J|A|C|S
COMMUNICATIONS

Published on Web 07/20/2002

NMR Study of 100 kDa HCV IRES RNA Using Segmental Isotope Labeling

Insil Kim, Peter J. Lukavsky, and Joseph D. Puglisi*

Department of Structural Biology, School of Medicine, Stanford University, Stanford, California 94305-5126

Received April 23, 2002

NMR has proven to be a powerful tool to determine structural and dynamic features of RNA molecules. However, the application of NMR to RNA has been limited to molecules less than 25 kDa because of resonance overlap and broadening. Deuterium substitution[1] and TROSY experiments[2] partially alleviate line broadening, but resonance overlap is still a major barrier to the study of large RNAs.[3] Because most biological RNAs are larger than 25 kDa, NMR investigations have focused on smaller RNA domains of a large RNA. Although this approach is aided by the high local thermodynamic stability of RNAs, the isolated domain may adopt a different structure than that in the full RNA. Segmental labeling of RNA, in which part of the RNA is labeled with stable isotope, reduces the complexity of spectra, and could potentially allow NMR studies of very large RNAs and their complexes. Unambiguous resonance assignments have been achieved through selective isotope labeling of RNAs using T4 DNA ligase, T4 RNA ligase, or chemical synthesis.[4−6] However, the application of these methods has been limited by complicated procedures or high costs. In this paper, we present a general method of segmental labeling with T4 RNA ligase and hammerhead ribozyme. Segmental labeling allowed the NMR characterization of domain II of the 100 kDa internal ribosomal entry site (IRES) of the Hepatitis C virus (HCV). The approach presented here demonstrates the power of NMR to study large RNAs of critical biological function.

T4 RNA ligase catalyzes the ATP-dependent joining of a 3′ "donor" RNA having a 5′-terminal monophosphate and a 5′ "acceptor" RNA terminating in a 3′-hydroxyl. The product of ligation has a standard 3′,5′ phosphodiester linkage.[7] The efficiency of ligation depends on the structure of the 5′ and 3′ ends, and it is very efficient for RNAs containing a single-stranded region as the site of ligation. To prevent intramolecular ligation and the formation of byproducts, both termini of the two fragments should be properly arranged. Both sides of the 5′-fragment should be dephosphorylated, and both sides of the 3′-fragment should be phosphorylated. Previously, laborious treatments of alkaline phosphatase, T4 polynucleotide kinase, and deprivation by $NaIO_4$ were executed prior to ligation, and in every step the products were purified by PAGE or HPLC.[5]

Here, we describe a simple one-step reaction to prepare RNA fragments for ligation using in vitro transcription and hammerhead ribozyme cleavage. The hammerhead ribozyme catalyses a transesterification reaction generating a cyclic 2′,3′-phosphodiester and a free 5′-hydroxyl terminus of the product RNAs.[8,9] Therefore, by using the 5′-hammerhead ribozyme, the RNA product has a 5′-hydroxyl terminus after transcription, and it can be used directly as an acceptor without the treatment with calf intestinal alkaline phosphatase. In contrast, by using 3′-hammerhead ribozyme, the RNA product has a 2′,3′-phosphodiester at the 3′-terminal. In addition, by using the GMP to prime the transcription reaction, the 5′-terminal phosphate present on the transcript is replaced with a 5′ monophosphate.[10] Therefore, by using the 3′-hammerhead ribozyme and GMP, the RNA product has 3′- and 5′-monophosphate termini after transcription, and it could be used directly as a donor without further treatment.

Both the 3′- and 5′-RNA fragments were prepared by in vitro transcription using T7 RNA polymerase. A 64 nucleotide RNA, which corresponds to position 40−104 of the HCV IRES, was transcribed with ^{15}N-labeled nucleotide triphosphates. To ensure formation of an oligonucleotide with 5′- and 3′-OH, a 5′-hammerhead sequence was used in cis in the plasmid construct. The hammerhead cleavage is complete after 2−4 h. The 3′-oligonucleotide, which corresponds to nucleotides 105−354 of the HCV IRES, was also synthesized using transcription from a plasmid with T7 RNA polymerase. To ensure that both the 5′- and the 3′-termini have monophosphate groups, the transcription was primed with GMP, and a 3′-hammerhead sequence was added to the plasmid construct. For this oligonucleotide, hammerhead cleavage yields a 2′,3′ cyclic phosphate, which is sufficient to block potential self-ligation. Both oligonucleotides were purified by denaturing gel electrophoresis, and the RNAs were electroeluted into Tris-borate-EDTA buffer.

The purified RNAs were used for ligation by T4 RNA ligase. Optimal conditions of ligation were screened in a series of small reaction. A 10x ligation buffer is added for a final condition of 50 mM Tris-HCl (pH7.5), 3.3 mM dithiothreitol, 2.5% PEG6000, and 60 mM ATP. The concentrations of RNA fragments, $MgCl_2$, PEG6000, and T4 RNA ligase should be optimized. For the HCV IRES, the optimized conditions were 70.7 nmol of 3′ fragment and 66.8 nmol of 5′ fragment mixed in a 40 mL reaction mixture containing 50 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 5.8% PEG6000, 3.3 mM dithiothreitol, 60 mM ATP, 0.006% BSA, and 2000U of T4 RNA ligase (from New England Biolabs, Beverly, MA). After incubation at 15° for 4 h, the T4 RNA ligase was heat-inactivated at 50° for 10 min. Approximately 50% of the RNA fragments were ligated in the above optimal condition (Figure 1).

The ^{1}H−^{15}N HMQC spectrum of the intact 310 nucleotide 100 kDa IRES RNA showed severe spectral overlap (Figure 2a). This spectrum was considerably simplified by the use of the segmentally (nts 40−104) ^{15}N-labeled RNA (Figure 2b). For comparison, a ^{1}H−^{15}N HMQC spectrum was acquired on a [^{13}C, ^{15}N]-labeled RNA corresponding only to domain II (nts 44−118) of the HCV IRES (Figure 2c). Even though the domain II is composed of 77 nt and its molecular weight is ca. 25 kDa, the imino resonances of uridine residues show large dispersion and good resolution in the ^{1}H−^{15}N HMQC spectrum. The imino resonances of uridine residues of domain II were assigned to those of four Watson−Crick A−U base pairs (U46−A116, U113−A50, U59−A109, U101−A66), one reverse hoogsteen (U95−A72), three G−U base pairs (U78−G88, U91−G75, U61−G107), one hoogsteen pair (U92−A74), and one

* To whom correspondence should be addressed. E-mail: puglisi@stanford.edu.

10.1021/ja026647w CCC: $22.00

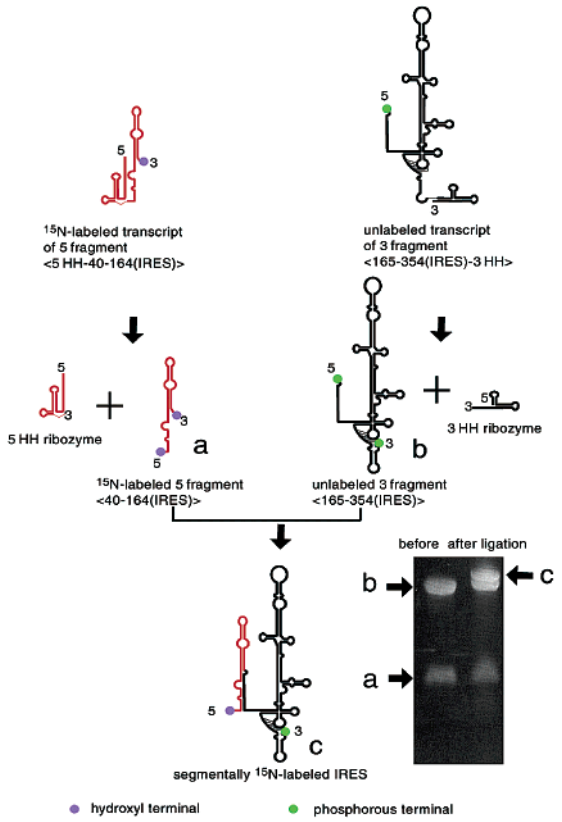

Figure 1. Scheme for the preparation of the segmentally ^{15}N-labeled IRES. The ^{15}N-labeled RNA fragment is shown by red lines, and the unlabeled RNA fragment is shown by black lines. In the lower right is a polyacrylamide gel that analyzes the result of the ligation reaction. a and b are the labeled and unlabeled fragments, respectively; c is the ligated product. Bands were visualized by UV shadowing.

U−U base pair (U103−U64). All of the imino resonances are also observed in the spectrum of the segmentally (nts 40−104) ^{15}N-labeled IRES without any significant chemical shift changes except the downfield imino resonance of U113 that is unlabeleled in the segmentally ^{15}N-labeled IRES. Two downfield imino resonances of U46 and U59 in the segmentally ^{15}N-labeled IRES are observed in the lower threshold of the ^{1}H−^{15}N HMQC spectrum (data not shown). In addition to these resonances, three additional resonances are observed in the segmentally^{15}N-labeled IRES that are also observed in the fully ^{15}N-labeled IRES. Two of the three additional peaks are assigned to the amino resonance of folded A99 and the imino resonance of U80 as determined from smaller RNAs derived from domain II (data not shown). The observation of these resonances may be because of reduced chemical exchange of domain II in the context of the full IRES. The third additional imino resonance may be derived from the imino resonance of U41 due to continuous stacking to the lower stem of domain II, which is not included in the 77nt construct of the domain II.

T1 and T2 values were similar for ^{15}N-labeled domain II oligonucleotide and segmentally ^{15}N-labeled IRES. These data strongly suggest that domain II of the IRES forms an independent structure and that long-range tertiary interactions with other parts of the IRES are not formed. This result supports the previous observation of small-angle X-ray scattering[11] and EM study,[12] which suggest that the HCV IRES does not form a globular structure.

In summary, we provide a simple and efficient method of segmental labeling of large RNA molecules. Although the full ^{15}N-labeled 100 kDa IRES molecule is too large to study directly using NMR because of severe signal overlap, we succeeded in observing the resonances of domain II by using an IRES with domain II selectively ^{15}N-labeled. Segmentally labeled RNAs have a wide range of potential applications in the determination of local properties of RNA domains, global conformation using residual dipolar couplings,[13] or the analysis of folding pathways.

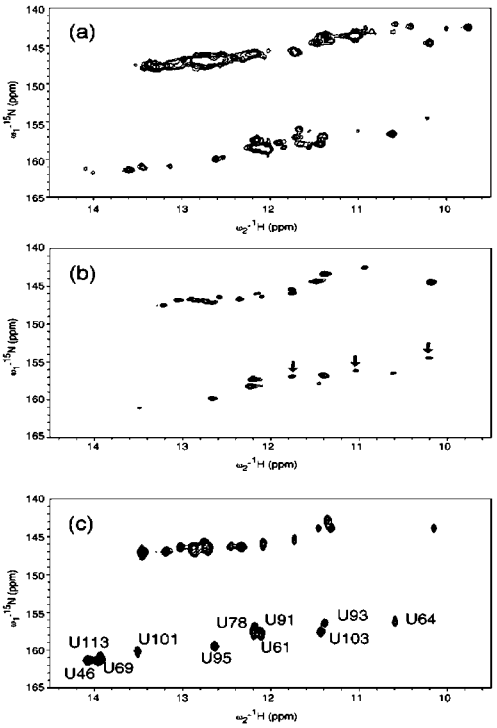

Figure 2. Imino regions of ^{1}H−^{15}N HMQC spectra of (a) 0.1 mM fully ^{15}N-labeled IRES, (b) 0.1 mM segmentally ^{15}N-labeled IRES, and (c) 0.54 mM [^{13}C, ^{15}N]-labled domain II. Three additional resonances observed in the segmentally ^{15}N-lablled IRES are indicated by arrows.

Acknowledgment. This work was supported by NIH grant AI47365, Eli Lilly, and the Hutchinson Foundation.

Supporting Information Available: Experimental details (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Nikonowicz, E. P. *Methods Enzymol.* **2001**, *338*, 320−341.
(2) Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wuthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13585−13590.
(3) Cromsigt, J.; van Buuren, B.; Schleucher, J.; Wijmenga, S. *Methods Enzymol.* **2001**, *338*, 371−399.
(4) Xu, J.; Lapham, J.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 44−48.
(5) Ohtsuki, T.; Kawai, G.; Watanabe, K. *J. Biochem.* **1998**, *124*, 28−34.
(6) Kim, I.; Muto, Y.; Watanabe, S.; Kitamura, A.; Futamura, Y.; Yokoyama, S.; Hosono, K.; Kawai, G.; Takaku, H.; Dohmae, N.; Takio, K.; Saskamoto, H.; Shimura, Y. *J. Biomol. NMR* **2000**, *17*, 153−65.
(7) Romaniuk, P. J.; Uhlenbeck, O. C. *Methods Enzymol.* **1983**, *100*, 52−59.
(8) Hutchins, C. J.; Rathjen, P. D.; Forster, A. C.; Symons, R. H. *Nucleic Acids Res.* **1986**, *14*, 3627−3640.
(9) Price, S. R.; Ito, N.; Oubridge, C.; Avis, J. M.; Nagai, K. *J. Mol. Biol.* **1995**, *249*, 398−408.
(10) Sampson, J. R.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1033−1037.
(11) Beales, L. P.; Rowlands, D. J.; Holzenburg, A. *RNA* **2001**, *7*, 661−670.
(12) Spahn, C. M.; Kieft, J. S.; Grassucci, R. A.; Penczek, P. A.; Zhou, K.; Doudna, J. A.; Frank, J. *Science* **2001**, *291*, 1959−1962.
(13) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.

JA026647W