ELSEVIER

J. Biochem. Biophys. Methods 42 (2000) 153–178

JOURNAL OF
biochemical and
biophysical
methods

www.elsevier.com/locate/jbbm

The identification of the A-type RNA helices in a 55mer RNA by selective incorporation of deuterium-labelled nucleotide residues (Uppsala NMR-window concept)

Tatiana Maltseva, Andras Földesi, Jyoti Chattopadhyaya*

Department of Bioorganic Chemistry, Box 581, Biomedical Center, University of Uppsala, S-751 23 Uppsala, Sweden

Received 1 December 1999; accepted 2 December 1999

Abstract

The 55-nt long RNA, modelling a three-way junction, with non-uniformly incorporated deuterated nucleotides has been synthesised in a pure form. The NMR-window part in this partially deuterated 55mer RNA consists of natural non-enriched nucleotide blocks at the three-way junction (shown in a square box in Fig. 2), whereas all other nucleotides of the rest of the molecule are partially deuterated (>97 atom% ^{2}H at C2′, C3′, C5′, C5, and ~ 50 atom% ^{2}H at C4′). The secondary structure of this 55mer RNA was determined by 2D ^{1}H NOESY spectroscopy in D_2O or in 10% D_2O–H_2O mixture. The use of deuterated building blocks in the specific region of the 55mer RNA allowed us to identify two distinct A-type RNA helices in a straightforward manner by observing connectivities of H1′ with the basepaired imino and the aromatic H2 of all adenosine nucleotides as the first step for the determination of its tertiary structure in a cost- and time-effective manner without employing any $^{13}C/^{15}N$ labelling. These two decameric helices involve 40 nucleotides, for which all non-exchangeable H1′, H6, H2, H8 and H5 protons (all 40 H1′, all 40 H6 or H8 aromatics, all seven H2 of adenine nucleotide and all four non-deuterated H5 of cytosines) as well as all 16 exchangeable imino protons (with the exception of four terminal basepairs) and 16 amino protons of cytosines have been assigned. Since all aromatic-H2′, H3′ as well as H5′/5″ crosspeaks from partially deuterated residues have been eliminated from the NMR spectra, the observation of natural nucleotide residues in the NMR window part has essentially been simplified. It has been found that the crosspeaks from the natural nucleotides located at the three-way junction in the NMR-window part show different degrees of

*Corresponding author. Fax: +46-185-544-95.
E-mail address: jyoti@bioorgchem.uu.se (J. Chattopadhyaya)
Website address: http://biorgchem.boc.uu.se

0165-022X/00/$ – see front matter
PII: S0165-022X(99)00057-3

line-broadening, thereby indicating that the various nucleotide residues have very different mobilities with respect to themselves as well as compared to other nucleotides in the helices. The assignment of H2′ and H3′ in the NMR-window part has been made based on NOESY and DQF-COSY crosspeaks. It is noteworthy that, even in this preliminary study, it has been possible to identify 10 H2′ out of total 14 and 9 H3′ out of 14. The data show that expanded AU containing a tract of 55mer RNA does not self-organise into a tight third helix, as the two decameric A-type helices, across the three-way junction which is evident from the absence of any additional imino protons, except those that already have been assigned for the two decameric helices.

Keywords: ^{2}H labelling; 55mer RNA; NMR; Structure; Selective ^{2}H incorporation

1. Introduction

The overall structure and dynamics of large RNAs are dictated by the local formation of various modules such as helices, hairpins, internal loops, pseudoknots and junctions [4]. The self-organisation of these elements dictates the folding of the RNA [3,42] into a functional tertiary structure. It has recently been shown [3,6,24,36,42] that the junction motif is important for the global structure of RNA. It has also been found that an appropriate folding of the junction is critically dictated by the nucleotide sequence [36] as well as coordination of metal ion [6] or any other ligands in the junction.

The structure of the three-way junction is not readily addressed by NMR spectroscopy mainly because of the conformational heterogeneity of RNA in a quasiphysiological condition [2,3,21,30], which is further complicated by severe resonance overlap problem within a very small stretch between 6.5 and 4.0 ppm. Recently, to solve the problem of overcrowding/overlapping resonances in ^{1}H NMR spectra, we have developed synthetic methodologies for stereospecific deuteration of sugar moieties in nucleosides [8]. The usefulness of these deuterated nucleosides have been shown by incorporating them into the specific domains in a non-uniform manner of a 21mer RNA [10] or in large oligo-DNAs [1,41] using the solid-phase synthesis protocol, or in a 31mer RNA [12] using DNA-dependent T7 RNA polymerase. Thus, these synthetic partially deuterated nucleotide blocks are segmentally incorporated into either oligo-DNA or -RNA by solid-phase chemistry to create ^{1}H NMR invisible parts and a short ^{1}H NMR visible part (the Uppsala 'NMR-window' concept), which is then used to extract the structural information [10,12]. The usefulness of this non-uniform deuterium labelling of oligo-RNA and -DNA is many-fold: (i) it simplifies the spectral crowding and the coupling patterns [9–12,25,41], (ii) reduces the line-broadening associated with ^{1}H dipolar relaxation as well as spin diffusion [1], (iii) gives much improved resolution and sensitivity of the proton resonances vicinal to the deuterated centers [25], (iv) increases nOe intensities [1,25] and (v) helps to probe dynamics by selective T_1 and T_2 measurements [26,27].

Although in a few conformational studies for large RNAs (40 nt or larger) with natural abundance [2,3,13,15,21,28,30,31] or with uniformly ^{13}C/^{15}N-labelled RNA

[7,13,14,17–19,23,32–35,39] the assignments within some of the large helices were performed, indicating that the helices around the junction were formed, it was often shown [3,21] that the catalytic core does not adopt a single conformation. Moreover, in the case of the hammerhead ribozyme [21] it was suggested [3,21] that there is considerable flexibility in the dynamic behaviour of the three-way junction. It is likely that this flexibility is an intrinsic property of such RNA junctions, which adopt a single biologically significant conformation upon binding to protein or other RNA. The slow conformational exchange between these conformations could severely complicate the application of NMR spectroscopy due to the large resonance broadening. It is also noteworthy that this is commonly encountered in the NMR structure determination of even a relatively small oligonucleotide such as a 20mer DNA duplex [41].

One of the ways to identify the A-type helices in the overall folding of the RNA is to establish its basepairing pattern. Recently [14] the selectively labelled ^{15}N-G,U nucleosides were used to identify that pattern in *E. coli* 5S rRNA (120 nucleotides). An alternative approach to identify the A-type helices is to establish the nOe connectivity between basepaired G imino protons and the sugar H1′. This connectivity [16] originates through a pathway from $G_{(imino)}$–$G_{(hydrogen\text{-}bonded\ amino)}$–$G_{(non\text{-}hydrogen\text{-}bonded\ amino)}$ to both the H1′ of the 3′-end of the same strand and of the opposite strand (Fig. 1). In our earlier work [10], we have demonstrated the possibility to improve the detection of this connectivity by using partially deuterated nucleoside blocks due to the elimination of the spin diffusion on H1′ protons because of substitution of the vicinal protons in the sugar moiety by deuterium. This has led us to show [10] that the 21mer RNA folds in both stem and loop because the A-type helix was easily distinguishable from the loop part by the typical nOe connectivity pattern.

Herein, we report an application of our NMR-window approach based on non-uniform and selective incorporation of the partially deuterated nucleotide residues into a 55mer RNA (three-way junction) [40] to demonstrate that (i) it is possible to distinguish its helices from the non-hydrogen bonded part, and (ii) our approach allows us to assign the nonexchangeable protons of a large RNA molecule in an unambiguous manner. Thus, it has emerged that the 55mer RNA has two distinct A-type RNA helices, which have been identified in a straightforward manner by observing connectivities of H1′ with the basepaired imino and H2 protons of all adenosine nucleotides as the first step for the determination of its tertiary structure in a cost- and time-effective manner without employing any ^{13}C/^{15}N labelling. The non-exchangeable H1′, H6, H2, H8 and H5 protons of the two decameric helices involving 40 nucleotides in the 55mer RNA were also assigned unambiguously, including 40 H1′, 40 H6 or H8 aromatics, seven H2 of adenine nucleotides and all four non-deuterated H5 of cytosines. Our data show no indication of the presence of a tight three-way junction under the condition of NMR measurement for the 55mer RNA [40].

2. Experimental

All figures and schemes described in this paper are available in colour format in authors website.

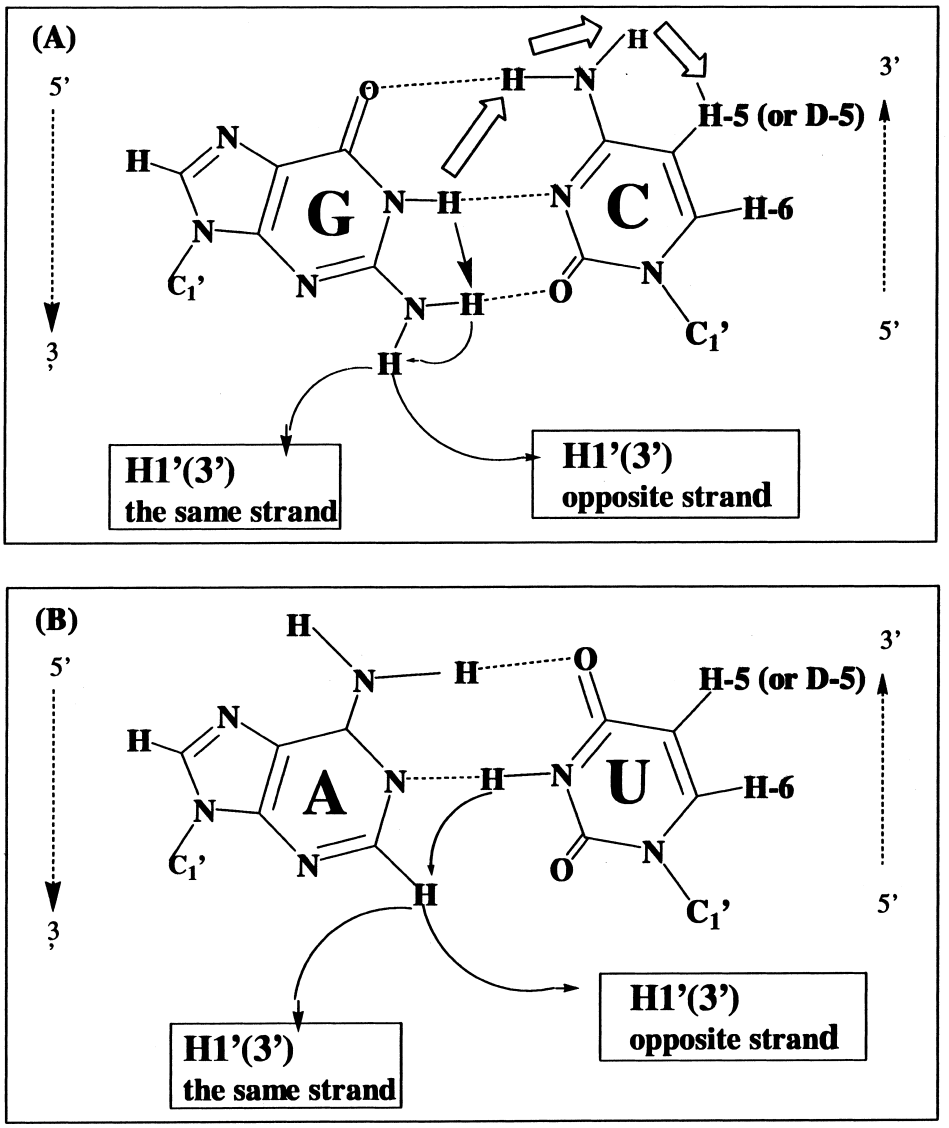


Fig. 1. Schematic presentation of (A) GC and (B) AU basepairs to show: (1) (pink arrows in panel A) the nOe relaxation pathways of Im proton of G to H1′ of the 3′-end nucleotide ($n + 1$) of both the same strand as well as with that of the opposite strand; (2) (red arrows in panel B) the nOe relaxation pathways of A(H2) proton of A with H1′ of the 3′-end nucleotide ($n + 1$) of both the same strand as well as with that of the opposite strand; and (3) (blue arrows on panel A) the spin diffusion pathway between imino and H5 of neighbouring C originating from GC basepair, H5↔H6↔4-NH_2↔N1H.

2.1. Synthesis and sample preparation

The 2′,3′,4′#,5′,5″-$^2H_5$-ribonucleoside blocks (A, U and G; >97 atom% deuterium enrichment at C2′, C3′ and C5′; ~50 atom% incorporation at C4′#) were prepared following our published procedure [8]. The 2′,3′,4′#,5′,5″-$^2H_5$-uridine block was deuterium labelled at C5 upon equilibration in ^2H_2O with K_2CO_3 as catalyst under reflux for 2 days. The 5,2′,3′,4′#,5′,5″-$^2H_6$-uridine derivative (>95 atom% deuterium enrichment at C5) was converted to $^2H_6$-cytidine via C^4-triazolyl chemistry during the ammonia deprotection step of the oligomers [20,29]. The 20 and 35mer RNAs were synthesized by the phosphoramidite protocol on an Expedite Nucleic Acid Synthesis System (PerSeptive Biosystems) using standard 1 μmol scale RNA program (8× for 20mer and 10× for 35mer). The 20 and 35mers were dissolved in equimolar ratio in a buffer containing 100 mM NaCl and 10 mM phosphate (pH 7.0). After lyophilisation from ^2H_2O (99.9 atom%) three times, the sample was dissolved in 200 μl ^2H_2O (99.996 atom%) or ^2H_2O:H_2O = 1:9 (v/v) to a final concentration of ~2 mM for the appropriate NMR measurements performed in a Shigemi tube.

2.2. Nuclear magnetic resonance spectroscopy

All ^1H NMR spectra were recorded on Bruker DRX-500 and 600 NMR spectrometers. Phase-sensitive NOESY experiments were performed at 298 and 303 K using the following parameters: mixing time was varied 0.3, 0.2, 0.1 and 0.6 s, 4K complex data points in t_2, 1K or 512 complex data points in t_1, a relaxation delay of 5 or 10 s, a sweep width of 10 ppm in both dimensions, acquisitions per FID were 64; a Lorenz apodization function for t_2, and a shifted sine-bell apodization function for the t_1 dimension were used. The data were zero-filled in t_1 to give 4×4K complex data points. The base correction has been separately done for all regions. The residual water resonance was saturated with very low power during the relaxation delay. Two-dimensional data sets for DQF-COSY spectra were collected in the phase-sensitive mode with the time-proportional phase incrementation without phosphorus decoupling. Typically 4K data points were collected for each 1K t_1 values in DQF-COSY experiments. The 4×2K data points were resolution enhanced by a shifted square sine-bell window function in both the t_1 and t_2 directions, then Fourier transformed and phase adjusted. Relaxation delays of 5 or 10 s were used. The data were collected with the nonspinning sample to avoid t_1 noise.

3. Results and discussion

3.1. Analysis of the DQF-COSY spectra of 55mer RNA in the NMR-window: does it adopt a single conformation at the three-way junction?

The schematic presentation of the partially deuterated 55mer RNA as a tight three-way junction is shown in Fig. 2. The nucleotide residues shown within the square box in Fig. 2 for the 55mer RNA are those with isotopic composition at the natural

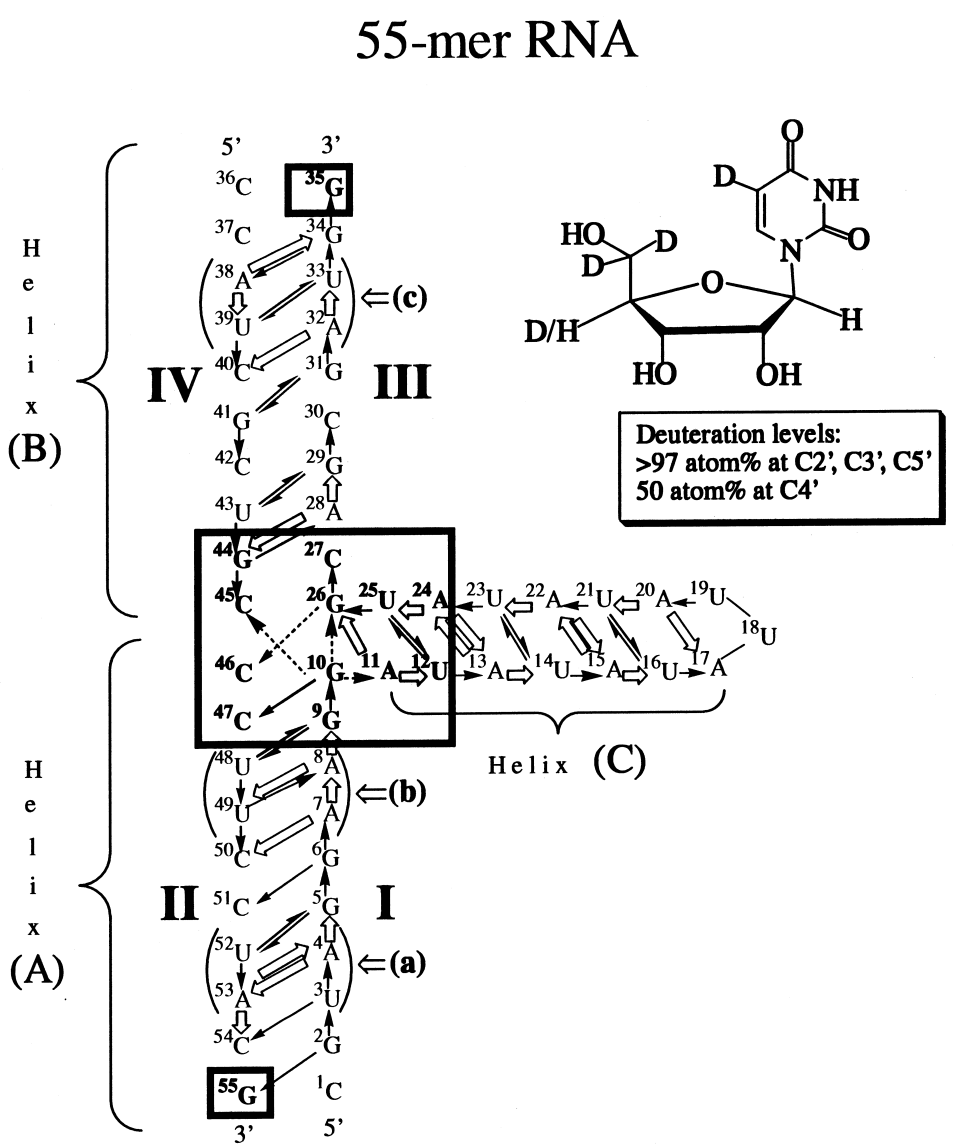


Fig. 2. The schematic presentation of the three-way junction RNA consisting of 55 nucleotide residues. All nucleosides are partially deuterated (diagram at the right corner) except for the residues presented in three square boxes (for ^{35}G, ^{55}G, as well residues at the core part: ^{44}G, ^{45}C, ^{46}C, ^{47}C, ^{9}G, ^{10}G, ^{11}A, ^{12}U, ^{24}A, ^{25}U, ^{26}G, ^{27}C). The sequence between ^{1}C and ^{10}G residues is denoted as strand (I), the sequence between ^{46}C and ^{55}G residues as strand (II), the sequence between ^{26}G and ^{35}G as strand (III), and the sequence between ^{36}C and ^{45}C as strand (IV). The 5′-UA tract (a), 5′-AA tract (b), and 5′AU tract (c) are shown in parentheses. The expected nOe between imino proton and H1′ of the 3′-end nucleoside ($n + 1$) of both the same as well as of opposite strand for a typical A-RNA helix are shown by solid arrows. The expected nOe between H2A and H1′ of the 3′-end nucleotide residue ($n + 1$) for both the same and the opposite strand are also shown but by arrows in the outline font. Note however that nOe are only found for helices (A) and (B).

abundance, whereas all other nucleotide residues outside the square box are partially enriched with deuterium atoms (>97 atom% ^{2}H at C2′, C3′, C5′, C5, and ~ 50 atom% ^{2}H at C4′). Such deuterium-enriched oligo-RNA has been designed [8] and synthesised [40] in order to minimise spectral overlap at the three-way junction to address if the 55mer RNA adopts a single conformation.

If the 55mer RNA adopted a single conformation as shown in Fig. 2, it is expected to show only six crosspeaks between H5 and H6 in the DQF-COSY spectra for the non-deuterated ^{45}C, ^{46}C, ^{47}C, ^{27}C, ^{25}U and ^{12}U residues which are located within the NMR-window region, since H5 proton is replaced by deuterium for all other nucleotide residues outside the NMR-window region (see Section 2). In DQF-COSY spectra at 600 MHz, these crosspeaks were not detectable; they were, however, observed at 500 MHz (Fig. 3A) in an experiment with the 1K data set in the F1 dimension. This is most

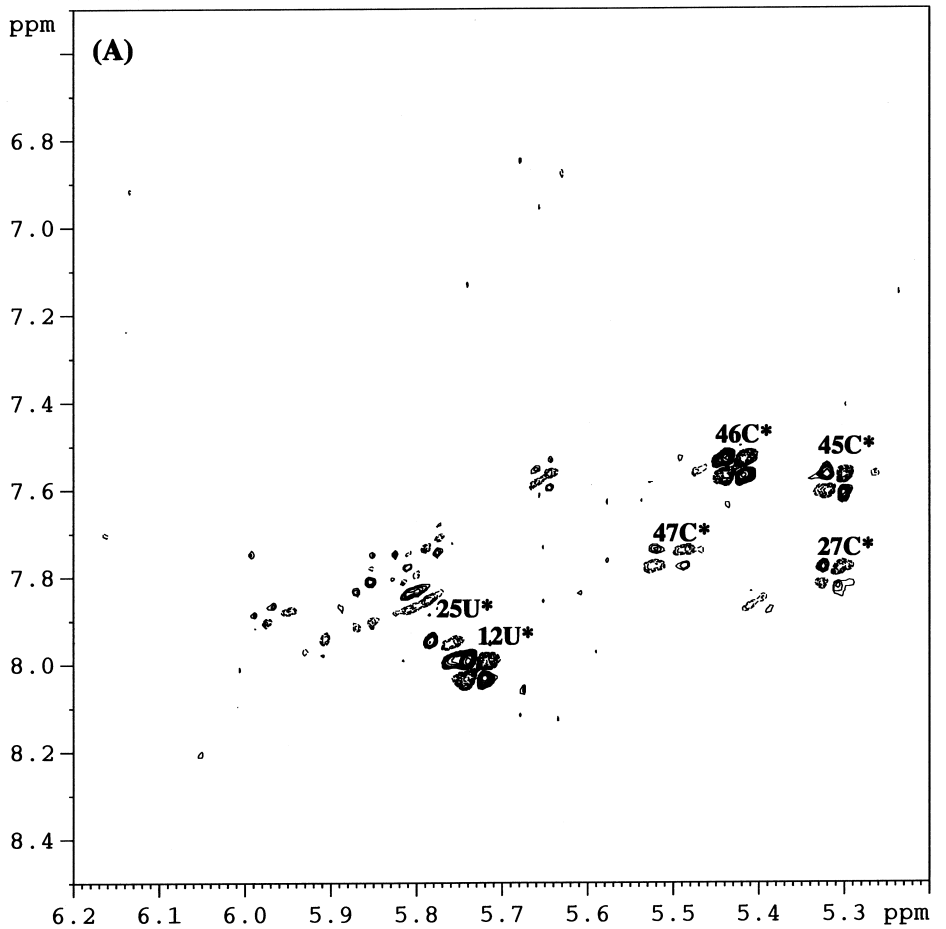


Fig. 3. Expanded DQF-COSY spectrum of aromatic–H5 region is shown in Panel (A), whereas Panel (B) shows the H1′–H2′ area of the 55mer RNA in D_2O at 25°C.

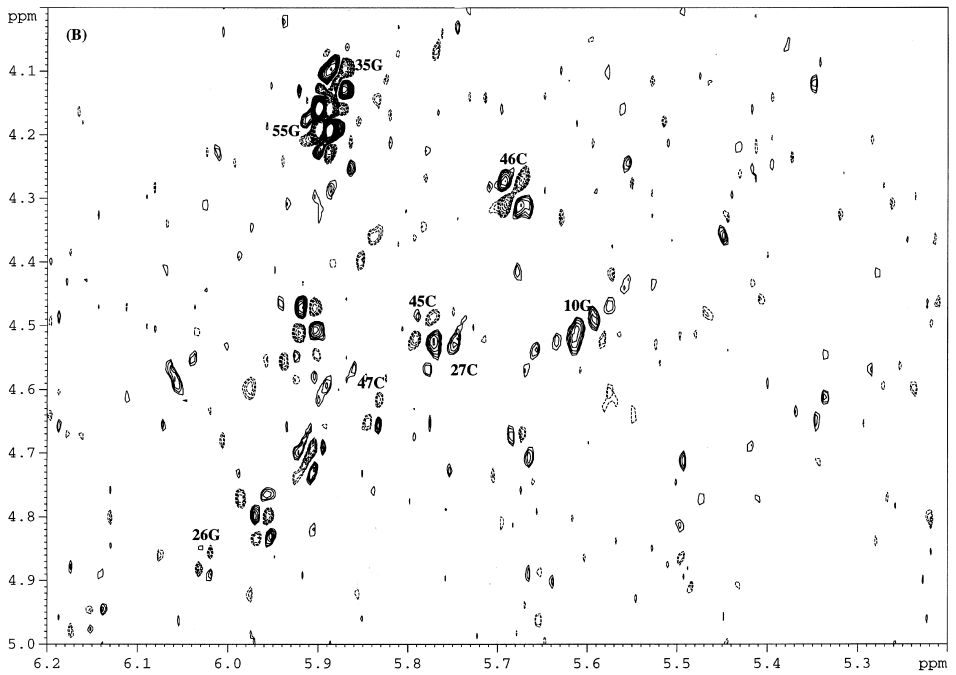


Fig. 3. (*continued*)

probably owing to a shorter T_2 relaxation time or a slow tumbling motion for this large 55mer RNA molecule. A similar observation has recently been made [22,23] where H5–H6 crosspeaks in DQF-COSY spectra have only been observed with a 400-MHz spectrometer. Fig. 3A shows six intense crosspeaks in the H5–H6 region: four of them are located at a more upfield region (5.5–5.2 ppm) and two at a more downfield part of the spectrum (5.8–5.7 ppm). At a higher threshold level, a few extra crosspeaks could be observed in the upfield part of the spectrum, but they were not clearly visible in the NOESY spectra, which may (or may not) indicate the presence of minor forms. The DQF-COSY data from the NMR window part of 55mer RNA clearly indicate that there is one single predominant conformation for the 55mer RNA.

3.2. Starting point of the assignment of 55mer RNA in its DQF-COSY spectra

In addition to the H5–H6 crosspeaks, 10 other crosspeaks could also be found in the H1′–H2′ area of the DQF-COSY spectra (Fig. 3B). It is well known that the $J_{1',2'}$ is close to zero [37,38] for most RNA sugar moieties with A-type conformation, hence no crosspeak is observable in the DQF-COSY spectra. We have, however, shown that the H1′–H2′ crosspeak of the terminal residue in a 21mer RNA hairpin [10] could be easily identified in DQF-COSY spectrum because it adopts ~1:1 North and South pseudo-rotational equilibrium, leading to a H1′–H2′ crosspeak with high intensity. This observation has led us to keep the terminal ^{35}G and ^{55}G residues non-deuterated in the 55mer RNA in order to have a clearly identifiable starting point in the proton resonance

assignment procedure. Indeed, we have observed two strong intensity crosspeaks amongst 10 observable H1′–H2′ crosspeaks (Fig. 3B) in the upfield region (crosspeaks at ~5.9/4.1 ppm), which have been attributed to H1′–H2′ of the terminal ^{35}G and ^{55}G residues. They have been used as the starting point in our assignment strategy for H1′–imino (see below). It is noteworthy that all the other eight H1′–H2′ crosspeaks have considerably lower intensity and are broadened because of the slow tumbling motion of 55mer in the NMR time scale. From our deuteration strategy, we can conclude that these observed H1′–H2′ crosspeaks belong to nucleotides located in the NMR-window part. Clearly, the presence of these crosspeaks indicates that corresponding sugars in this junction do not adopt a North-type conformation.

3.3. Stems (A) and (B) of 55mer RNA adopt the A-type helix with a single conformation: identification of the imino and A(H2) resonances

As a starting point in our assignment exercise for the 55mer RNA, we assumed (Fig. 2) that it folds itself into a structure consisting of three A-type RNA stems: two A-helix type decamers (A) (strands I + II) and (B) (strands III + IV), whereas the third 7-bp helix (C) is created by AU basepairs with a U loop. In Fig. 4, the downfield region of the ^{1}H-NMR spectrum of 55mer RNA in H_2O/D_2O is presented where the imino proton resonances are shown in the temperature range of 5–65°C. In general [30] the hydrogen-bonded imino protons of GC basepairs absorb between ~13.5–12 ppm, AU basepairs between ~15–13 ppm and the non-hydrogen-bonded iminos of hairpin around ~11–9 ppm. A careful examination of Fig. 4 shows that the numbers of resonances observed in the 5–30°C range at pH 7.0 are less than expected on the basis of the assumed secondary structure shown in Fig. 2. Indeed, in the imino proton region of AU basepair (15–13 ppm) at 25°C there are only seven resonances with unequal intensities. Moreover, there are no resonances between 11–9 ppm. A possible reason for this discrepancy could be that the AU type basepairs in stem (C) have a higher flexibility compared with AU basepairs in stems (A) and (B) because the latter are better stacked between GC basepairs compared to the former. It is noteworthy that the titration of 55mer RNA solution by Mg^{2+} did not have any influence on the imino proton spectra, suggesting that Mg^{2+} ion does not induce any additional stabilisation of the 55mer RNA structure. The temperature dependence of the imino protons, however, shows that the helices are kept intact up to 50°C; however, they completely melt at 65°C.

There are three independent lines of evidence which support the formation of a continuous stack of basepairs in decamer stems (A) and (B) (Fig. 2), but not in the heptameric AU stem-hairpin (C): (i) the nOe crosspeaks between the various imino (Im) protons; (ii) the nOe crosspeaks between imino protons and H1′ protons of the 3′ adjacent neighbours from the same and the opposite strands; and (iii) the nOe crosspeaks between imino and A(H2) protons. These are schematically presented in Scheme 1.

For an A-type RNA conformation, the substitution of the H2′ and H3′ by ^{2}H in the deuterated sugar residues will eliminate two strongest relaxation pathways, H1′–H2′ and H1′–H3′, which will inevitably lead to narrowing of the line width and increasing of the transverse relaxation time of the H1′ proton, compared with the natural counterpart. This in turn will increase the sensitivity of experiment with respect to nOe between AH(2)–H1′ and Im–H1′.

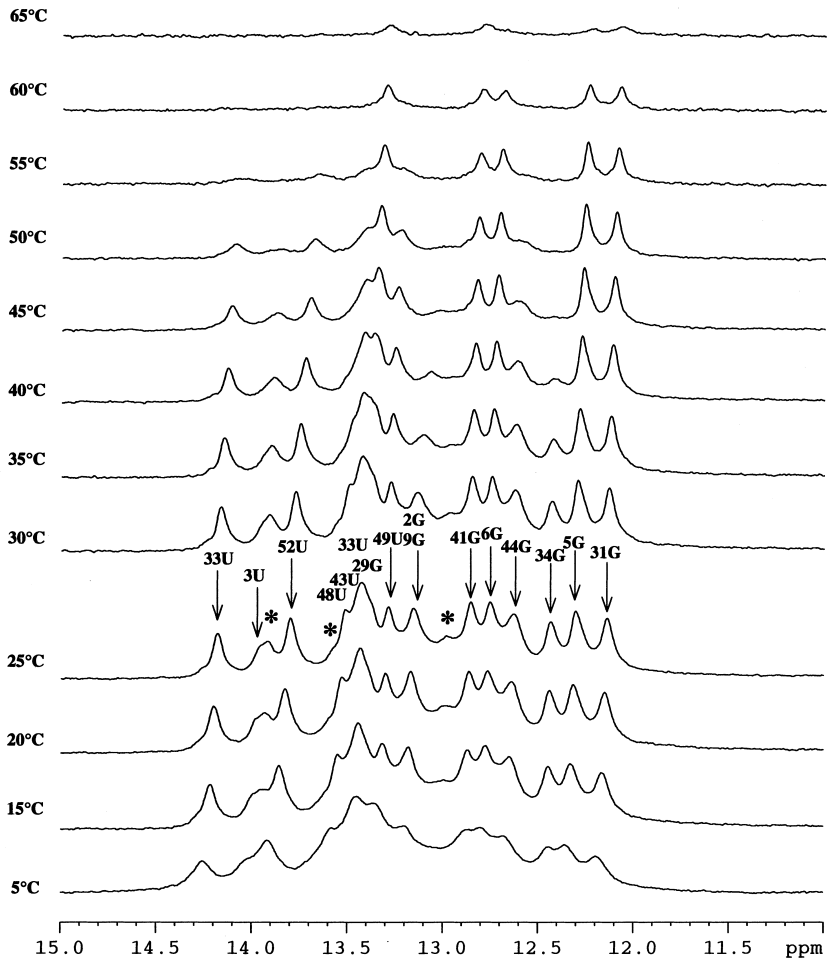
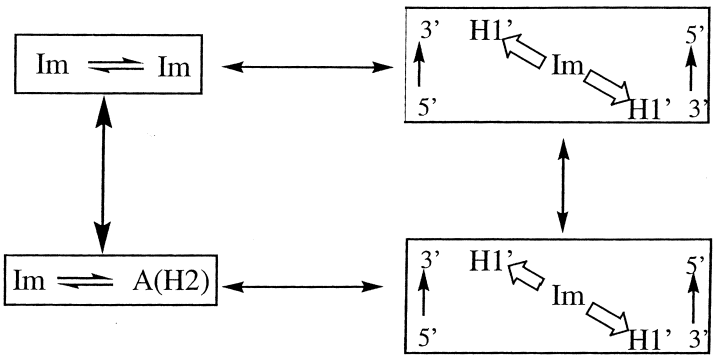


Fig. 4. One-dimensional ^{1}H spectra of the imino proton region from 15 to 11 ppm of the 55mer RNA at different temperatures: 5–65°C. The assignments of the imino proton resonances are presented on top of the spectra at 25°C. The red stars (*) show the resonances which have not been assigned.

The presence of these independent fingerprints of the A-RNA type structures for stems (A) and (B) (Fig. 2) is further illustrated below:

3.3.1. Imino–imino nOe crosspeaks

In Fig. 5A, the nOe crosspeaks between imino–imino protons of 55mer RNA are shown at 10°C. Note that the number of these crosspeaks does not increase with the decrease of temperature down to 0°C. Instead, some crosspeaks in fact have disappeared owing to the line broadening (Fig. 4). At temperatures around 25°C or higher, few crosspeaks (data not shown) as well as few resonances (Fig. 4) have disappeared because of the increase of the exchange rate between the imino protons and water.



Scheme 1.

Two sets (14 nonequivalent imino protons) of sequential imino–imino proton connectivities have been found in 14.5–12 ppm region in the NOESY spectra of 55mer RNA, which are shown by black and red lines in Fig. 5A. It is noteworthy that all of these imino proton resonances stretched over whole region between 14.5 and 12 ppm where the imino protons of both GC and AU basepairs normally appear. This allowed us at the outset to assume that these sequential connectivities apparently belong to two decameric helices (A) and (B) rather than to the stem (C) containing only AU basepair (Fig. 2).

3.3.2. Imino–H1′ nOe crosspeaks

In order to discriminate which of these (Fig. 5A) two sequential imino–imino pathways belong to which of the decameric helices (A) or (B), we have used additional information based on the nOe crosspeaks between imino and the H1′ proton of 3′-end nucleotide residue ($n + 1$) for both the same and the opposite strand. Solid lines in Fig. 5B connect these contacts.

Since we have unambiguously assigned the chemical shifts of non-deuterated ^{35}G(H1′) and ^{55}G(H1′) residues (around 5.9 ppm) through the H1′–H2′ crosspeaks in the DQF-COSY spectra (Fig. 3B), as the starting point of the assignment procedure, we can easily locate the imino protons of ^{34}G and ^{2}G nucleotides in the imino part of the NOESY spectra. Indeed, we have observed two strong H1′–Im with H1′ at 5.9 ppm (Fig. 5B). Since the imino protons of ^{34}G and ^{2}G (Fig. 2) can potentially show two crosspeaks with H1′ in the imino–H1′ area in the NOESY spectra (i.e., ^{34}G(1H) with ^{35}G(H1′) and ^{38}A(H1′); ^{2}G(1H) with ^{55}G(H1′) and ^{3}U(H1′)), two alternative assignments for the imino protons of ^{34}G(1H) and ^{2}G(1H) at ~12.5 ppm and ~13.2 ppm are possible.

The crosspeak at ~12.5 ppm clearly belongs to a single imino proton resonance. The intensity of this resonance reduces at 35°C (Fig. 4). Additionally, it shows only one crosspeak in NOESY spectra (Fig. 5A) with the imino proton of the neighbouring residue, indicating that it most probably belongs to one of the residues located toward the termini. This is furthermore evident from the fact that the imino protons of the

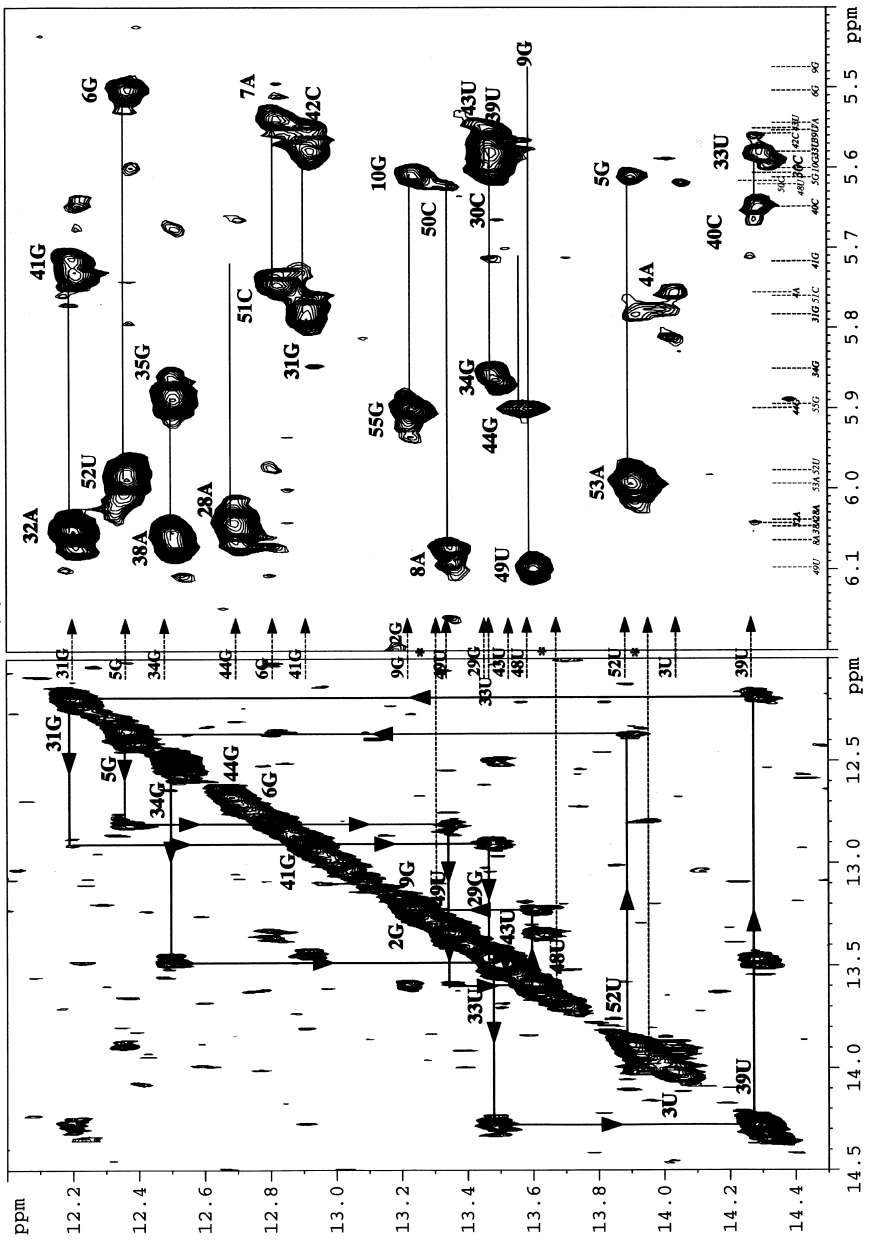

(A)
(B)
ppm
12.2
12.4
12.6
12.8
13.0
13.2
13.4
13.6
13.8
14.0
14.2
14.4
14.5
14.0
13.5
13.0
12.5
ppm
6.1
6.0
5.9
5.8
5.7
5.6
5.5
ppm
31G
5G
34G
44G
6G
41G
9G
2G
49U
29G
43U
48U
33U
52U
3U
39U
32A
52U
38A
28A
41G
6G
35G
51C
7A
42C
31G
10G
55G
8A
50C
34G
44G
49U
30C
43U
39U
9G
53A
5G
4A
40C
33U

Fig. 5.

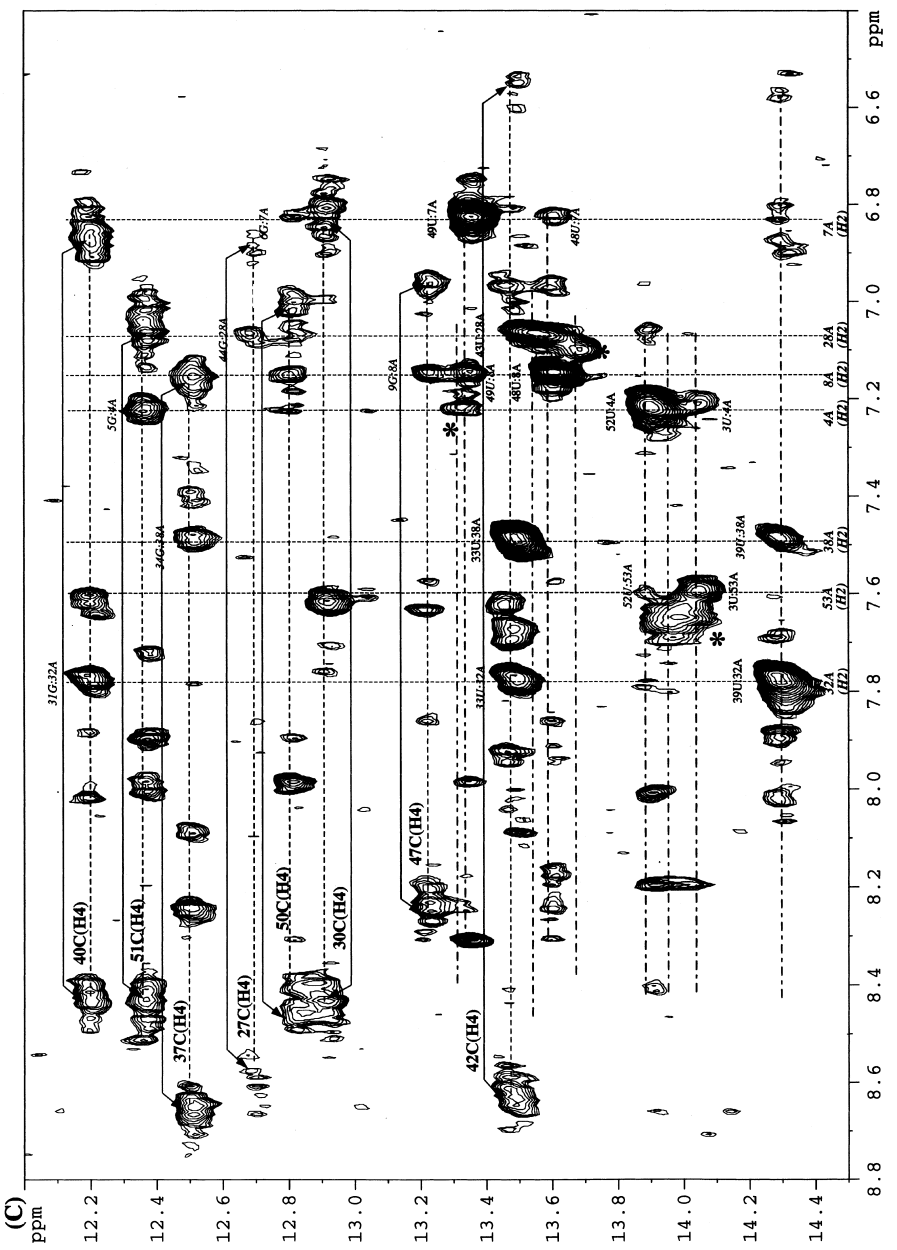


Fig. 5. (*continued*)

terminal bases, ^{35}G(1H) and ^{55}G(1H), exchange with solvent more rapidly long before the interior basepairs do.

Finally, the imino resonance at ~12.5 ppm is assigned to ^{34}G(1H) because the crosspeak between imino and H1′ (at ~12.5 ppm) is observed at very much more downfield (~6.05 ppm) than a similar imino/H1′ crosspeak at ~13.2 ppm (Fig. 5B). Furthermore, this assignment for ^{34}G(1H) allows us to sequentially assign the imino protons starting from ^{34}G(1H) up to the ^{43}U(3H) (Fig. 5A), which is matched with all G(1H) located between 13.5 and 12 ppm and U(3H) located between 15 and 13 ppm as generally expected [30]. This assignment is also corroborated by computer modelling of the decamer helices (A) and (B) basing on the observation that almost all imino protons show typical A-type [10,16] sequential imino–H1′ ($n + 1$) traces as presented in Fig. 5B. According to this model, there is an essential difference in the distances between different imino–imino protons for the 5′UA [(a) in Fig. 2], 5′AA [(b) in Fig. 2] and 5′AU tracts [(c) in Fig. 2]. The model has predicted the nOe sequential contacts (< 5 Å between imino–imino protons) for the decamer helix (B) (i.e., strands III and IV) starting from ^{34}G(1H) through ^{33}U(3H) and ^{39}U(3H) of 5′AU tract up to ^{43}U(3H) with disruption between ^{43}U(3H) and ^{44}G(1H) bases. For the second decameric helix (A) (i.e., strands I and II) the nOe sequential contacts are expected to start from ^{2}G(1H) to ^{3}U(3H), followed by interruption at the 5′UA tract between ^{3}U(3H) and ^{52}U(3H) (> 5.7 Å), and then regaining the nOe sequential contacts up to ^{10}G(1H). The expected nOe pattern of helix (B) fits nicely with the observed nOe patterns shown in Fig. 5A (black line). Noteworthy that the imino–imino crosspeak assigned to ^{34}G(1H)–^{33}U(3H) (Fig. 5A) is not observed with more dilute sample (2.5 times) but imino–imino ^{33}U(3H)–^{39}U(3H) is still very intense, which again fitted well with the A-type model for helix (B).

The spin diffusion crosspeak between imino and H5 of neighbouring C originating from GC basepair, H5↔H6↔4-NH_2↔N1H (Fig. 1), is commonly observed for small or medium sized A-type RNA structures (such as 21mer RNA [10] and 31mer RNA [12]). Unfortunately, this approach is less sensitive for an RNA molecule as large as 55mer because the H5 proton has a shorter relaxation time due to strong nOe with H6 (as observed in DQF-COSY spectra). This kind of spin diffusion nOe may arise for the central four GC basepairs, ^{26}G–^{45}C, ^{44}G–^{27}C, ^{9}G–^{47}C, ^{10}G–^{46}C, in the 55mer RNA

Fig. 5. NOESY spectra of 55mer RNA in H_2O/D_2O at 10°C. (Note the clear sharpness of crosspeaks owing to suppression of the spin diffusion in the deuterated sugar moieties!). In panel (A) the imino–imino proton region is presented. Two sequential connectivities have been found and labelled by red and black lines. The assignments of the imino proton resonances are shown on top diagonal peaks and on the right side of the panel with an arrow, sequential number and capital letter. In panel (B) the imino–H1′ proton region is presented. Two crosspeaks between imino and H1′ ($n + 1$) for the same and the opposite strands are connected by the solid line. The assignments of the H1′ protons are shown near the crosspeaks by number and capital letter. In panel (C) the imino–aromatic proton region is presented. Cross-section line through imino proton in F2 dimension is shown by a dashed black line for G(1H) imino protons and by a dashed blue line with dots for U(3H) imino protons. The amino protons of C are shown by a black solid line with arrows and labelled with number and capital letter. The A(H2) protons labelled at the bottom of the panel and cross-sections through H2 chemical shift in F1 dimension are shown by dashed red lines. The crosspeaks between imino (G(1H) or U(3H) with A(H2) are labelled near the crosspeaks. Three unassigned crosspeaks which potentially could belong to U(3H)–A(H2) of stem (C) labelled by the red symbol (*).

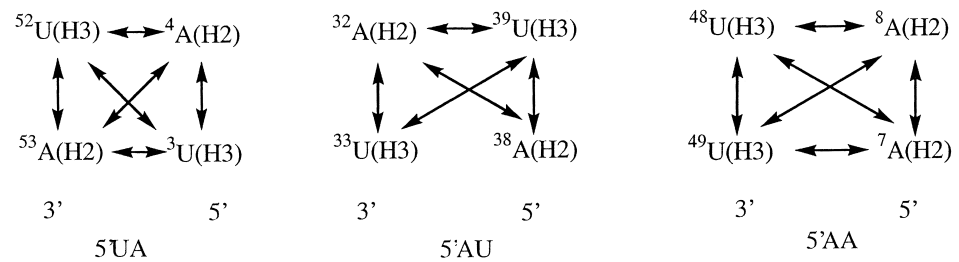
because the H5 protons for all other C residues outside the NMR-window have been replaced by deuterium. Nevertheless, the only resonance at 12.75 ppm, which does not show any interaction with any other imino protons, shows the nOe crosspeaks with H5 of neighbouring C. Additionally, the resonance at 12.75 ppm starts to broaden only above 50°C, whereas the imino resonance for ^{34}G at ~12.5 ppm exchanges completely at ~40°C, suggesting a higher stability of the former imino proton compared to the latter because of its relatively poor accessibility to the bulk water. This limited the number of candidates for assignment to two basepairs: ^{44}G–^{27}C and ^{9}G–^{47}C. Finally, the resonance at 12.75 ppm has been assigned to ^{44}G(1H), near the putative three-way junction, because of the crosspeak between ^{44}G(1H) and ^{28}A(H2) (Fig. 5C) (see below).

The assignment of a second sequential imino pathway (Fig. 5A, red line) is not clearly straightforward as for stem (B) and could not be performed only on Im–Im and Im–H1′ nOe. It needs additional information. In Fig. 5A (red line) the assignment of protons belonging to helix (A) (i.e., strands I and II) is shown to start from ^{3}U(3H) (~14 ppm) to ^{9}G(1H), which has been confirmed through assignment of A(H2) as well as by aromatic(Ar)–H1′ connectivity (see below).

3.3.3. Imino and adenine H2 protons

The trouble in detecting the H5–Im crosspeaks for large RNA makes the assignment of A(H2) protons especially important for two reasons: (i) H2A protons show strong nOe with both imino protons, and (ii) with H1′ of the 3′-end nucleotide ($n + 1$) of both the same strand as well as with that of the opposite strand (shown in Fig. 1 by red arrows). This is particularly important in partially deuterated residues where C1′ still has a proton but all other protons in the sugar moiety are replaced by deuterium. Owing to this, the crosspeaks between A(H2) and H1′ are very strong in NOESY spectra because of elimination of the extra relaxation pathway of H1′ [1,41]. These are readily observable in conventional NOESY spectra or could be filtered out from other crosspeaks in Ar–H1′ area through a relaxation filter in NOESY-type experiment.

First of all we go through A(H2)–Im interaction: according to the secondary structure of 55mer RNA (Fig. 2), seven U(H3)–A(H2) peaks are expected for two decamer helices (A) and (B). Six A(H2) protons belong to the 5′AU, 5′UA and 5′AA tracts and are expected to show interaction between imino proton of their own basepair and with the imino proton of the next basepair. This means that each tract should be represented by four crosspeaks as shown in Scheme 2.



Scheme 2.

Fig. 5C shows the NOESY spectra of 55mer RNA at 300 ms mixing time at 10°C in the area where imino–aromatic/amino proton crosspeaks are presented. As expected, the crosspeaks between imino U(H3) and A(H2) are the strongest even at 25°C and at a shorter mixing time (100 ms). Even if the distances between A(H2)–A(H2) or U(H3)–U(H3) protons are estimated to be as large as 5 Å (as was estimated from the model of 5′UA tract), four crosspeaks still should be observable (Fig. 5C) due to the short distances between A(H2) and U(H3) stacked basepairs. It was also possible to distinguish the 5′AA tract from the 5′AU and 5′UA tracts through the comparison of the intensities of nOe crosspeaks between ^{7}A(H2)↔^{8}A(H2) in the aromatic–aromatic region of the NOESY spectrum. The nOe crosspeak ^{7}A(H2)↔^{8}A(H2) is very strong compared with those of ^{32}A(H2)↔^{38}A(H2) and ^{4}A(H2)↔^{53}A(H2). Moreover, ^{53}A(H2), ^{4}A(H2), ^{32}A(H2), ^{38}A(H2), ^{7}A(H2), ^{8}A(H2) protons exhibit crosspeaks with the neighbouring imino protons of the GC basepairs ^{2}G(H1), ^{5}G(H1), ^{31}G(H1), ^{34}G(H1), ^{6}G(H1), ^{9}G(H1), respectively (Fig. 5C). This confirms the sequential Im–Im proton connectivities as established above. It also proves that the chemical shifts of ^{9}G(H1) and ^{2}G(H1) are similar (Fig. 5A,C).

The ^{28}AH(2) has no crosspeak with any of the U(H3) imino protons except with its own partner ^{43}U; it has, however, two crosspeaks (Fig. 5C) with two neighbouring imino protons of GC basepair, ^{44}G(H1) and ^{29}G(H1), thereby confirming the assignment of the ^{44}G(H1) imino proton, as discussed above.

The analysis of NOESY spectra in imino–imino, imino–A(H2) and imino–H1′ regions (Fig. 5) shows that there is no obvious resonance which could be assigned to the imino proton of U(H3) of the stem (C) of 55mer RNA or to ^{26}G(H1) and ^{10}G(H1) at all studied temperatures (0–30°C). We have found only three residual crosspeaks between imino–A(H2) (Fig. 5C, labelled by red stars) which possibly belong to the AU basepair of stem (C), but in NOESY spectra in D_2O these candidates for A(H2) do not show any interaction with H1′ as expected because of the absence of imino protons under our experimental conditions (Fig. 1).

The nOes from G(H1)/U(H3) imino protons to the 3′-neighbouring H1′ in the same strand and the H1′ across the opposite strand in the 3′ direction exactly correspond to those expected [10,16] for an A-type helical geometry (Fig. 1). For all assigned imino protons belonging to the stems (A) and (B) there are two crosspeaks between imino proton and H1′ protons (Fig. 5B). This allowed us to identify the chemical shifts of at least 28 H1′ resonances (Fig. 5B, at bottom of panel). Additionally, the nOes from A(H2) to the 3′-neighbouring H1′ on the same strand and the H1′ from the opposite strand in the 3′ direction (Fig. 6) gave 14 H1′ resonances. It is also noteworthy that at mixing time of 300 ms, A(H2) also shows week nOe with H1′ of its own residue providing information for stems (A) and (B), it also helped to assign seven additional H1′ chemical shifts. These two unambiguous data sets of H1′ obtained on the basis of assignment of imino as well as A(H2) resonances presented a final opportunity for crosschecking according to Scheme 1. The unambiguous assignment of 17 H1′ resonances of 17 nucleotides are as follows from the total of 40 H1′ protons of helix (A) and (B) (Fig. 2): (1) ^{38}A(H1′) both from ^{34}G(H1)–^{38}A(H1′) and weak ^{38}A(H2)–^{38}A(H1′); (2) ^{34}G(H1′) both from ^{33}U(H3)–^{34}G(H1′) and ^{38}A(H2)–^{34}G(H1′); (3)

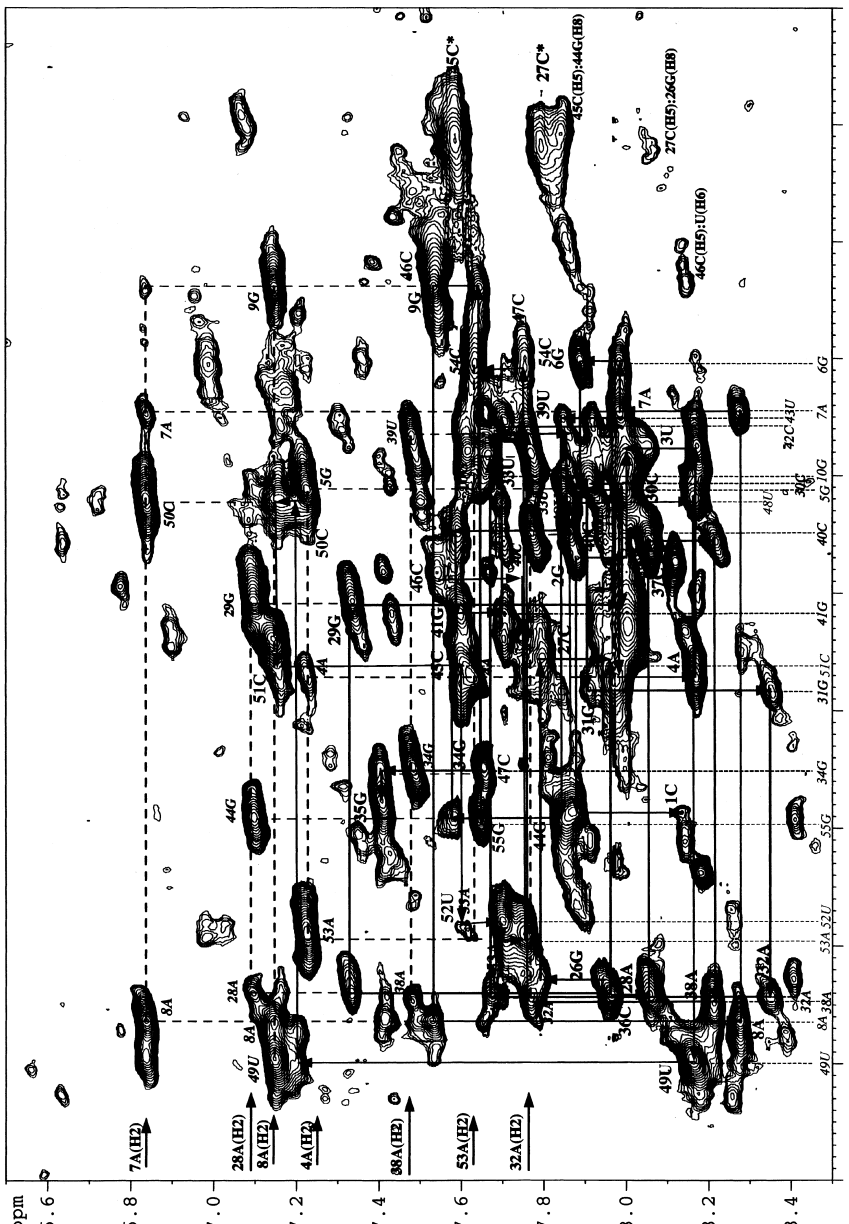


Fig. 6. NOESY spectra of 55mer RNA in D_2O at 30°C in aromatic–H1′ region. Four sequential connectivities, 5′-(H6/H8)$_i$↔(H1′)$_i$↔(H6/H8)$_{i+1}$, have been found for helices (A) (strand (I) red line strand (II) pink line) and helices (B) (strand (III) dark blue line and strand (IV) light blue line). The assignments of the (H6/H8)$_i$–(H1′)$_i$ crosspeaks are shown at the top by sequential number and capital letter. The cross-section in F2 dimension through A(H2) protons is labelled by a green dashed line. The assignments of A(H2) protons are presented near these green lines with a green arrow and sequential number and capital letter from the left side of panel. The bottom part of the panel shows the assigned H1′ protons which have been confirmed in the imino–H1′ region of the NOESY spectra (see Fig. 5B). The crosspeaks between 5′-C(H5) $_i$–(H6/H8)$_{i-1}$ are labelled in black.

^{39}U(H1′) both from ^{38}A(H2)–^{39}U(H1′) and ^{33}U(H3)–^{39}U(H1′); (4) ^{33}U(H1′) both from ^{32}A(H2)–^{33}U(H1′) and ^{39}U(H3)–^{33}U(H1′); (5) ^{40}C(H1′) both from ^{32}A(H2)–^{40}C(H1′) and ^{39}U(H3)–^{40}C(H1′); (6) ^{32}A(H1′) both from ^{31}G(H1)–^{32}A(H1′) and weak ^{32}A(H2)–^{32}A(H1′); (7) ^{28}A(H1′) both from ^{44}G(H1)–^{28}A(H1′) and weak ^{28}A(H2)–^{28}A(H1′); (8) ^{44}G(H1′) both from ^{43}U(H3)–^{44}G(H1′) and ^{28}A(H2)–^{44}G(H1′); (9) ^{29}G(H1′) both from ^{43}U(H3)–^{29}G(H1′) and ^{28}A(H2)–^{29}G(H1′); (10) ^{9}G(H1′) both from ^{48}U(H3)–^{9}G(H1′) and ^{8}A(H2)–^{9}G(H1′); (11) ^{49}U(H1′) both from ^{8}A(H2)–^{49}U(H1′) and ^{48}U(H3)–^{49}U(H1′); (12) ^{8}A(H1′) three data set ^{49}U(H3)–^{32}A(H1′), ^{7}A(H2)–^{8}A(H1′) and weak ^{8}A(H2)–^{8}A(H1′); (13) ^{7}A(H1′) both set ^{6}G(H1)–^{32}A(H1′) and weak ^{7}A(H2)–^{7}A(H1′); (14) ^{50}C(H1′) both from ^{49}U(H3)–^{50}C(H1′) and ^{7}A(H2)–^{50}C(H1′); (15) ^{9}G(H1′) both from ^{52}U(H3)–^{5}G(H1′) and ^{4}A(H2)–^{5}G(H1′); (16) ^{4}A(H1′) three data set ^{3}U(H3)–^{4}A(H1′), ^{53}A(H2)–^{4}A(H1′) and weak ^{4}A(H2)–^{4}A(H1′); (17) ^{53}A(H1′) three data set ^{52}U(H3)–^{53}A(H1′), ^{4}A(H2)–^{53}A(H1′) and weak ^{53}A(H2)–^{53}A(H1′); (18) ^{54}C(H1′) both from ^{3}U(H3)–^{54}C(H1′) and ^{53}A(H2)–^{54}C(H1′). Based on these data, the assignment of the rest of the crosspeaks from nOe imino and A(H2) with H1′ gave four more H1′ chemical shifts of the following residues: ^{35}G(H1′), ^{41}G(H1′), ^{45}C(H1′), ^{51}C(H1′). Additionally, there are pairs of H1′, whose chemical shifts are known from nOe with imino protons, but they have two alternative assignments which can be assigned by crosschecking Ar–H1′ nOe crosspeaks (such as ^{30}C(H1′) or ^{43}U(H1′); ^{10}G(H1′) or ^{48}U(H1′); ^{6}G(H1′) or ^{52}U(H1′); ^{3}U(H1′) or ^{55}G(H1′); ^{31}G(H1′) or ^{42}C(H1′)). From all 40 H1′ resonances of (A) and (B) stems, the H1′ chemical shifts of only few residues, however, could not be assigned based on the nOes from imino and A(H2) resonances only. These include (i) two terminal residues from both 5′ end ^{1}C(H1′), ^{2}G(H1′) and ^{37}C(H1′), ^{36}C(H1′) and (ii) four residues in the NMR-window part of 55mer RNA, ^{26}G(H1′), ^{27}C(H1′), ^{46}C(H1′), ^{47}C(H1′). The labels for all assigned H1′ protons are presented in Figs. 5B and 6.

3.4. Assignment based on non-exchangeable H1′ and aromatic protons

To confirm the assignment of H1′ as well as to make firm assignment of the aromatic protons, the sequential connectivity of 5′ $(H6/H8)_i \leftrightarrow (H1')_i \leftrightarrow (H6/H8)_{i+1}$ is required in order to be internally consistent with all data presented above. The overlap of the crosspeaks in the aromatic–H1′/H5 area of the NOESY spectrum of the 55mer RNA (Fig. 6) is very severe: the A(H2) resonances are severely superimposed with H6/H8 in the aromatic region. Compared with the B-type of DNA helix where the nOe crosspeaks of A(H2) to H1′ are very weak [5] for the A-type RNA helix, they have middle to strong intensity. Even for stems (A) and (B) for 55mer RNA, the number of aromatic resonances according to the primary sequence (Fig. 2) are expected to be 23 for helix A and 24 for helix B. The A(H2) can be distinguished from H6/H8 resonances by absence of nOe (or very weak) with other H2′, H3′, H4′ sugar protons or by relaxation filtering experiment. For a partially deuterated residue, there is a single crosspeak between H8/H6–H1′ and H4′ with narrowing line because of the absence of a spin diffusion pathway through H2′ and H3′. This phenomenon has been observed by us [1,41] for 20mer DNA (B-type of conformation of helix), where a stretch of 12 basepairs in the

NMR-window part were left fully protonated but other nucleotide residues have partially deuterated sugar residues.

The sequential assignment in the H8/H6↔H1′ region (Fig. 6) of the 55mer RNA starts from the non-deuterated nucleotide at the 3′-end of strands (II) and (III). This is because of the fact that the H1′ proton at the 3′-end was identified (i) by its lack of a nOe crosspeak to a 3′ neighbouring base proton [37,38] and, most importantly, (ii) by the presence of the strong H1′–H2′ crosspeaks in DQF-COSY spectra (Fig. 3B). The chemical shifts of aromatic H8 of ^{35}G (H8) and ^{55}G (H8) can be easily identified through analysis of nOes between H8 with H2′, H3, and H4′. It is also known [37,38] that the sugar protons at the 3′-end have very much up-field shifts because of the absence of the 3′-phosphate moiety.

The sequential assignment paths from ^{35}G (H8, H1′) to ^{26}G (H8, H1′) (Fig. 6, dark blue line) and ^{55}G (H8, H1′) to ^{46}G (H8, H1′) (Fig. 6, pink line) have been fully accomplished in this work. The aromatic H6 of C in the NMR window is identified from the H5–H6 crosspeak in DQF-COSY spectra (Fig. 3A). The chemical shifts of H1′ of ^{26}G(H1′), ^{45}C(H1′), ^{46}C(H1′), ^{47}C(H1′) were confirmed through the H1′–H2′ crosspeaks (Fig. 3B). This is due to the larger contribution of the South-type conformation (i.e., $J_{1',2'} \approx 7$–8 Hz) [37,38], making the H1′–H2′ crosspeaks visible in the NMR-window. The flexible nature of nucleotides in the NMR-window part is also evident from the absence of imino protons which are involved in quick exchange with bulk water.

The sequential paths for strand (I) from ^{1}C(H6, H1′) to ^{10}G(H8, H1′) (Fig. 6, light blue line) and ^{36}C(H6, H1′) to ^{45}C(H6, H1′) in strand (IV) (Fig. 6, red line) cannot be fully accomplished without additional information such as the knowledge of the H1′ resonances obtained from the studies of exchangeable protons. This is because the starting point of assignment at the 5′-end was difficult to recognise because of overlapping crosspeaks in the aromatic and H1′ areas. The ^{45}C(H6) in the NMR-window part was identified from H5–H6 crosspeaks in DQF-COSY spectra (Fig. 3A), which were then used as the starting points for further sequential assignment.

3.5. The non-exchangeable protons of stem (C) in D_2O

After assignment of all aromatic and H1′ protons of stems (A) and (B), we found a considerable number of nOe crosspeaks in the aromatic–H1′ area which were unaccounted and potentially could belong to the H2/H8/H6 and H1′ protons of the stem (C) (Fig. 6). As has been shown above from the analysis of imino protons, there is no evidence to suggest that this stem has stable AU basepairs. Nevertheless, a partial sequential $5'(H6/H8)_i \leftrightarrow (H1')_i \leftrightarrow (H6/H8)_{i+1}$ connectivity can be established. There is also clear indication of the presence of A(H2) crosspeaks with H1′ which can possibly belong to AU tract of stem (C). It is noteworthy that these crosspeaks have much lower intensity than the crosspeaks belonging to the helices (A) and (B). Moreover some of them are as sharp as the corresponding peaks from helices (A) and (B), but some of them have serious line broadening. This shows that the residues in stem (C) experience different mobilities and conformational exchange process.

Some useful information on the conformation adopted by junction could come from

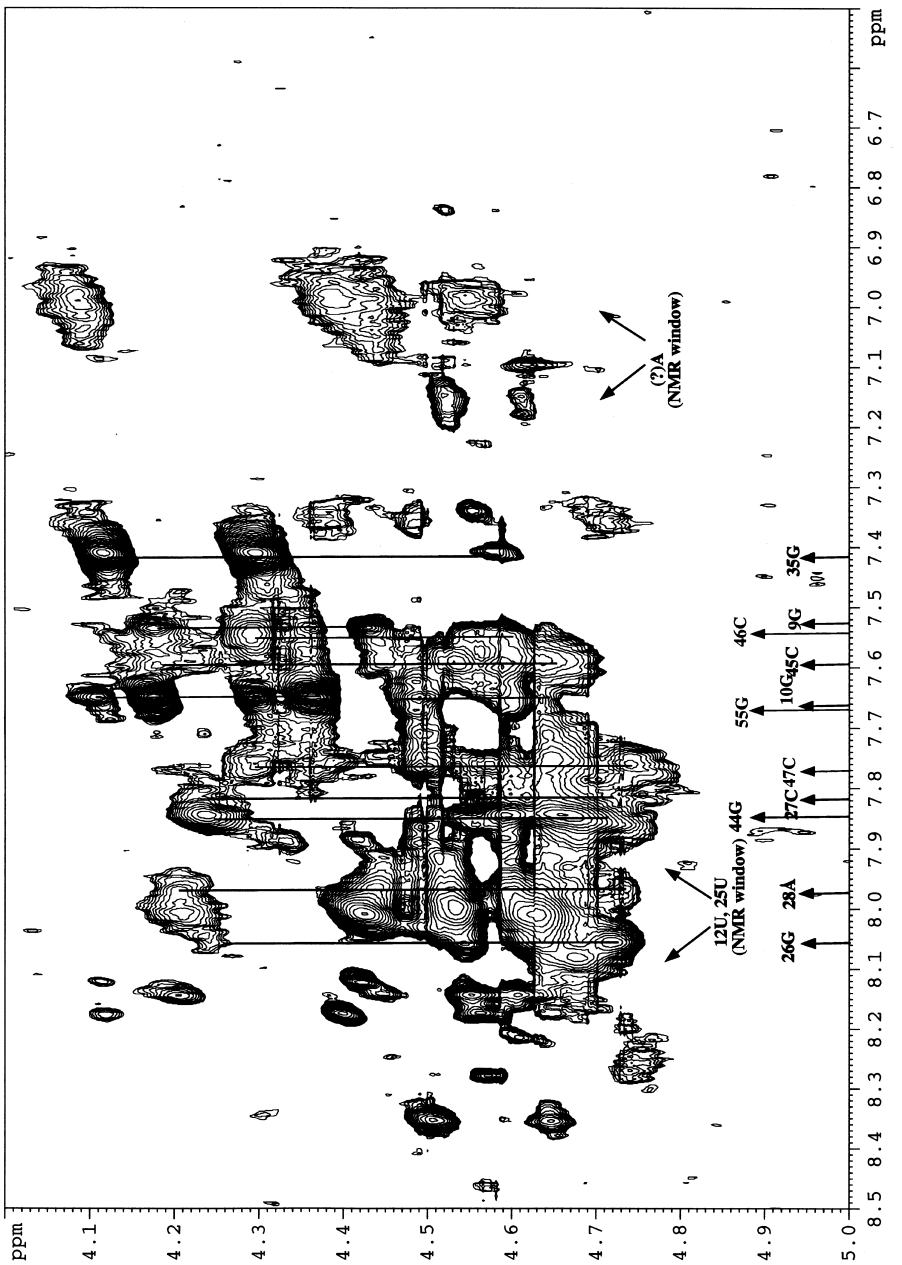


Fig. 7. NOESY spectra of aromatic–H2′/H3′/H4′ region (from the NMR-window part) of 55mer RNA in D_2O at 30°C. The cross-sections in F1 dimension through aromatic protons belong to the nucleotide residues of the NMR window, which are labelled by arrow with colour as well as by sequential number and capital letter from the bottom side of the panel, and used for sequential assignment in Fig. 5. The crosspeaks belonging to the residues of the AU tract in the NMR window region of stem (C) are presented by black arrows.

the crosspeaks between $C(H5)_i \leftrightarrow H8/H6_{i-1}$ such as $^{45}C(H5)$–$^{44}G(H8)$, $^{27}C(H5)$–$^{26}G(H8)$: there is no indication of nOe crosspeak between $^{46}C(H5)$ and $^{45}G(H6)$. Instead the crosspeak between $^{46}C(H5)$ and aromatic H6 of U *not belonging* to the NMR-window part is observed (Fig. 6). Moreover, there are two sets of crosspeaks between $^{10}C(H8)$ and $^{11}A/^{12}U$ (both of them are fully protonated). *These data suggest that 55mer RNA does not adopt a tight junction with three helices, instead it adopts only two stems*.

3.6. Non-exchangeable H2′, H3′, H4′ protons obtained from NOESY spectra of 55mer RNA in D_2O

Protonated residues in the NMR-window located near the helical three-way junction are shown in the square box in Fig. 2. For all other partially deuterated residues, H2 and H3′ protons are replaced by deuterium and hence are not observable in the aromatic–H2′/H3′ (Fig. 7) and H1′–H2′, H3′ areas. Following are our conclusions from the analysis of the expended part of the NOESY spectra.

1. Various nucleotide residues within the NMR-window part have various mobilities and conformational exchange processes. For example, the H8–H2′, H3′ crosspeaks of ^{9}G are very intense and sharp but the same crosspeaks for ^{44}G are more broadened.
2. There are two sets of very broadened crosspeaks, which belong to the NMR-window residues, but they do not show any nOe contact with other NMR-window residues. These crosspeaks are not assignable in this study. The present investigation, however, suggests that these residues are under slow conformational exchange and are not involved in any stacking interaction.
3. The severe line broadening of crosspeaks belonging to the protons in the NMR-window suggests that there is considerably strong spin diffusion between all sugar protons in the fully protonated residues. Although these crosspeaks have been used for assignment but the elimination of spin diffusion for residues belonging to the NMR-window is desirable in order to perform any quantitative extraction of data.
4. As mentioned above, H1′ and H4′ protons were left partially undeuterated in the deuterated nucleotide residues outside the NMR-window area. Thus the H1′ protons of these deuterated residues are no longer involved in the spin diffusion pathway because of elimination of H2′ and H3′. Hence, they show sharper linewidth compared with the H1′ of the fully protonated residues. This useful property has been used to prove the A-type RNA character in helices (A) and (B) (see above). Nevertheless, the partially deuterated C4′ (i.e., residual H4′) in a large RNA, as our 55mer, essentially complicate spectra by overlapping with the resonances of the residues of the non-deuterated NMR-window part in the aromatic–H2′, H3′, H4′, H5′, H5″ (Fig. 7) and specially in the H1′–H2′, H3′, H4′, H5′, H5″ areas of the spectrum (data not shown). Hence, a new synthetic approach has been undertaken to synthesize building blocks incorporating full substitution at C4′ with deuterium.

The chemical shifts of H1′, H4′, H8/H6, H2, imino protons are presented in Table 1.

Table 1
^{1}H NMR chemical shifts of non-exchangeable protons at 30°C and exchangeable protons at 10°C

Stem	Strand	Residue	H8/H6	H5/H2	NH	NH_2 [e]	H1′	H2′	H3′
		1C	8.14	d	b	c	5.89	d	d
		2G	7.89	b	12.22	b	5.67	d	d
		3U	8.01	d	14.00	b	5.58	d	d
Helix	(I)	4A	8.17	7.23	b	b	5.77	d	d
(A)		5G	7.98	b	12.35	b	5.60	d	d
		6G	7.89	b	12.82	b	5.50	d	d
		7A	7.99	6.06	b	b	5.55	d	d
		8A	8.28	6.10	b	b	6.06	d	d
		9G	7.54	b	13.22	b	5.44	4.43	4.18
		10G	7.65	b	a	b	5.61	4.52	4.60
		11A	a	a	b	b	a	cr	cr
		12U	7.98	5.77	a	b	5.93*	cr	cr
		13A	a	a	b	b	a	d	d
		14U	a	d	a	b	a	d	d
		15A	a	a	b	b	a	d	d
		16U	a	d	a	b	a	d	d
		17A	a	a	b	b	a	d	d
Helix		18U	a	d	a	b	a	d	d
(C)		19U	a	d	a	b	a	d	d
		20A	a	a	b	b	a	d	d
		21U	a	d	a	b	a	d	d
		22A	a	a	b	b	a	d	d
		23U	a	d	a	b	a	d	d
		24A	a	a	b	b		cr	cr
		25U	8.01	5.73	a	b	5.59*	cr	cr
		26G	8.06	–	c	b	6.03	4.87	4.66
		27C	7.79	5.32	b	8.57/6.90	5.75	4.54	4.64
Helix		28A	7.97	7.10	b	b	6.04	d	d
(B)	(III)	29G	7.34	b	13.45	b	5.71	d	d
		30C	8.01	d	b	8.43/6.84	5.61	d	d
		31G	7.91	b	12.20	b	5.78	d	d
		32A	8.35	7.77	b	b	6.05	d	d
		33U	7.67	d	13.45	b	5.58	d	d
		34G	7.65	b	12.50	b	5.85	d	d
		35G	7.41	b	c	b	5.88	4.10	4.29
		36C	7.97	d	b	c	6.04	d	d
		37C	8.06	d	b	8.64/7.15	5.66	d	d
		38A	8.22	7.48	b	b	6.05	d	d
Helix		39U	7.77	d	14.27	b	5.57	d	d
(B)	(IV)	40C	7.87	d	b	8.42/6.88	5.65	d	d
		41G	7.60	b	12.90	b	5.72	d	d
		42C	7.71	d	b	8.62/6.58	5.56	d	d
		43U	7.99	d	13.50	b	5.55	d	d
		44G	7.85	b	12.67	b	5.89	4.59	4.66
		45C	7.59	5.31	b	c	5.78	4.51	4.65
		46C	7.55	5.43	b	c	5.68	4.29	4.54
		47C	7.75	5.50	b	8.23/6.97	5.85	4.63	4.74
		48U	7.86	d	13.59	b	5.62	d	d

Table 1. Continued

Stem	Strand	Residue	H8/H6	H5/H2	NH	NH_2[e]	H1′	H2′	H3′
Helix		49U	8.17	d	13.34	b	6.09	d	d
(A)	(II)	50C	7.20	d	b	8.43/6.99	5.62	d	d
		51C	7.15	d	b	8.42/7.03	5.76	d	d
		52U	7.61	d	13.89	b	5.98	d	d
		53A	7.70	7.64	b	b	5.99	d	d
		54C	7.76	d	b	a	5.51	d	d
		55G	7.65	b	c	b	5.89	4.17	4.36

a, chemical shift could not defined; b, proton does not exist; c, the imino and amino protons of the terminal residue, which are in quick exchange with water; d, proton is replaced by deuterium; e, two amino protons of cytosine.

Despite the large size of the helices (40 nucleotides, 20 basepairs) the chemical shift of H1′ of both helices (A) and (B) are accurately defined for most of the residues (labelled in Table 1). Nevertheless the chemical shifts of some aromatic protons are ambiguous because of the overcrowded crosspeaks in the aromatic and H1′ area. The information from crosspeaks located in the aromatic–aromatic area did not improve the possibility to assign the aromatic protons in a dramatic manner because of isochronous or nearly isochronous chemical shifts and very weak nOe crosspeaks, which is also true for H1′–H1′ or H5 crosspeaks.

4. Conclusion

In deuterated residues the spin diffusion of H1′ proton is restricted owing to the substitution of H2′ and H3′ protons by deuterium. This useful property of deuterated nucleotides has earlier been used by us in solving the structure of selectively deuterated 21mer [10] and 31mer [12] RNA and 20mer DNA [1,41]. In this study, we have shown that for deuterated residues, the A(H2)–H1′, Im–H1′ and Im–Im crosspeaks in NOESY, together with H1′–H2′ crosspeaks for the terminal residue in DQF-COSY, seem to provide adequate information in order to be able to identify the A-type RNA helices unambiguously within a large folded RNA. This approach has been illustrated in this paper on 55mer RNA which is folded [40] in a three-way junction. To the best of our knowledge, the size of the RNA dealt with in the present work is still one of the largest which is possible to deal with by solution NMR spectroscopy in conjunction with a non-uniform deuterium-labelling technique. To tackle the conformation of a large RNA as used in the present study, the uniform ^{13}C labelling of all nucleotide residues will most probably not help very much to overcome the trouble of poor dispersion of the $^{13}C(H1')$ resonances. One way to overcome this problem is to label the sugar moiety by ^{13}C in the NMR-window part (i.e., non-uniform labelling) strategically placed at chosen sites. Clearly, the information about helices presented in 55mer RNA will allow us to optimise our choice of the biologically interesting site for selective incorporation of isotopes.

Acknowledgements

The authors thank Professor H. Seliger and Dr P. Weiler (Sektion Polymere, Universität Ulm, Albert-Einstein-Allee 11, D-89081 Ulm, Germany) for collaborative preparation of the RNA samples in the Uppsala lab. Thanks are due to the Swedish Board for Technical Development (NUTEK) to JC, the Swedish Natural Science Research Council (NFR contract # K-KU 12067-300 to AF and K-AA/Ku04626-321 to JC) and the Swedish Research Council for Engineering Sciences (TFR) to JC for generous financial support. The authors also thank Wallenbergsstiftelsen, University of Uppsala and Forskningsrådsnämnden (FRN) for their financial supports for the purchase of the 500 and 600 MHz NMR spectrometers.

References

[1] Agback P, Maltseva TV, Yamakage S-I, Nilsson FPR, Földesi A, Chattopadhyaya J. The differences in the T_2 relaxation rates of the protons in the partially deuterated and fully protonated sugar residues in a large oligo-DNA ('NMR-window') gives complementary structural information. Nucleic Acids Res 1994;22:1404–12.

[2] Allain FH-T, Varani G. How accurately and precisely can RNA structure be determined by NMR? J Mol Biol 1997;267:338–51.

[3] Batey RT, Williamson JR. Effects of polyvalent cations on the folding of an rRNA three-way junction and binding of ribosomal protein S15. RNA 1998;4:984–97.

[4] Chastain M, Tinoco Jr. I. Structural elements in RNA. Prog Nucleic Acids Res Mol Biol 1991;41:131–77.

[5] Chou S-H, Flynn P, Reid B. Solid phase synthesis and high-resolution NMR studies of two synthetic double-helical RNA dodecamers: r(CGCGAAUUCGCG) and r(CGCGUAUACGCG). Biochemistry 1989;28:2422–35.

[6] Dahm SC, Uhlenbeck OC. Role of divalent metal ions in the hammerhead RNA cleavage reaction. Biochemistry 1991;30:9464–9.

[7] Dieckmann T, Suzuki E, Nakamura GK, Feigon J. Solution structure of an ATP-binding RNA aptamer reveals a novel fold. RNA 1996;2:628–40.

[8] Földesi A, Nilson FPR, Glemarec C, Gioeli C, Chattopadhyaya J. Synthesis of 1′$^{\#}$, 2′,3′,4′$^{\#}$,5′,5″-$^{2}H_6$-β-D-ribonucleosides and 1′$^{\#}$,2′,2′,3′,4′$^{\#}$,5′,5″-$^{2}H_7$-β-D-2′-deoxyribonucleosides for selective suppression of proton resonances in partially-deuterated oligo-DNA oligo-RNA and in 2,5A core (^{1}H-NMR window). Tetrahedron 1992;48:9033–72.

[9] Földesi A, Yamakage S-I, Maltseva TV, Nilsson FPR, Agback P, Chattopadhyaya J. Partially-deuterated nucleotide residues in large DNA duplex simplify the spectral overlap and provide both the J-coupling and nOe informations by the 'NMR-window' approach. Tetrahedron 1995;51:10065–92.

[10] Földesi A, Yamakage S-I, Nilsson FPR, Maltseva TV, Chattopadhyaya J. The use of non-uniform deuterium labelling ('NMR-window') to study the NMR structure of a 21mer RNA hairpin. Nucleic Acids Res 1996;24:1187–94.

[11] Földesi A, Maltseva TV, Dinya Z, Chattopadhyaya J. Synthesis of partially-deuterated 2′-deoxyribonucleoside blocks and their incorporation into an oligo-DNA for simplification of overcrowding and selective enhancement of resolution and sensitivity in the ^{1}H-NMR spectra. Tetrahedron 1998;54:14487–514.

[12] Glemarec C, Kufel J, Földesi A, Maltseva TV, Sandström A, Kirsebom LA, Chattopadhyaya J. The NMR structure of 31mer RNA domain of *Escherichia coli* Rnase P RNA using its non-uniformly deuterium labeled counterpart (the 'NMR-window' concept). Nucleic Acids Res 1996;24:2022–35.

[13] Greenbaum NL, Radhakrishnan I, Hirsh D, Patel DJ. Determination of the folding topology of the SL1 RNA from *Caenorhabditis elegans* by multidimensional heteronuclear NMR. J Mol Biol 1995;252:314–27.

[14] Grune M, Görlach M, Soskic V, Klussmann SM, Bald R, Furste JP, Erdmann VA, Brown LR. Initial analysis of 750 MHz NMR spectra of selectively ^{15}N-G,U-labelled *E. coli* 5S rRNA. FEBS Lett 1996;385:114–8.
[15] Heus HA, Pardi A. Nuclear magnetic resonance studies of the hammerhead ribozyme domain. Secondary structure formation and magnesium ion dependence. J Mol Biol 1991;217:113–24.
[16] Heus HA, Pardi A. Novel ^{1}H nuclear magnetic resonance assignment procedure for RNA duplexes. J Am Chem Soc 1991;113:4360–1.
[17] Jiang F, Fiang R, Live D, Kumar RA, Patel DJ. RNA folding topology and intermolecular contact in the AMP-RNA aptamer complex. Biochemistry 1996;35:13250–66.
[18] Jiang F, Kumar RA, Jones RA, Patel DJ. Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. Nature 1996;382:183–6.
[19] Jiang F, Patel DJ, Zhang X, Zhao H, Jones RA. Specific labeling approaches to guanine and adenine imino and amino protons in the AMP-RNA aptamer complex. J Biomol NMR 1997;9:55–62.
[20] Kellenbach E R, van den Elst H, Boelens R, van der Marel GA, van Boom JH, Kaptein R. A convenient synthesis of DNA fragments nitrogen-15 labelled at the exocyclic cytosine amino group. Recl Trav Chim Pays-Bas 1991;110:387–8.
[21] Kieft JS, Tinoco Jr. I. Solution structure of a metal-binding site in the major groove of RNA complexed with cobalt (III) hexamine. Structure 1997;5:713–21.
[22] Kolk M, van der Graaf M, Wijmenga SS, Pleij CWA, Heus HA, Hilbers CW. NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA. Science 1998;280:434–8.
[23] Kolk M, Wijmenga SS, Heus HA, Hilbers CW. On the NMR structure determination of a 44n RNA pseudoknot: assignment strategies and derivation of torsion angle restraints. J Biomol NMR 1998;12:423–33.
[24] Long DM, Uhlenbeck OC. Self-cleaving catalytic RNA. FASEB J 1993;7:25–30.
[25] Maltseva TV, Földesi A, Chattopadhyaya J. Partially-deuterated oligo-DNA reduces overcrowding, enhance resolution and sensitivity and provide improved NMR constraints for structure elucidation of oligo-DNA. Tetrahedron 1998;54:14515–28.
[26] Maltseva TV, Földesi A, Chattopadhyaya J. T_1 and T_2 relaxations of the ^{13}C nuclei of deuterium-labeled nucleosides. Magn Res Chem 1998;36:227–39.
[27] Maltseva TV, Földesi A, Chattopadhyaya J. Measurement of T_1 relaxation rates of natural abundant ^{13}C at C2′ in a non-uniformly deuterium-labelled oligo-DNA. J Chem Soc Perkin Trans 1998c;22:2689–94.
[28] Odai O, Kodama H, Hiroaki H, Sakata T, Tanaka T, Uesugi S. Synthesis and NMR study of ribooligonucleotides forming a hammerhead-type RNA enzyme system. Nucleic Acids Res 1990;18:5955–60.
[29] Reese CB, Ubasawa A. Reaction between 1-arenesulphonyl-3-nitro-1,2,4-triazoles and nucleoside base residues. Elucidation of the nature of side-reactions during oligonucleotide synthesis. Tetrahedron Lett 1980;21:2265–8.
[30] Sarma RH, Sarma MH, Rein R, Shibata M, Setlik RS, Ornstein RL, Kazin AL, Cairo A, Tomasi TB. Secondary structure in solution of two anti-HIV-1 hammerhead ribozymes as investigated by two-dimensional ^{1}H 500 MHz NMR spectroscopy in water. FEBS Lett 1995;357:317–23.
[31] Schmitz U, Freymann DM, James TL, Keenan RJ, Vinayak R, Walter P. NMR studies of the most conserved RNA domain of the mammalian signal recognition particle (SRP). RNA 1996;2:1213–27.
[32] Simorre J-P, Zimmermann GR, Mueller L, Pardi A. Triple-resonance experiments for assignment of adenine base resonances in ^{13}C/^{15}N-labeled RNA. J Am Chem Soc 1996;118:5316–7.
[33] Simorre J-P, Legault P, Hangar AB, Michiels P, Pardi A. A conformational change in the catalytic core of the hammerhead ribozyme upon cleavage of an RNA substrate. Biochemistry 1997;36:518–25.
[34] Simorre J-P, Legault P, Baidya N, Uhlenbeck OC, Maloney L, Wincott F, Usman N, Beigelman L, Pardi A. Structural variation induced by different nucleotides at the cleavage site of the hammerhead ribozyme. Biochemistry 1998;37:4034–44.
[35] Sklenar V, Dieckmann T, Butcher SE, Feigon J. Optimization of triple-resonance HCN experiments for application to large RNA oligonucleotides. J Magn Res 1998;130:119–24.
[36] Uhlenbeck OC. A small catalytic oligonucleotide. Nature 1987;328:596–600.
[37] Varani G, Tinoco Jr. I. RNA structure and NMR spectroscopy. Q Rev Biophys 1991;24:479–532.
[38] Varani G, Aboul-ela F, Allain FH-T. NMR investigation of RNA structure. Prog NMR Spectrosc 1996;29:51–127.

[39] Wallis NG, Dardel F, Blanquet S. Heteronuclear NMR studies of the interactions of ^{15}N-labeled methionine-specific transfer RNAs with methionyl-tRNA transformylase. Biochemistry 1995;34:7668–77.
[40] Weiler P. Ph.D thesis. Charakterisierung von komplexen RNA-strukturen: Übergänge Zwischen Internen Loopstrukturen und Verzweigungen. Ulm University, Germany, 1998.
[41] Yamakage S-I, Maltseva TV, Nilsson FPR, Földesi A, Chattopadhyaya J. Deuteration of sugar protons simplify NMR assignments and structure determination of large oligonucleotide by the ^{1}H-NMR window approach. Nucleic Acids Res 1993;21:5005–11.
[42] Zarrinkar PP, Williamson JR. The kinetic folding pathway of the *Tetrahymena ribozyme* reveals possible similarities between RNA and protein folding. Nature Struct Biol 1996;3:432–8.