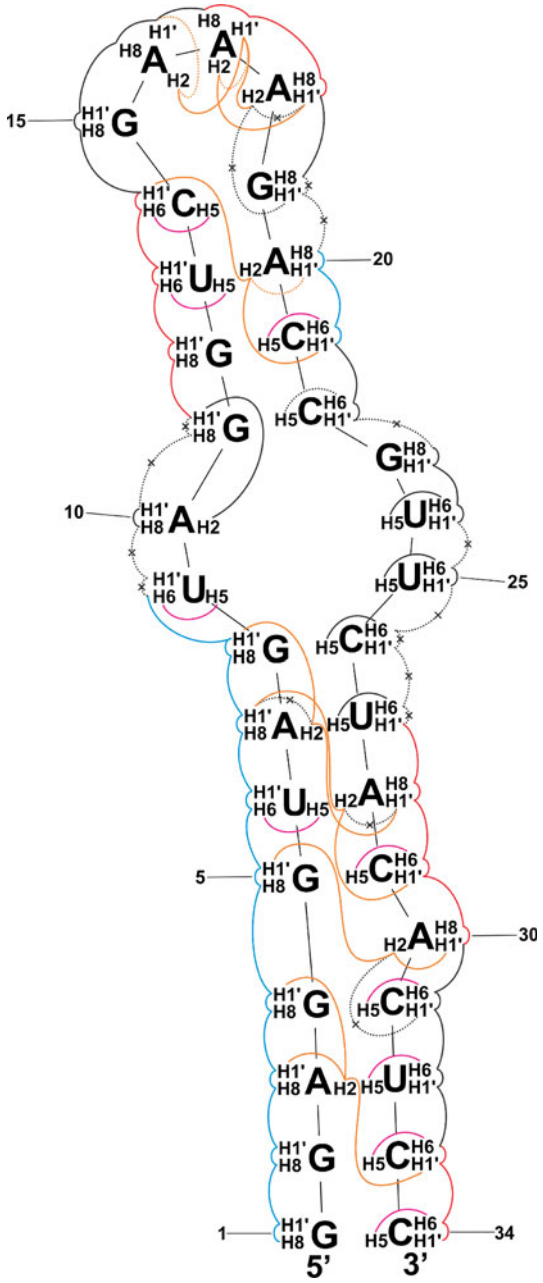


Fig. 4 The schematic presentation of 34-nt RNA showing the NOE correlations observed in the 4D C,C-edited NOESY spectrum. *Blue lines* indicate the two partial sequential walks originating from the characteristic H8/C8 peaks of G1 and G19, while *red lines* indicate additional four found sequential walks. *Orange lines* indicate assigned signals from H2 protons, whereas *magenta lines* indicate assigned signals between H5 and H6 protons. Very weak peaks are marked with the *dashed lines* and missing peaks are marked with the *dashed lines* and × signs. *Black lines* represent peaks present in 4D spectrum which are hard to assign employing this experiment only

C8 (of adenines and guanines) from H6/C6 (of cytosines and uracils) cross-peaks on the basis of peak sign in the constant-time aromatic HSQC. During the walk every residue can be classified to one of these two classes (A/G or C/U), which helps to match the walk into the sequence. Additionally, specific contacts to H2/C2 of adenine can be utilized. Assigned adenine H2/C2 resonances can effectively support the assignment of the broken sequential walks. H2/C2 signals of the adenines involved in the AU Watson–Crick base pairs (A3, A7, A20 and A28) and of the bulged A30 residue were determined from the strong cross-peaks in the imino-H2 proton region of the 2D NOESY in H_2O. One observes the ribose planes of the 4D C,C-edited NOESY spectrum at the chemical shifts of each of H2/C2. The planes show typically three cross-peaks for correlations between H2 and anomeric protons: a very weak intra-residual peak, one to the following residue and one across the strand (marked in orange in Fig. 4). Additionally, H5/C5 resonances of cytosines and uracils, which are easily identified by strong cross-peaks between H5 and H6 protons, help to complete the sequential walks (marked in magenta in Fig. 4).

Overall, 89 % of the aromatic (H6 and H8) and anomeric proton peaks of 34-nt RNA were assigned in the 4D C,C-edited NOESY spectrum. The missing resonances belong to the residues in the internal loop or to the proximate residues (see Fig. 4). In our previous study we have shown that the asymmetric internal loop adopts two energetically comparable families of structures, which both satisfy NMR data (Cevec et al. 2010). This explains why these resonances are not observable in the 4D C,C-edited NOESY spectrum.

To conclude, we have demonstrated the utility of high-resolution selective 4D C,C-edited NOESY experiment in sequential assignment of RNAs. The strategy proposed here effectively complements more sensitive (but crowded) 2D homonuclear NOESY spectra by resolving overlapped cross-peaks in two different ^{13}C dimensions. We showed that a single 4D experiment can replace two complementary ^{13}C-edited 3D spectra that had to be recorded traditionally in an equal time. Enhanced resolution of 4D spectrum enables reliable assignment of RNAs which often feature limited dispersion of ^{1}H and ^{13}C chemical shifts. The non-uniformly sampled experiment was recorded using only 14 % of data points required conventionally. Significantly larger time savings are expected at higher B_0 fields (scaling with B_0^3), higher intrinsic sensitivity (as there are less sampling points needed) or when substantial resolution enhancement (i.e. extended sampling of the indirect time-domain) is required, e.g. for larger RNAs. The issue of non-uniform sampling artifacts is circumvented by virtually complete suppression of diagonal peaks, which allows rapid reconstruction of high-resolution four-dimensional spectrum without compromising sensitivity.

Acknowledgments This work was supported by Bio-NMR project funded by European Commission's 7th Framework Program (contract No. 1618630) and the Slovenian Research Agency, the Ministry of Higher Education, Science and Technology of the Republic of Slovenia [P1-0242 and J1-4020]. J. S. thanks Polish National Science Centre for the financial support with the Grant No. 2012/05/N/ST4/01120. The study was carried out at the Biological and Chemical Research Centre, University of Warsaw, established within the project co-financed by European Union from the European Regional Development Fund under the Operational Programme Innovative Economy, 2007–2013.

References

Brutscher B, Simorre JP (2001) Transverse relaxation optimized HCN experiment for nucleic acids: combining the advantages of TROSY and MQ spin evolution. J Biomol NMR 21:367–372

Brutscher B, Boisbouvier J, Pardi A, Marion D, Simorre JP (1998) Improved sensitivity and resolution in ^{1}H-^{13}C NMR experiments of RNA. J Am Chem Soc 120:11845–11851

Cevec M, Thibaudeau C, Plavec J (2010) NMR structure of the let-7 miRNA interacting with the site LCS1 of lin-41 mRNA from Caenorhabditis elegans. Nucleic Acids Res 38:7814–7821

Coggins BE, Zhou P (2008) High resolution 4-D spectroscopy with sparse concentric shell sampling and FFT-CLEAN. J Biomol NMR 42:225–239

Diercks T, Truffault V, Coles M, Millett O (2010) Diagonal-Free 3D/4D HN, HN-TROSY-NOESY-TROSY. J Am Chem Soc 132:2138–2139

Esteller M (2011) Non-coding RNAs in human disease. Nat Rev Genet 12:861–874

Fiala R, Jiang F, Sklenář V (1998) Sensitivity optimized HCN and HCNCH experiments for $^{13}C/^{15}N$ labeled oligonucleotides. J Biomol NMR 12:373–383

Fiala R, Czernek J, Sklenář V (2000) Transverse relaxation optimized triple-resonance NMR experiments for nucleic acids. J Biomol NMR 16:291–302

Fürtig B, Richter C, Wöhnert J, Schwalbe H (2003) NMR spectroscopy of RNA. Chembiochem 4:936–962

Goddard TD, Kneller DG (2008) SPARKY 3: University of California, San Francisco, http://www.cgl.ucsf.edu/home/sparky

Hyberts SG, Frueh DP, Arthanari H, Wagner G (2009) FM reconstruction of non-uniformly sampled protein NMR data at higher dimensions and optimization by distillation. J Biomol NMR 45:283–294

Kay LE, Clore GM, Bax A, Gronenborn AM (1990) Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1-beta in solution. Science 249:411–414

Kazimierczuk K, Misiak M, Stanek J, Zawadzka-Kazimierczuk A, Koźmiński W (2012) Generalized Fourier transform for non-uniform sampled data. Top Curr Chem 316:79–124

Kupče Ē, Freeman R (1995) Adiabatic pulses for wide-band inversion and broad-band decoupling. J Magn Reson Ser A 115:273–276

Kupče Ē, Boyd J, Campbell ID (1995) Short selective pulses for biochemical applications. J Magn Reson Ser B 106:300–303

Lu K, Miyazaki Y, Summers MF (2010) Isotope labeling strategies for NMR studies of RNA. J Biomol NMR 46:113–125

Luan T, Jaravine V, Yee A, Arrowsmith CH, Orekhov VY (2005) Optimization of resolution and sensitivity of 4D NOESY using multi-dimensional decomposition. J Biomol NMR 33:1–14

Marino JP, Schwalbe H, Anklin C, Bermel W, Crothers DM, Griesinger C (1994) Three-dimensional triple-resonance ^{1}H, ^{13}C, ^{31}P experiment: sequential through-bond correlation of ribose protons and intervening phosphorus along the RNA oligonucleotide backbone. J Am Chem Soc 116:6472–6473

Marino JP, Diener JL, Moore PB, Griesinger C (1997) Multiple-quantum coherence dramatically enhances the sensitivity of CH and CH2 correlations in uniformly ^{13}C-labeled RNA. J Am Chem Soc 119:7361–7366

Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. Nat Rev Genet 10:155–159

Mobli M, Stern AS, Bermel W, King GF, Hoch JC (2010) A non-uniformly sampled 4D HCC(CO)NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment. J Magn Reson 204:160–164

Nikonowicz EP, Pardi A (1993) An efficient procedure for assignment of the proton, carbon and nitrogen resonances in $^{13}C/^{15}N$ labeled nucleic acids. J Mol Biol 232:1141–1156

Riek R, Pervushin K, Fernandez C, Kainosho M, Wüthrich K (2001) [^{13}C, ^{13}C]- and [^{13}C, ^{1}H]-TROSY in a triple resonance experiment for ribose-base and intrabase correlations in nucleic acids. J Am Chem Soc 123:658–664

Stanek J, Augustyniak R, Koźminski W (2012) Suppression of sampling artefacts in high-resolution four-dimensional NMR spectra using signal separation algorithm. J Magn Reson 214: 91–102

Stanek J, Nowakowski M, Saxena S, Ruszczyńska-Bartnik K, Ejchart A, Koźmiński W (2013) Selective diagonal-free ^{13}C, ^{13}C-edited aliphatic–aromatic NOESY experiment with non-uniform sampling. J Biomol NMR 56:217–226

Varani G, Aboulela F, Allain FHT (1996) NMR investigation of RNA structure. Prog Nucl Magn Reson Spectrosc 29:51–127

Vuister GW, Clore GM, Gronenborn AM, Powers R, Garrett DS, Tschudin R, Bax A (1993) Increased resolution and improved spectral quality in four-dimensional $^{13}C/^{13}C$-separated HMQC-NOESY-HMQC spectra using pulsed-field gradients. J Magn Reson Ser B 101:210–213

Wen J, Zhou P, Wu JH (2012) Efficient acquisition of high-resolution 4-D diagonal-suppressed methyl–methyl NOESY for large proteins. J Magn Reson 218:128–132

Werner-Allen JW, Coggins BE, Zhou P (2010) Fast acquisition of high resolution 4-D amide–amide NOESY with diagonal suppression, sparse sampling and FFT-CLEAN. J Magn Reson 204:173–178

Wijmenga SS, van Buuren BNM (1998) The use of NMR methods for conformational studies of nucleic acids. Prog Nucl Magn Reson Spectrosc 32:287–387

Xia YL, Man D, Zhu G (2001) 3D H-aro-NOESY-CH_3NH and C-aro-NOESY-CH_3NH experiments for double labeled proteins. J Biomol NMR 19:355–360

Xu Y, Lin Z, Ho C, Yang D (2005) A general strategy for the assignment of aliphatic side-chain resonances of uniformly ^{13}C, ^{15}N-labeled large proteins. J Am Chem Soc 127:11920–11921

Xu YQ, Zheng Y, Fan JS, Yang DW (2006) A new strategy for structure determination of large proteins in solution without deuteration. Nat Methods 3:931–937