Site-specific isotope labeling of long RNA for structural and mechanistic studies

Ikumi Kawahara1,2, Kaichiro Haruta1,*, Yuta Ashihara1, Daichi Yamanaka1, Mituhiro Kuriyama1, Naoko Toki1, Yoshinori Kondo1, Kenta Teruya3, Junya Ishikawa4, Hiroyuki Furuta4, Yoshiya Ikawa4, Chojiro Kojima2,5 and Yoshiyuki Tanaka1,*

1Graduate School of Pharmaceutical Sciences, Tohoku University, Sendai 980-8578, 2Graduate School of Biological Sciences, NAIST, Ikoma 630-0192, 3Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto 603-8334, 4Department of Chemistry and Biochemistry, Graduate School of Engineering, Kyushu University, Fukuoka 819-0395 and 5Institute for Protein Research, Osaka University, Suita 565-0871, Japan

Received December 18, 2010; Revised October 11, 2011; Accepted October 12, 2011

ABSTRACT

A site-specific isotope labeling technique of long RNA molecules was established. This technique is comprised of two simple enzymatic reactions, namely a guanosine transfer reaction of group I self-splicing introns and a ligation with T4 DNA ligase. The trans-acting group I self-splicing intron with its external cofactor, ‘isotopically labeled guanosine 5'-monophosphate’ (5'-GMP), steadily gave a 5'-residue-labeled RNA fragment. This key reaction, in combination with a ligation of 5'-remainder non-labeled sequence, allowed us to prepare a site-specifically labeled RNA molecule in a high yield, and its production was confirmed with 15N NMR spectroscopy. Such a site-specifically labeled RNA molecule can be used to detect a molecular interaction and to probe chemical features of catalytically/structurally important residues with NMR spectroscopy and possibly Raman spectroscopy.

INTRODUCTION

In mechanistic studies of functional RNA, the signals from the functional residues (pinpointed studies) are important for characterizing their chemical nature and for elucidating the mechanism of action of the RNA (1). For this purpose, isotope labeling of functionally/structurally important residues enables us to extract pinpoint information from the functional RNA, by using NMR/Raman spectroscopy and mass spectrometry. Currently, site-specific isotope labeling of RNA molecules has only been achieved with chemical syntheses (2–11), since an alternative enzymatic RNA synthesis is not compatible with site-specific modifications. However, the chemical synthesis is applicable only to ‘short’ RNA/DNA oligomers (typically <20 nt) (2–17), due to low yields of long RNA/DNA chains. Therefore, to directly access spectral data from important residues in ‘long’ functional RNAs and to skip the time-consuming assignment step, we established a site-specific isotope labeling technique of RNA molecules without any limitation of the RNA length (Figures 1 and 2).

In NMR spectroscopy of RNAs, the isotope labeling technique [i.e. uniform (18,19) and nucleotide-specific (20–23) ones] expanded the applicable molecular size limit. Recently, the segmental isotope labeling technique of RNA (24–26) further enlarged this size limit. Although these techniques were oriented to 3D structure determination, a site-specific labeling technique of DNA/RNA oligomers is oriented to their chemical characterizations, such as the identification of the hydrogen bonding (2,5,12–15,27,28) and metalation site (3,4,6–8,11,16,17) in a pinpoint manner. Specifically, direct evidence of the 15N–15N J-coupling across the hydrogen bond (H2JNN) (12) and the HgII-mediated 15N–15N J-coupling (H2JNN) in T-HgII-T base pairs (8,16,17), were obtained by using appropriate site-specifically labeled DNA/RNA oligomers (for H2JNN, 15N–HgII–15N and 15N–HgII–14N, respectively). More importantly, the derived fine spectral data were utilized to evaluate the strength of the hydrogen bond (5,13–15,27,28) and N-metal bond (8,17). Therefore, in order to apply such fine NMR spectra to any functional RNA molecule, site-specific labeling techniques without any applicable size-limit are becoming indispensable day by day.

*To whom correspondence should be addressed. Tel/Fax: +81 22 217 5917; Email: tanaka@mail.pharm.tohoku.ac.jp
Correspondence may also be addressed to Kaichiro Haruta. Tel: +81 438 64 2314; Fax: +81 438 64 2372; Email: Kaichiro.Haruta@mitsui-chem.co.jp

© The Author(s) 2011. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
MATERIALS AND METHODS

Synthesis of RNA oligomers for site-specific labeling of the hammerhead ribozyme

The sequences of the used RNA/DNA oligomers are highlighted with yellow background in Figure 2d. RNA oligomers (5’-fragment, 3’-fragment precursor and substrate strand for the hammerhead ribozyme) were synthesized by a DNA/RNA synthesizer (ABI model 392, CA, USA). The C17 residue at the cleavage site of the hammerhead ribozyme substrate (inhibitor) was substituted with 2’-O-methylcytidine to prevent the cleavage reaction with the hammerhead ribozyme (Figures 2a and 3a). The DNA bridge for the ligation reaction was purchased from TSUKUBA OLIGO SERVICE Co., Ltd. (Tsukuba, Japan). RNA oligomers were purified on an anion-exchange column (mono-Q; GE Healthcare UK, Ltd., Buckinghamshire, England) with a linear NaCl gradient (0–2 M) under denaturing conditions (8 M urea). Excess NaCl and urea were washed out using an ultrafiltration device (Amicon Ultra-15 3000 MWCO; Millipore, MA, USA).

Preparation of non-labeled hammerhead ribozyme

The non-labeled hammerhead ribozyme was prepared by in vitro transcription. The template gene of the full-length hammerhead ribozyme was constructed in a pCR®.2.1-TOPO® vector (TOPO® TA Cloning kit; Invitrogen, CA, USA), using synthetic DNA oligomers containing the T7 promoter and the coding sequence. Using this plasmid, PCR amplification was performed with the following primers. Forward primer: 5’-GCGTAATACGACCTCACTATAGGATGTACTACCAGCTGA-3’ and reverse primer: 5’-mGmGCGTTTCGTCTATTTGGGACTC-3’. To prevent an overrun of T7 RNA polymerase, two 2’-O-methylguanosine (mG) residues were added to the 5’ end of the reverse primer (29,30). The RNA oligomer was directly transcribed from
the PCR product using MEGA shortscript™ kit (Applied Biosystems, CA, USA). The transcript was purified on an anion-exchange column (mono-Q), with a linear NaCl gradient (0–2 M) in denaturing conditions (8 M urea), and desalted by a gel filtration column (TSK-GEL G3000PW; TOSOH, Tokyo, Japan).

Preparation of Tetrahymena group I intron and the general procedure for the optimization of the guanosine transfer

Trans-acting Tetrahymena group I intron (31–51) (the processing enzyme for the guanosine transfer reaction) was transcribed in vitro, using T7 RNA polymerase. For PCR amplification of the template DNA for in vitro transcription, we used a plasmid (pTZIVSU) (42,47) containing wild-type Tetrahymena group I intron (Figure 2b and Supplementary Figure S1), together with the following primers. By using the following primers, we obtained the trans-acting Tetrahymena group I intron (Figure 2c).

Forward primer:

\[ 5'-\text{GAAGAGGCGTATACGACCTACTATAGGGATCGGAGATCTCAAAGTTATCACGGCATGACG3'} \]

Reverse primer:

\[ 5'-\text{GTACTCCAAAATCAATATAATATCTTACGCATACAAATTAG3'} \]

The Tetrahymena group I intron was transcribed from the PCR product using the script MAX™ Thermo T7 Transcription kit (TOYOBO, Osaka, Japan). To remove RNA polymerase, rNTPs and pyrophosphate, the transcript was cleaned by phenol–chloroform extraction and dialysis using an ultrafiltration device (Amicon Ultra-15 3000 MWCO) or a cellulose dialysis tube (Spectra/Por® Dialysis Membrane MWCO: 3500, Spectrum Laboratory Inc., Rancho Dominguez, USA). Without this treatment, residual guanosine 5'-triphosphate (5'-GTP) promoted the guanosine transfer reaction without adding any external guanosine 5'-monophosphate (5'-GMP; Supplementary Figure S2, lanes 4–6). In contrast, when including the dialysis step this reaction did not occur (Supplementary Figure S2, lanes 7–9). Experiments using a cellulose dialysis tube resulted in better exclusion of residual non-labeled 5'-GTP and higher enzymatic activity, compared with ultrafiltration dialysis. Hence, we were able to reduce the concentration of group I intron to 1.4 μM, and this result is presented in Supplementary Figure S2.

Guanosine transfer reaction for the NMR sample

Prior to the guanosine transfer reaction, the Tetrahymena group I intron was refolded using the following procedure. Dialyzed Tetrahymena group I intron (14 μM) was heated to 80°C and slowly cooled down to 58°C (reaction temperature) in a 3-ml solution (30 mM Tris–HCl pH 7.2, 10 mM NH₄Cl and 3 mM MgCl₂). Following Tetrahymena group I intron refolding, 50 μM 3'-fragment precursor and 2 mM uniformly 15N-labeled 5'-GMP (Cambridge Isotope Laboratories, MA, USA) were added to the solution and incubated for 4 h (Figure 1c). The derived product (labeled 3'-fragment) was purified on an anion-exchange column (mono-Q) with a linear NaCl gradient (0–2 M) under denaturing conditions (8 M urea). This was followed by ethanol precipitation and resuspension in TE buffer (10 mM Tris–HCl pH 7.5 and 1 mM EDTA).

Ligation reaction for the NMR sample

The 3'- and 5'-fragments were ligated with T4 DNA ligase (Fermentas, Ontario, Canada), using a DNA bridge (Figure 2d) as a complementary strand for the ligation.
The ligation reaction of the hammerhead ribozyme was basically performed as described in the literature (24), although the optimization of the RNA/DNA strand ratio is possible (see 'Preparation of site-specifically 15N-labeled 76-nt RNA molecule [hXBPI mRNA (516–591)]' for modified conditions). The three fragments were annealed in 2.5 ml of the pre-reaction mixture (100 μM 5'-fragment, 50 μM 3'-fragment, 75 μM DNA bridge, 40 mM Tris–HCl pH 7.6, 10 mM DTT, 10 mM MgCl2 and 0.5 mM 5'-ATP) that was heated to 95°C and slowly cooled down to 37°C (24). The T4 DNA ligase (850 units per 2.5 ml of pre-reaction mixture) was then added and incubated at 37°C for 12 h (Figure 1d). This was followed by a DNase I (TURBO DNase; Applied Biosystems, CA, USA) digestion of the DNA bridge. The ligated product was purified on an anion-exchange column (mono-Q) as described above. The resulting product was desalted by a gel filtration column (TSK-GEL G3000PW).

**NMR measurements**

We recorded NMR spectra of substrate-free hammerhead ribozymes, with and without isotope labeling, as well as substrate-bound labeled ribozyme (Figure 3, Supplementary Figures S3 and S4). NMR spectra of substrate-free labeled/non-labeled ribozymes were recorded in a solution containing 0.5 mM ribozyme, 10 mM sodium phosphate pH 7.0 and 50 mM NaCl in 5% D2O (Figure 3; Supplementary Figures S3 and S4). The NMR spectrum of the substrate-bound ribozyme was recorded in the same solution, but with 0.4 mM of the 1:1 ribozyme–substrate complex (Supplementary Figure S3). All NMR samples were heated to 58°C for 3 min and cooled to room temperature prior to NMR measurements. One-dimensional 1H NMR spectra were recorded on Bruker DRX800 spectrometer and 1H-15N HSQC spectra were acquired on Bruker Avance500 spectrometer, using cryo-cooled probes. All spectra were recorded at 293 K. Spectra were processed with the program NMRPipe (version 3.0) (52).

**Optimizations of the guanosine transfer and ligation reactions**

Optimizations of the guanosine transfer reaction with trans-acting Tetrahymena group I intron were performed in terms of concentrations of Mg2+, group I intron and 5'-GMP (Supplementary Figures S5 and S6), temperature (Supplementary Figure S7) and reaction volume (Supplementary Figure S8). The dialyzed group I intron (0.5–25 μM) was heated to 80°C and slowly cooled down to the reaction temperature (48–68°C) in a solution containing 30 mM Tris–HCl pH 7.2, 10 mM NH4Cl and MgCl2 (3–30 mM). After the refolding of Tetrahymena group I intron, 50 μM 3'-fragment precursor and 5'-GMP (0.05–2 mM) were added to the solution, and the reaction mixture was incubated at the temperature indicated in each experiment. All optimization experiments were carried out in 12 μl (Supplementary Figures S5–S7) except for the scale-up experiment.

Figure 3. Labeled RNA (hammerhead ribozyme) and its NMR spectra. (A) Hammerhead ribozyme (magenta, red and green) and substrate (grey) with its cleavage site (blue arrowhead). (B) One-dimensional 1H NMR spectrum of the authentic non-labeled sample. (C) One-dimensional 1H NMR spectrum of the labeled sample. (D) Nitrogen-15-edited 1D 1H NMR spectrum of the labeled sample. (E) 1H-15N HSQC spectrum of the labeled sample. The doublet imino proton signal of G10.1 is highlighted in the spectrum (c).
Preparation of site-specifically $^{15}$N-labeled 76-nt RNA molecule [hXBPI mRNA (516–591)]

As an example of a site-specifically $^{15}$N-labeled long RNA molecule, hXBPI mRNA (516–591) was prepared (Supplementary Figure S11). Its sequence and labeled position (G544) are shown in Supplementary Figure S11a. The 3'-fragment precursor (Supplementary Figure S11b) was processed with a group I intron in the solution containing 50 $\mu$M 3'-fragment precursor, 2.8 $\mu$M group I intron, 30 $\mu$M Tris–HCl pH 7.2, 10 mM NH$_4$Cl and 3 mM MgCl$_2$ (Supplementary Figure S11c). The resulting solution containing guanosine-transferred 3'-fragment was directly subjected to a ligation reaction without purification (Supplementary Figure S11f) in combination with the 5'-fragment (Supplementary Figure S11c) and the complementary DNA bridge (Supplementary Figure S11d). The solution conditions for the ligation were 52 $\mu$M 5'-fragment, 50 $\mu$M 3'-fragment, 75 $\mu$M DNA bridge, 40 mM Tris–HCl pH 7.6, 10 mM DTT, 10 mM MgCl$_2$, 0.5 mM 5'-ATP and 0.5 unit/$\mu$L T4 DNA ligase.

The derived full-length 76-nt RNA was isolated with PAGE and electro-elution. $^{1}$H–$^{15}$N HSQC and $^{15}$N-edited $^{1}$H NMR spectra were recorded with Bruker Avance-I 800 spectrometer, using cryo-cooled probe (Supplementary Figures S12 and S13). The solution for NMR measurements contains 0.2 mM 76-nt RNA molecule, 10 mM sodium phosphate pH 7.0 and 50 mM NaCl, in 10% D$_2$O. All spectra were recorded at 303 K. Spectra were processed with the program NMRPipe (version 3.0) (52).

RESULTS

To produce site-specifically labeled RNA molecules, we employed an enzymatic activity of group I introns (31–51), specifically, the cleavage reaction of the 5' exon–intron junction with an external guanosine cofactor (the first step of the self-splicing reaction; Figures 1b, 2d and Supplementary Figure S1). If this reaction takes place with isotopically labeled guanosine or its 5'-mono/di/triphosphate form, one could introduce a labeled guanosine residue at the 5'-end of the intron itself (Supplementary Figure S1). To make this reaction into an intermolecular reaction and obtain labeled RNA at its 5'-end, we employed a trans-acting Tetrahymena group I intron (36,39,41,44) (Figure 2c) instead of a wild-type cis-acting enzyme (Figures 2b and Supplementary Figure S1).

We prepared this trans-acting group I intron by in vitro transcription. The derived group I intron was dialyzed to eliminate co-existing non-labeled guanosine 5’-triphosphate and subjected to the guanosine transfer reaction (Supplementary Figure S2). Finally, we postulated that the remainder of the non-labeled 5'-fragment could be enzymatically ligated (10,24–26) (Figures 1b and 2d) to obtain a single-residue labeled RNA molecule.

We chose a hammerhead ribozyme from Schistosoma mansoni (53) as an example for functional RNA, and selected G10.1 [metal ion-binding site (3,6–8,11,54,55)] as the labeled site (Figure 2a). The starting materials (RNA fragments) were prepared according to the synthetic plan (Figure 2d). We then optimized the conditions for the guanosine transfer in terms of concentrations of Mg$^{2+}$, group I intron and 5'-GMP (Supplementary Figures S5 and S6), temperature (Supplementary Figure S7) and reaction volume (Supplementary Figure S8). After the optimizations, we obtained a high yield of the 3'-fragment with $^{15}$N-labeled guanosine at the 5'-end (Figure 1c) (gel-based yield: ~80%, isolated yield after purification: ~50%). This labeled fragment was ligated with a non-labeled 5'-fragment, using T4 DNA ligase (Figure 1d), resulting in a site-specifically labeled RNA molecule with a 34% overall yield, including sample loss due to purification steps.

In order to confirm the incorporation of the labeled guanosine residue, we recorded NMR spectra of the labeled RNA (Figure 3; Supplementary Figures S3 and S4). In the $^{1}$H–$^{15}$N HSQC and $^{15}$N-edited $^{1}$H NMR spectra, we observed a unique H1–N1 cross peak and an imino proton resonance, respectively (Figure 3d and e), demonstrating that labeled guanosine was definitely incorporated in the RNA strand. Importantly, the observed unique cross peak unambiguously indicated that a single guanosine residue was labeled (Figure 3). Furthermore, following the addition of a substrate (inhibitor) strand against the labeled strand (hammerhead ribozyme) under low ionic conditions, we observed twin peaks of H8–N7 correlation (Supplementary Figure S3). Because only a single guanosine was labeled, we concluded that these peaks arose from G10.1 (Supplementary Figure S3), possibly due to the difference in substrate-binding states (substrate-free and bound forms) or a structural polymorphism.

In order to examine if this method is compatible with other sequences and much longer RNA, we produced a site-specifically labeled 76-nt RNA [hXBPI mRNA (516–591)]. By using our labeling method, this 76-nt RNA molecule with labeled guanosine at the G544 site was prepared in a total yield of 19% (Supplementary Figure S11). In $^{1}$H–$^{15}$N HSQC spectrum, a single cross peak was observed for this labeled RNA, which again indicates that a single residue labeled RNA was successfully produced (Supplementary Figure S12). Here, we demonstrated that our labeling method is compatible with long RNA and robust for sequence variation.

DISCUSSION

In this study, we established for the first time a site-specific labeling technique of RNA molecules without any applicable size-limit. We emphasize that our method is...
compatible with T4 RNA ligases I and II, as ligating enzymes (Supplementary Figure S9). In addition, an adenosine residue can also be labeled with our method, given that a mutant group I intron has been shown to promote an adenosine transfer reaction (46). Actually, we confirmed that this adenosine transfer reaction proceeds (Supplementary Figure S14).

Furthermore, our simple, two-step enzymatic reaction protocol gives a high yield of the hammerhead ribozyme, namely, total yield: 34%; guanosine transfer: ~80% (a conversion rate based on a gel image) and ~50% (isolated yield including an HPLC purification loss); ligation: ~70% (isolated yield including an HPLC purification loss). In the case of 76-nt RNA molecule, we found that the ligation with T4 DNA ligase can be performed successively after the guanosine transfer reaction without any purification of the guanosine-transferred RNA fragment. This simpler protocol gave 19% total yield even for such a long 76-nt RNA molecule. In any case, it was demonstrated that our enzymatic method can be applied for RNA molecules with different sequences and chain lengths. Thus, our labeling method is robust and adaptable to any size of RNA molecules. However, caution must be taken to the selection of Anti-Internal Guide Sequence (Anti-IGS) so as to avoid a secondary structure formation of 3'-fragment precursor (substrate for guanosine transfer reaction) which inhibits the reaction (Supplementary Figure S15). For the reaction scale, we confirmed that the reactions proceeded at 0.15 μmol scale in 3 ml reaction volume for guanosine transfer (Supplementary Figure S8) and at 0.15 μmol scale in 3 ml reaction volume for ligation (Supplementary Figure S10) without affecting the yield.

Next, we compared our enzymatic method with chemical synthesis. Technically, it is possible to obtain site-specifically labeled long RNA by chemical synthesis and ligation technique, where a site-specifically labeled short RNA oligomer is chemically synthesized and ligated with (a) non-labeled RNA oligomer(s). However, for the syntheses of site-specifically labeled oligonucleotides, one should have techniques of organic synthesis (2–6,9,12–17) that biochemists and structural biologists don’t usually have. In contrast, our enzymatic pathway can be performed with commercially available sources (15N-labeled NMP, an RNA polymerase and a ligase), as well as fundamental biochemical techniques. Thus, this methodology has opened a new avenue to the preparation of site-specifically labeled long RNA molecule.

From a spectral point of view, the assignments of long RNA molecules are very time-consuming. However, with site-specifically labeled RNA molecules, one can directly access the spectral data of the desired site, e.g. catalytic residues and interfaces of intermolecular complexes. This progress will provide us with rapid assessment for mechanistic studies of functional, large RNA molecules.

In conclusion, we established a ribozyme-assisted single-guanosine-labeling technique of RNA molecules. Using this labeled RNA molecule, we were able to rapidly extract an NMR signal from a specific residue to explore complexation states (substrate-free and bound forms) and possibly electronic states of functional residues. It should also be mentioned that the derived site-specifically labeled RNA can be applicable to Raman spectroscopy and mass spectrometry. Thus, our method opened the door to analyze the mechanism of the biologically important RNA molecules.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online: Supplementary Figures 1–15.

**ACKNOWLEDGEMENTS**

This work was performed using the NMR spectrometer with the ultra-high magnetic field under the Cooperative Research Program of Institute for Protein Research, Osaka University.

**FUNDING**

Funding for open access charge: Human Frontier Science Program (HFSPO) from HFSPO, France, and a Grant-in-Aid for Scientific Research (C) (20550145) from the Ministry of Education, Culture, Sports, Science and Technology, Japan; Chiiki-Innovation and Senryaku-teki-Kiban-gijutu-Koudoka-Shienjigyō from Ministry of Economy, Trade and Industry, Japan.

Conflict of interest statement. None declared.

**REFERENCES**


