Journal of the Korean Magnetic Resonance Society **2014**, *18,* 5-9
DOI 10.6564/JKMRS.2014.18.1.005

NMR methods for structural analysis of RNA: a Review

Nak-Kyoon Kim[1*], Yun-Sik Nam[1], and Kang-Bong Lee[1]

[1]Advanced Analysis Center, Korea Institute of Science and Technology, 39-1, Hawolgok-Dong, Seongbuk-Gu, Seoul, 136-791, Republic of Korea

Received May 1, 2014; Revised May 16, 2014; Accepted June 10, 2014

Abstract

In last three decades, RNA molecules have been revealed to play the central roles in many cellular processes. Structural understanding of RNA molecules is essential to interpret their functions, and many biophysical techniques have been adopted for this purpose. NMR spectroscopy is a powerful tool to study structures and dynamics of RNA molecules, and it has been further applied to study tertiary interactions between RNA molecules, RNA-protein, and RNA-small molecules. This short article accounts for the general methods for NMR sample preparations, and also partially covers the resonance assignments of structured RNA molecules.

Keywords NMR, RNA, structure, method

Introduction

In addition to their central roles in translating genetic information to produce proteins by mRNA and tRNA, RNA molecules have been evolved to perform a variety of cellular functions such as chemical catalysis and regulations.[1-3] Recent developments in biophysical methods such as x-ray crystallography, electron microscopy (EM), small-angle scattering (SAXS), fluorescence resonance energy transfer (FRET), electron paramagnetic resonance (EPR) and NMR spectroscopy have provided different types of structural information on biomolecules such as proteins and nucleic acids.[4-5] Among them, NMR spectroscopy is a unique technique that provides the atomic level structural information of RNA molecules in liquid state. NMR technique can also provide dynamics information on timescales from picosecond to second ranges.

For NMR studies on RNA molecules, we need to determine whether the RNA of interest is feasible to study using NMR spectroscopy. First, the size of RNA should be considered. Well dispersed NOESY spectra is critical for measuring NOEs to get distance information between spatially apart two protons, but the large RNA molecules are difficult to study with NMR methods because of the intrinsically serious spectral overlap. The reasonable length of RNA for NMR structure determination is generally < 60 nts (< ~18kDa), but if we just need other information such as dynamics than structure determination, a bigger sized RNA (~100 nts) can be analyzed. Next, given that the sequence of the RNA is known, the secondary structure of RNA should be predicted. Several RNA folding programs such as Mfold[6] are available on the web, where the secondary structural information of RNA can be calculated by providing the RNA sequence.

In order to determine the RNA structures with NMR spectroscopy, which is the same case for x-ray crystallography and other structural method, obtaining a single RNA conformation is important. Therefore, when the designed RNA construct has

* Address correspondence to: **Nak-Kyoon Kim,** [1]Advanced Analysis Center, Korea Institute of Science and Technology, 39-1, Hawolgok-Dong, Seongbuk-Gu, Seoul, 136-791, Republic of Korea, Tel: 82-2-958-5996; Fax: 82-2-958-5969; E-mail: nkkim@kist.re.kr

multiple conformations, designing a new RNA construct that results in only one conformation, will save the overall time for structural analysis. This structural homogeneity will prevent the spectral cloudiness and ambiguity of the resonance frequencies caused by multiple conformers.

In the following text, the general procedure to prepare RNA samples using *in vitro* transcription with T7 RNA polymerase and some NMR experiments required for RNA assignments will be described.

RNA sample preparation for NMR studies

A large quantity (0.5-1mM) of homogeneous RNA is required to obtain a good NMR spectrum. RNA can be obtained either by chemical synthesis or by *in vitro* transcription using a DNA template and DNA-dependent RNA polymerase such as T7 RNA polymerase. RNA also can be purified from cell extract, but we don't cover the last case in this article. For the small RNA molecules whose lengths are less than 10 nucleotides, it is recommended to purchase a chemically synthesized RNA from commercially available sources because of the low *in vitro* transcription efficiency with RNA polymerases. In many cases, the structure of small-sized RNA can be analyzed without ^{13}C or ^{15}N isotope labeling. However, even for the small RNA molecules, when RNA contains a repeating sequence that causes spectral overlap, the ^{13}C or ^{15}N isotope labeled RNA is required. For the RNA sizes from 10-60 nts, usually RNA is transcribed *in vitro*, since a large amount of chemically synthesized big RNA is quite expensive even for the non- isotope labeled RNA.

For *in vitro* transcription with T7 RNA polymerase, a DNA template is required. The promoter sequence needs to be included at the 3'-end of the template sequence for polymerase recognition (Fig. 1). For the RNA molecule with its size between 10 and 60 nts, a synthetic single-stranded DNA with a promoter region is commonly used. In this case, DNA with complementary sequence (Top sequence, Fig. 1) for the promoter sequence should be separately prepared. It is often desirable to use a double-stranded DNA including the promoter region, which could increase the transcription yield. If possible, it is also helpful to

Figure 1. DNA template design for *in vitro* RNA transcription with T7 RNA polymerase. An annealed form of DNA duplexed with RNA sequences is shown.

include 2'-OMe group at the last one or two nucleotides of the 5'- end of DNA template, which could prevent the RNA 3'- end heterogeneity from N+1 mer RNA. Also, it is desirable to design an RNA sequence starting with at least one or two guanine residues at the 5'- end for higher transcription yield. As the RNA size becomes longer (>60 nts), it becomes more difficult and expensive to obtain a large amount of synthetic DNA template due to the decrease in phosphoramidite coupling efficiency. In this case, a small amount of DNA template is first chemically synthesized and cloned into a high copy plasmid such as pUC19. The cloned DNA in a plasmid is then transformed into a bacterial cell such as *E.Coli*. After DNA amplification from the cell growth, a large quantity of DNA can be purified by using the commercially available Giga-prep DNA preparation kits. In order to prepare an RNA sample required for NMR analysis, a large amount of T7 RNA polymerase is required. However, this enzyme is quite expensive for this purpose, so it is recommended to obtain a plasmid DNA for T7 RNA polymerase and to produce it in the laboratory.

Now, given that enough DNA template and T7 RNA polymerase are ready, RNA can be synthesized *in vitro*. The common RNA transcription mixture includes 1x transcription buffer, 2.5 mM DTT, 10-30 mM $MgCl_2$, 2-4 mM rNTP, 0.2-1 μM annealed DNA template, and T7 RNA polymerase. Before the synthesis of a large quantity (10–20 mL) of RNA, a small-scaled (50–100 μL) test transcription in

microtubes is recommended, which will guide to find a best transcription condition. Commonly, 5-10 different test reaction conditions with various concentrations of $MgCl_2$, rNTPs, and T7 RNA polymerase are designed. The test transcription can be performed in either a water bath or a heating block at 37°C. Additionally RNase inhibitor and inorganic pyrophosphatase can also be added to the reaction mixture to enhance the transcription efficiency. After 1-2 hour incubation, the transcription results are analyzed using a small-size PAGE. Gel is stained with toluidine blue and the optimal transcription condition is determined by identifying the most intense RNA band in the gel.

Once the optimal transcription condition is selected, a large-scale transcription can be performed with 10-20 mL size in a 50 mL sterile centrifuge tube, which can generate one to three ~1mM RNA samples for NMR studies. The transcription yield will vary depending on the size and sequence of RNA. The reaction is incubated for 6-8 hours at 37°C, and quenched by adding EDTA whose amount is equivalent to Mg^{2+} ions. The 3M sodium acetate buffer with pH 5.2 is also added for efficient ethanol precipitation. Next, RNA is ethanol-precipitated overnight at -20 °C, and the RNA pellet is obtained from centrifugation at 12000 – 14000 rpm at 4 °C. The pellet contains a lot of salt from the precipitation procedure, so it is recommended to wash the pellet with 70% ethanol, or to exchange the re-suspended RNA solution with water for one or two times. The RNA is purified using the sequencing-gel-sized big gel and 10-20% polyacrylamide denaturing gel. The percentage of the polyacrylamide is chosen based on the size of RNA.

The desired RNA band is excised and eluted from the electroelution chamber in elutrap system (Whatman).

The purified RNA is then filtered, and the solution is completely exchanged to the desired NMR buffer using centrifugal membrane tubes. For RNA refolding, before the final step of concentration, the RNA is diluted to 10-20 μM and heated at 95 °C for 5 minutes in water bath, followed by cooling in ice for ~ 30 minutes.

Finally, the RNA is concentrated to the volume with a concentration of 0.5–1 mM required for NMR analysis. For the ^{13}C-, ^{15}N- single labeled, or the ^{13}C-^{15}N- double labeled RNA *in vitro* transcription, the labeled rNTPs are used, and the same procedure is used as shown in non-labeled RNA transcription.

NMR analysis of RNA

An optimal RNA buffer condition should be determined to obtain a sharp NMR line width, non-overlapping spectra, and the homogeneous single RNA conformation. First, the RNA sample is prepared in 95% H_2O and 5% D_2O, and the imino region (9-15ppm) of 1D NMR spectra is investigated The 11-echo or watergate water suppression pulse sequences are usually used.[7]

In general a series of salt and pH titrations should be performed at several temperatures. The pH of the RNA buffer is usually adjusted to pH 5.5 using HCl, and increased by 0.5 up to pH 7.0 using NaOH. The higher pH would degrade RNA, and also induce the decrease in proton signals. Titrations can be performed at low RNA concentrations (100-200 μM). When the optimal pH is decided, monovalent metal ions such as Na^+ and K^+ should also be added, since addition of salt stabilizes the negatively charged RNA backbone and the RNA tertiary structures. A common range of the salt concentration is from 50mM to 300mM. However, too much salt should be prevented, since it causes NMR line broadening and difficulties in shimming. A phosphate buffer is recommended for RNA analysis, since the pH is not affected much by the temperature changes, and also it does not generate any proton signals.

Once the optimal buffer condition is determined, a series of NMR spectra are required for structural analyses. First, most of the RNA proton resonances should be assigned. Exchangeable protons in the base (iminos H1/H3 in G/U, and aminos H41/H42 in C), non-exchangeable proton in the base (H2 in A, H5/H6 in C and H8 in A/G) and in the sugar (H1', H2', H3', H4' and H5'/H5''), are assigned. The 2'-hydroxyl protons can also be assigned in 2D NOESY. Second, the NOEs and several dihedral angles should also be measured for structural

calculations. Non ^{13}C- or ^{15}N- labeled RNA can be used to obtain initial structural information.

The secondary structure of RNA can be recognized from the sequential walk in the imino proton region of 2D NOESY. Each resonance in the imino region (9-15 ppm) stands for either Watson-Crick (A-U or G-C) or non-canonical (e.g. G-U wobble) base-pairings in the helical region of RNA. The non-exchangeable protons in the base and sugar moiety can also be assigned. The sequential assignments for H8/H6(N) → H1'(N) → H8/6(N+1) in the helical region of RNA can be achieved in 2D NOESY. The H2 of adenine and H5 of uracil or cytosine can be assigned in this spectrum. 2D TOCSY is useful to confirm the H5 and H6 assignments, and 2D HCCH-COSY is used for H1' and H2' assignments.

For the unambiguous proton assignment, the uniformly labeled or base-selective ^{13}C-, ^{15}N- labeled RNA samples are required. For the bigger sized RNA (>50 nt), the partially deuterated (H2/H6/H8, H1', H2', but D3', D4', D5'/D5'') RNA samples can be used. The imino and amino exchangeable protons can be assigned from 2D ^{15}N-correlated CPMG NOESY, 2D ^{1}H-^{15}N HMQC, 2D H5(C5C4N)H spectra.[8] Direct hydrogen binding interaction can be detected from 2D HNN COSY spectra.[9] The assignments for all non-exchangeable protons can be achieved from analysis of 2D ^{1}H-^{13}C HSQC, 2D HCCH-COSY, and 3D HCCH-TOCSY.[10,11] A suite of 2D-filtered/edited proton NOESY (F2f, F1fF2e, F1fF2f, and F1eF2e) experiments on base-specific ^{13}C,^{15}N- labeled RNAs are very useful to resolve ambiguous assignments in overlapped regions and to obtain NOE restraints.[12]

The sugar pucker is determined as C2′-*endo*, if the residues shows strong H1'-H2' cross-peaks in 2D DQF-COSY.[13] $^{3}J_{H2'P}$ and $^{3}J_{CP}$ can be measured using ^{31}P spin echo difference 2D CT-HSQCs to determine the ε dihedral angles for the loop residues.[14] The base conformation is determined to be *syn* ($\chi = 60 \pm 30°$), if the residue showed an intra-nucleotide strong H1'-H8 NOE and a weak H2'-H8 NOE in a 2D NOESY. Further NMR experiments such as residual dipolar couple can be performed for obtaining more restraints that are required for structural calculations.

Discussion

Recently, RNA molecules have gained enormous attention due to the new discoveries of their biological roles in the cell. RNA can function alone or with other biomolecules such as RNA and proteins. In order to achieve their roles, the functional RNAs have a variety of interesting tertiary structures, and the function of these RNA molecules can be understood by investigating their structures. NMR spectroscopy has a unique advantage such that the atomic level structures of RNA molecules in liquid state and the dynamics information can be obtained.

This article shortly summarized how the RNA samples are prepared for NMR studies, and how the RNA assignments are performed. The listed NMR experiments are essential, and there are more complicated NMR methods that are designed for special assignments for specific elements in RNA.

Acknowledgements

This study was supported by KIST grant (No. 2V03611, 2E24490) and by Global research lab (No. 2U04660).

References

1. J.F. Atkins, R.F. Gesteland, T.R. Cech, RNA worlds: From Life's origins to diversity in gene regulation (2010)

2. M.V. Rodnina, *Curr. Opin. Struct. Biol.* 23, 595. (2013)
3. J. Nandakumar, T.R.Cech, *Nat. Rev. Mol. Cel.l Biol.* 14, 69. (2013)
4. Y.-G. Choi, H.-E. Kim, J.-H Lee, *J. Korean Magn. Reson. Soc.* 17, 76. (2014)
5. J.-H. Jeong, Y.-G. Park, S.-S. Choi, Y. Kim, *J. Korean Magn. Reson. Soc.* 17, 81. (2014)
6. M. Zuker, *Nucleic Acids Res.* 31, 3406. (2003)
7. V. Sklenar, A. Bax, *J. Magn. Reson.* 74, 469. (1987)
8. J. Wohnert, R. Ramachandran, M. Gorlach, L.R. Brown, *J. Magn. Reson.* 139, 430. (1999)
9. A.J. Dingley, L. Nisius, F. Cordier, S. Grzesiek, *Nat. Protoc*. 3, 242. (2008)
10. J. Cromsigt, B. van Buuren, J. Schleucher, S. Wijmenga, *Methods Enzymol.* 338, 371. (2001)
11. T.Dieckmann, J. Feigon, *J. Biomol. NMR* 9, 259. (1997)
12. R.D.Peterson, C.A.Theimer, H. Wu, J. Feigon, *J. Biomol. NMR* 28, 59. (2004)
13. G.Varani, F. Aboulela, F.H.T. Allain, *Prog. NMR Spectrosc.* 29, 51. (1996)
14. L.E.Kay, P.Keifer, T.Saarinen, *J. Am. Chem. Soc*. 114, 10663. (1992)