Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 44–48, January 1996
Biochemistry

Determining RNA solution structure by segmental isotopic labeling and NMR: Application to *Caenorhabditis elegans* spliced leader RNA 1

JING XU*, JON LAPHAM†, AND DONALD M. CROTHERS*†

Departments of *Molecular Biophysics and Biochemistry and †Chemistry, Yale University, New Haven, CT 06511

Contributed by Donald M. Crothers, September 21, 1995

ABSTRACT **Recent developments in multidimensional heteronuclear NMR spectroscopy and large-scale synthesis of uniformly ^{13}C- and ^{15}N-labeled oligonucleotides have greatly improved the prospects for determination of the solution structure of RNA. However, there are circumstances in which it may be advantageous to label only a segment of the entire RNA chain. For example, in a larger RNA molecule the structural question of interest may reside in a localized domain. Labeling only the corresponding nucleotides simplifies the spectrum and resonance assignments because one can filter proton spectra for coupling to ^{13}C and ^{15}N. Another example is in resolving alternative secondary structure models that are indistinguishable in imino proton connectivities. Here we report a general method for enzymatic synthesis of quantities of segmentally labeled RNA molecules required for NMR spectroscopy. We use the method to distinguish definitively two competing secondary structure models for the 5′ half of *Caenorhabditis elegans* spliced leader RNA by comparison of the two-dimensional $\{^{15}N\}^1H$ heteronuclear multiple quantum correlation spectrum of the uniformly labeled sample with that of a segmentally labeled sample. The method requires relatively small samples; solutions in the 200–300 μM concentration range, with a total of 30 nmol or ≈40 μg of RNA in ≈150 μl, give strong NMR signals in a short accumulation time. The method can be adapted to label an internal segment of a larger RNA chain for study of localized structural problems. This definitive approach provides an alternative to the more common enzymatic and chemical footprinting methods for determination of RNA secondary structure.**

Recent developments in the synthesis of uniformly ^{13}C-/^{15}N-labeled RNA (1, 2) have advanced the application of multidimensional heteronuclear NMR (3, 4) to RNA structural studies. For small RNA molecules, it is relatively rare to find that both the ^1H and ^{15}N chemical shifts of two imino or amino groups are degenerate or that both the ^1H and ^{13}C chemical shifts of two CH groups are degenerate, so uniform labeling is usually sufficient to resolve the ambiguities caused by overlap in ^1H spectroscopy.

However, when dealing with RNAs larger than ≈35 nucleotides, the limitations of uniform labeling become apparent because of two fundamental problems. The first is associated with the twin difficulties of extensive spectral overlap due to the larger number of resonance and poorer intrinsic spectral resolution because of the increasing rotational correlation time. Many spectral editing techniques have been developed to simplify the spectra further (5).

The second problem, which will be the focus of this paper, reflects the increased diversity of secondary structures that are accessible to longer and more complicated RNAs (6). For short oligonucleotides, such diversity is very limited and usually there is only one plausible secondary structure. For large oligonucleotides, however, there may be several possible secondary structures for one defined sequence. The problems caused by this diversity can be subdivided into two classes. First, multiple conformations of the RNA may exist, complicating the spectra and making their interpretation difficult. Therefore, for NMR studies, every effort is made in sample preparation to ensure that only one major conformation exists under the conditions studied. In the second class, only one major conformation exists but it may be difficult to define the correct secondary structure based on the exchangeable proton spectra. The difficulties in this second case arise from two independent sources—namely, spectral overlap, which is frequently encountered, and imino pathway degeneracy, which refers to the situation in which two distinct secondary structures have indistinguishable imino proton connectivities. If a defined sequence has two secondary structures that are degenerate in their imino pathway, there is no way to distinguish them based on a conventional exchangeable proton two-dimensional (2D) nuclear Overhauser effect spectroscopy (NOESY) spectrum.

One approach to resolving these problems is isotopic labeling of a specific region of an RNA, which we refer to as segmental labeling. Only a subset of the imino protons will show up in the spectrum of such a sample, since proton resonance signals can be selected according to their covalently bonded isotopes by spectral editing techniques. Here we describe a procedure (Fig. 1) for making segmentally labeled RNA samples in up to milligram quantities, as required for NMR studies. First, milligram amounts of uniformly labeled and unlabeled RNA were obtained by *in vitro* T7 RNA polymerase transcription (7). Second, the full-length RNA molecules were sequence-specifically cleaved by RNase H (refs. 8 and 9; J.L. and D.M.C., unpublished work), using as a cleavage guide a complementary 2′-*O*-methyloligonucleotide containing a tetradeoxynucleotide. Finally, the labeled and corresponding unlabeled segments were recombined by ligation using T4 DNA ligase and a cDNA template (11).

The utility of segmental labeling in NMR studies was demonstrated by determination of the secondary structures of the 5′ half of *Caenorhabditis elegans* spliced leader RNA 1 (5CESL) (Fig. 2), an important component in trans-splicing in *C. elegans* (14, 15). Although there have been many studies on the process of trans-splicing (14, 16), little is known about the structure and interactions in the trans-spliceosome. Our objective is to study the structure of SL RNA as a first step toward understanding the structure/function relationship. For this purpose, it would be ideal to solve the structure of the full-length *C. elegans* SL RNA 1. However, the length of the RNA is 105 nucleotides, too large to be solved by NMR.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: 2D, two-dimensional; NOE(SY), nuclear Overhauser effect (spectroscopy); SL, spliced leader; 5CESL, 5′ half of *Caenorhabditis elegans* spliced leader RNA; HMQC, heteronuclear multiple quantum correlation.

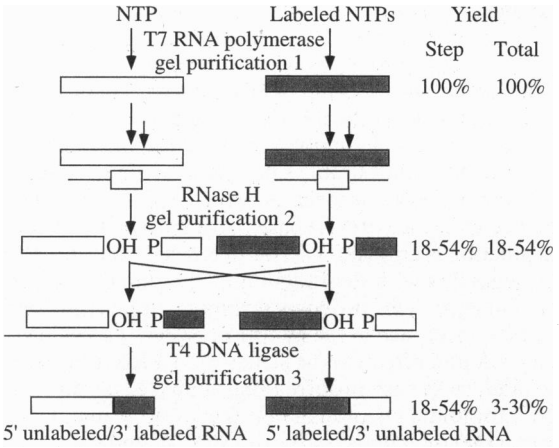

FIG. 1. Flow chart for synthesizing segmentally labeled RNA samples.

Therefore, we chose 5CESL as the subject for NMR studies for the following reasons: (*i*) This segment covers the first stem–loop region in the proposed secondary structure of full-length *C. elegans* SL RNA 1 (15), which most likely involves the intramolecular interactions at the spliced site. (*ii*) The RNA fold program (6) predicts the same hairpin model (Fig. 2*A*) as the most stable secondary structure for this segment alone. This information and an earlier study on trypanosome SL RNA (17) indicated that this segment is likely to be structurally independent of the 3′ half of the SL RNA 1. (*iii*) The first stem–loop has two competing secondary structures in trypanosome spliced leader RNAs (17), suggesting caution in assigning a secondary structure in the *C. elegans* case. Our objective in this work was to characterize definitely the secondary structure of 5CESL. The results document the ability of straightforward NMR measurements to define unambiguously the secondary structure of RNA molecules.

MATERIALS AND METHODS

Materials. [^{15}N]Ammonium chloride (98+%) was obtained from Cambridge Isotope Laboratories (Cambridge, MA). RNase-free DNase I was obtained from Boehringer Mannheim. Inorganic pyrophosphatase, nuclease P1, phospho*enol*-pyruvate, myokinase from chicken muscle, guanylate kinase from porcine brain, nucleoside monophosphate kinase from beef liver, and pyruvate kinase from rabbit muscle were obtained from Sigma. RNase H from *Escherichia coli* was obtained from Pharmacia or Sigma. T4 DNA ligase was obtained from New England Biolabs. Nucleic acid purification (NAP) columns were obtained from Pharmacia. Microcon, Centricon, and Centriprep 3-kDa and 10-kDa cutoff centrifugal concentrators were obtained from Amicon. Affi-Gel 601 boronate-derivative polyacrylamide gel came from Bio-Rad. The Vydac nucleotide analysis column was from Rainin (Woburn, MA). NMR tubes were from Wilmad (Buena, NJ) and Shigemi (Tokyo). The 2′-*O*-methyloligonucleotide RNA–DNA chimera was ordered from the Keck Biotechnology Research Laboratory at Yale University.

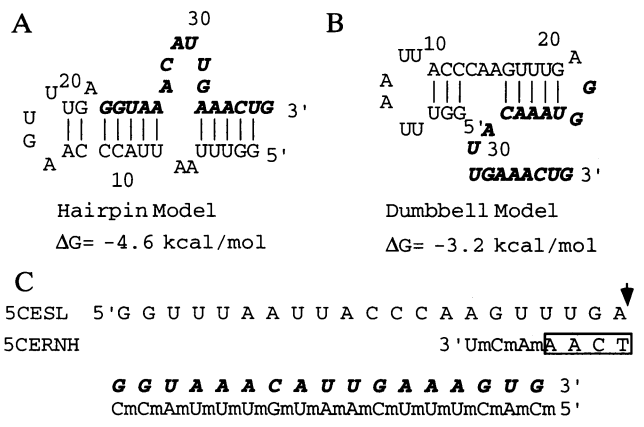

FIG. 2. Alternative secondary structures of 5CESL. (*A*) Hairpin model. (*B*) Dumbbell model. The MFOLD program, version 2.0 by Zuker and Jaeger (6, 12, 13), was used for devising the possible secondary structures. The ΔG values at 25°C were calculated. The boldface, italic segment was ^{15}N-labeled in the segmentally labeled sample. (*C*) Sequence of the chimeric 2′-*O*-methyloligonucleotide 5CERNH. Arrow, expected cutting site by RNase H.

Preparation of [^{15}N]NTPs. [^{15}N]NTPs were prepared as described (1, 2), except for some modifications in cell extraction. Approximately 10 g of *E. coli* cells were harvested from a 1-liter culture. The cell pellet was resuspended in 10 ml of STE (100 mM NaCl/10 mM Tris·HCl/1 mM EDTA, pH 8.0). A solution of 10% SDS was added to the cell suspension to a final concentration of 0.5%. The cell suspension was then ruptured thoroughly with a Branson sonifier 450 sonicator, followed by a 10-min incubation on ice. STE equilibrated phenol was preheated to 65°C and 5 ml was added to the cell lysate. The mixture was incubated at 65°C for 30 min with periodic vigorous shaking and then 5 ml of 24:1 chloroform/isoamyl alcohol was added. The mixture was Vortex mixed followed by centrifugation at 6000 rpm in a Beckman JA-10 rotor for 10 min. The aqueous phase was removed and the organic phase was extracted once with 4 ml of STE. The two aqueous phases were pooled, extracted once with an equal volume of 24:1 chloroform/isoamyl alcohol, and then ethanol precipitated. The precipitated pellet was lyophilized to dryness and resuspended in a minimal amount of autoclaved H_2O. A typical yield of 5000 A_{260} units was obtained from 10 g of wet packed cells. The nucleic acids were digested with P1 nuclease, separated into ribo- and deoxyribonucleotides, and triphosphorylated as described (1, 2).

Preparation of Uniformly Labeled/Unlabeled RNAs. RNAs were synthesized by *in vitro* transcription with T7 RNA polymerase using synthetic oligonucleotide templates as described (7). T7 RNA polymerase was isolated from *E. coli* strain BL21/PAR1219 according to a procedure of Grodberg and Dunn (18). Deoxyoligonucleotides were synthesized on an Applied Biosystems model 380B DNA synthesizer. Typically, 40–50 ml of transcription gave 0.4 μmol of unlabeled 5CESL RNA after gel purification and 50–60 ml of transcription gave 0.4 μmol of isotopically labeled RNA.

Site-Specific RNA Hydrolysis. This method was derived from a previously described procedure (refs. 8 and 9; J.L. and D.M.C., unpublished work). The sequence of the 2′-*O*-methyloligonucleotide–DNA chimera used is shown in Fig. 2*C*, as is the predicted RNase H cutting site. The synthesized chimera was desalted by NAP25 column and concentrated to ≈6 mM by Speed-Vac. The RNA (≈3 mM) and the corresponding chimera (≈6 mM) were mixed at a molar ratio of 1:1.4 to ensure that all the RNAs were bound by the chimera, heated to 90°C, and slowly cooled to 4°C. The reaction mixtures included 20 mM Hepes·KOH (pH 8.0), 50 mM KCl, 10 mM $MgCl_2$, 1 mM dithiothreitol (DTT), RNA–chimera complex (≈3 mM), and RNase H (32 units/μl). The mixture was incubated at 37°C for 2 h or more. The progress of the reaction was monitored by electrophoresis on a Hoefer Scientific minigel. The reaction efficiency was ≈90% for 5CESL RNA. The products were purified by PAGE without prior desalting, recovered by extensive crush-and-soak, concentrated, and desalted with a Centricon concentrator. The maximum recovery rate was 70% for the chimera and 60% for RNA segments.

Cross Ligation and Purification of Segmentally Labeled Oligonucleotides. The labeled segment and the corresponding unlabeled segment were ligated by T4 DNA ligase (11) using as a cDNA template the DNA used for transcription. The mixture of labeled segment, corresponding unlabeled segment, and the DNA template (1:2:1.5) was heated to 90°C and slowly cooled to 4°C. The ligation reaction mixtures included 2–3 mM annealed complex, 50 mM Tris·HCl, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 50 mg of bovine serum albumin per ml, and 40,000 units of T4 DNA ligase per ml. The mixture was incubated at room temperature for 24–72 h. Ligation efficiencies were 60–90%. Higher concentrations of T4 DNA ligase did not accelerate the reaction. The reaction mixture was purified by 20% denaturing PAGE. The highest recovery was 85% for DNA template and 60% for RNA.

NMR Sample Preparation. All the NMR samples were placed in 10 mM sodium phosphate, 50 mM NaCl, and 0.05 mM EDTA (pH 6.4) buffer either by dialysis or by Centricon centrifugation. The unlabeled samples were dissolved in a vol of 350 μl in 5-mm Wilmad NMR tubes; the uniformly or segmentally labeled samples were in a vol of 140 μl in 5-mm Shigemi tubes.

NMR Spectroscopy. All experiments were performed on a GE Omega 500 spectrometer with *x, y, z* pulsed-field gradients using a 5-mm Bruker carbon, nitrogen, phosphorus triple-resonance probe. All the NMR data were processed on Silicon Graphics computers using the FELIX program from Biosym Technologies (San Diego). One-dimensional ^1H experiments were acquired with a gradient-enhanced jump-return spin-echo sequence for water suppression (19, 20). The 2D NOESY spectra in H_2O were obtained in the pure absorption mode (21) using a standard pulse sequence (22), except that the last 90° pulse was replaced with a gradient-enhanced jump-return spin-echo sequence for water suppression (19, 20). The 2D $\{^{15}N\}^1H$ heteronuclear multiple quantum correlation (HMQC) was collected for imino protons using a gradient-enhanced jump-return spin-echo pulse sequence (23).

RESULTS AND DISCUSSION

Preparation of Segmentally Labeled Oligonucleotides. The challenge of segmental labeling is to make such a sample in the quantities required for NMR studies. Although enzymatic methods to ligate segments of RNA in picogram quantities are available (11, 24), scaling up to NMR quantities is not trivial. The simplest approach to producing a segmentally labeled RNA would be to transcribe the two halves of the RNA separately and then ligate them. However, T7 RNA polymerase has a strong dependence of transcription yield on the 5′ end sequence of the RNA (25). This limitation on polymerase is usually circumvented by modification of the 5′ end sequence of the RNA being studied to achieve a good transcription yield. When preparing the 3′ half of the RNA, such an approach is generally impractical because any modification made in the 5′-end sequence of the 3′ segment will be incorporated into the middle of the RNA molecule after ligation. Also, a large (often up to 50%) portion of the transcription product varies in length at the 3′ end. If this occurs for the 5′ RNA used in a ligation reaction, the yield will drop significantly because only the correct length RNA will be used in the ligation reaction. In the procedure described here, an adaptation of the site-specific RNase H cleavage reaction (refs. 8 and 9; J.L. and D.M.C., unpublished work) alleviated these problems. The products of the RNase H cleavage are a 3′-hydroxyl and a 5′-monophosphate at the hydrolysis site, which are the correct substrates for subsequent ligation. Although the sequence and secondary structure of the RNA might affect RNase H cleavage (9), our experience so far has not yielded examples of such behavior.

The success of the procedure also relies on the fact that DNA template-directed RNA ligation retains accuracy and high efficiency (up to 90%) when scaling up to NMR quantities. The advantage of using T4 DNA ligase (11) over T4 RNA ligase (24) is at least 2-fold. First, since T4 DNA ligase will ligate junctions only in double-stranded regions, there is little tendency to have undesirable side products such as circularized and oligomeric RNA molecules, which appear with T4 RNA ligase. Second, since precise base pairing to the DNA template at the junction is critical for RNA–RNA ligation with T4 DNA ligase, segments with degradation at the junction site will be selected against. This property is important because even one nucleotide difference in the middle of an RNA sequence could have a profound effect on the structure. T4 RNA ligase, on the other hand, lacks such proofreading ability. Ligation efficiency is higher for joining two small nonstructured segments than for ligating larger segments with extensive secondary structure.

However, yield is still an important factor in the application of this procedure. The stepwise yields in Fig. 1 reflect the range of experience we have had, and the range of total yields are products of the yields per individual step. We do not in general find serious losses in the RNase H cleavage or T4 DNA ligase joining steps, but there is a significant loss of product during each gel purification. Chromatographic methods could in principle be used to purify RNA after transcription, but we had limited success getting single nucleotide resolution with this approach, causing difficulty in distinguishing the correct product band after RNase H cutting. Furthermore, efficient chromatographic methods to replace the second and the third gel purification steps are difficult to find because strongly denaturing conditions are required to disrupt either the chimera–RNA or the DNA–RNA complex. Fortunately, extensive crush-and-soak extraction from gel slices increases the recovery rate of RNA to ≈60% in each step, which leads to a maximum total yield of ≈30%. It is obvious that efficient transcription can offset the yield problem caused by loss of product during gel purification. Relatively small amounts of sample are required because the $\{^{15}N\}^1H$ HMQC experiment is very sensitive; high quality spectra can be obtained in 2–3 h of NMR time using a 0.2 mM sample (30 nmol or ≈40 μg of RNA in 140 μl). This reduces the labor and expense in sample preparation if the only purpose of segmental labeling is to define the secondary structure unambiguously. Since ligation of more than two pieces of RNA is possible with T4 DNA ligase, an RNA sample in which only an internal portion is labeled while the rest is unlabeled can be obtained in a similar way. In summary, the procedure we described here can be generally used for the purpose of isotopic labeling of a specific region of an RNA molecule in milligram quantities.

Application to the Determination of the Secondary Structure of 5CESL. The usual first objective in an RNA structural study is to define the secondary structure. (We found for 5CESL that standard enzymatic digestion methods were complicated by formation of RNA dimers under the divalent ion conditions required for most of the digestion reactions.) As mentioned before, a hairpin model (Fig. 2*A*) was predicted as the most stable secondary structure. However, a parallel structural study on the 22-nucleotide 5′ miniexon segment (15) showed that it can form an unexpectedly stable stem–loop structure (J.X., unpublished data). The derivative melting profile of 5CESL has two peaks, one at 41°C and the other at 50°C (in 10 mM sodium phosphate, pH 6.4/50 mM NaCl/0.05 mM EDTA; data not shown). The melting temperature of the miniexon in the same buffer is 48°C (J.X., unpublished data), which is very close to the higher melting peak of 5CESL. This observation forces consideration of a suboptimal secondary structure for 5CESL, a "dumbbell" model, which incorporates the likely secondary structure of the miniexon (Fig. 2*B*). Although the calculated free energies suggest that the dumbbell model is less stable than the hairpin

model, the credibility of the predictions is limited because the energy data set used in the calculation still needs improvement.

Distinguishing the two possible secondary structures for 5CESL by conventional NMR was difficult. In the hairpin model, two imino pathways, G2–U3–U4–U5 and U8–U9–U24–G23–G22, were expected, whereas in the dumbbell model, two different imino pathways, G1–G2–U3 and G16–U17–U18–U19, were expected. However, these two sets of imino pathways both have G–G–U and G–U–U connectivities. If some of the imino protons exchange rapidly with solvent water, these two structures could be degenerate in their imino pathways. This turned out to be the case. There were 10 imino peaks in the one-dimensional proton spectrum (Fig. 3). Imino protons 1, 2, 6, and 8 (numbered by their increasing resonance frequency) were very unstable with increasing temperature and disappeared by 20°C, indicating fast exchange of those protons with water. Only two short imino walks could be identified, G(7)–G(5)–U(3) and G(4)–U(1)–U(2), based on the 2D H_2O NOESY spectrum (Fig. 4). The observed intensity of a NOE peak between two imino protons depends on how close they are and how stable the imino protons are. A longer mixing time is advantageous for developing NOEs, but imino proton exchange with solvent protons decreases the intensity of the signal detected. Various mixing times from 85 to 250 ms were tried at 0.5°C. The crosspeak between U(1) and U(2) was observable only when an 85-ms mixing time was used because the U(2) imino proton exchanges quickly. We could not identify the connectivities of U(6) and U(8)/G(9/10) to other protons, most likely because of rapid exchange and the difficulty in finding an optimal mixing time. The G·U base pair resonances could arise either from G1·U37 in the hairpin model or from G20·U24 in the dumbbell model.

The {^{15}N}^{1}H HMQC of uniformly labeled 5CESL (Fig. 5*A*) clearly identified the base type of each imino proton. Particularly, U(3) and G(4), which were difficult to identify in the 2D H_2O NOESY because of the spectral overlap, were well resolved by ^{15}N chemical shift. U(6) was clearly identified as a U imino proton. G(9) and G(10), which had almost identical chemical shifts, were also resolved by dispersion in the nitrogen dimension. This experiment complemented the assignment of the 2D H_2O NOESY. However, it did not provide sufficient additional information to distinguish the two secondary structural models.

Analysis of the data and the sequence suggested that segmental labeling of the molecule might solve the problem. Only two segments in the whole sequence, G1–G2–U3 and G22–G23–U24, could give rise to the G(7)–G(5)–U(3) connectivity observed in 2D H_2O NOESY. Dividing the whole molecule into two halves, G1–A21 and G22–G38, and labeling only one half should give us a clear answer (Fig. 3). For instance, assuming that the hairpin model is correct, when only the 3′ half of the molecule (G23–G38) is ^{15}N labeled, the U(3), G(5), G(7), and U(8) imino protons will be labeled by ^{15}N

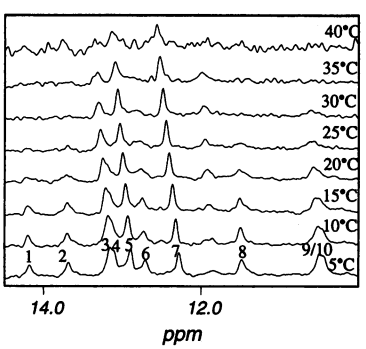

FIG. 3. Temperature dependence of imino spectra of 5CESL. The 10 imino protons observed at 5°C are numbered.

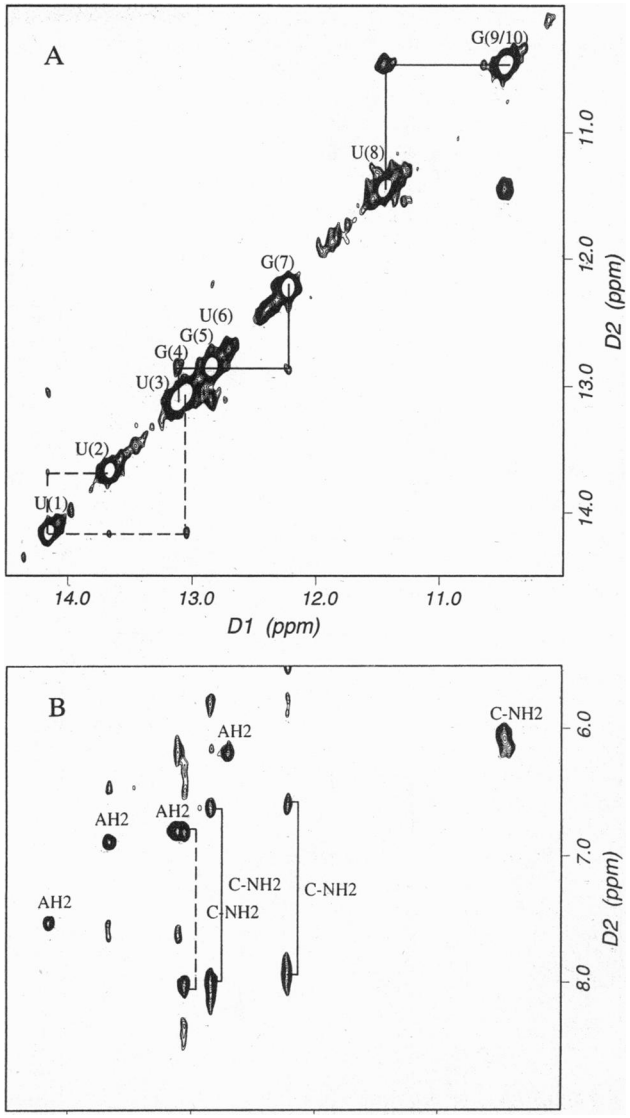

FIG. 4. A portion of the exchangeable proton 2D H_2O NOESY spectrum of 5CESL showing imino–imino (*A*) and imino–amino (*B*) regions. The two imino–imino sequential connectivities are shown by solid and dashed lines. Spectrum was acquired in 22 h at 0.5°C. The acquisition parameters included 2048 complex points in t2, 240 t1 increments, 128 scans per t1 increment, and a mixing time of 85 ms. The sample concentration was 1.2 mM. Prior to Fourier transformation the free induction decays (FIDs) were apodized with a sinebell window function (size = 64) to remove the residual water signal (26) and a sinebell square window function (75°, 512 points). The spectrum was apodized with a sinebell square window function (75°, 64 points) and zero-filled in t1 to give a final 2K × 2K real data matrix.

while the U(1), U(2), and G(4) imino protons will not; if the dumbbell model is correct, U(3), G(5), and G(7) will not be labeled by ^{15}N. As shown in the 2D {^{15}N}^{1}H HMQC spectrum of U5/^{15}N3 5CESL (Fig. 5*B*), only the U(3), G(5), G(7), and U(8) cross peaks were retained. Therefore, the hairpin model is the correct secondary structure for 5CESL.

Conclusion. In this study, we were able to distinguish two distinct secondary structures of a defined sequence with degenerate imino pathways. Although such problems have not to our knowledge been reported before, they will probably not be uncommon when longer RNA sequences are studied by NMR. The inherent advantage of segmental labeling is that one knows which portion of the primary sequence gives rise to the signal, thus dramatically reducing the complexity in data interpretation.

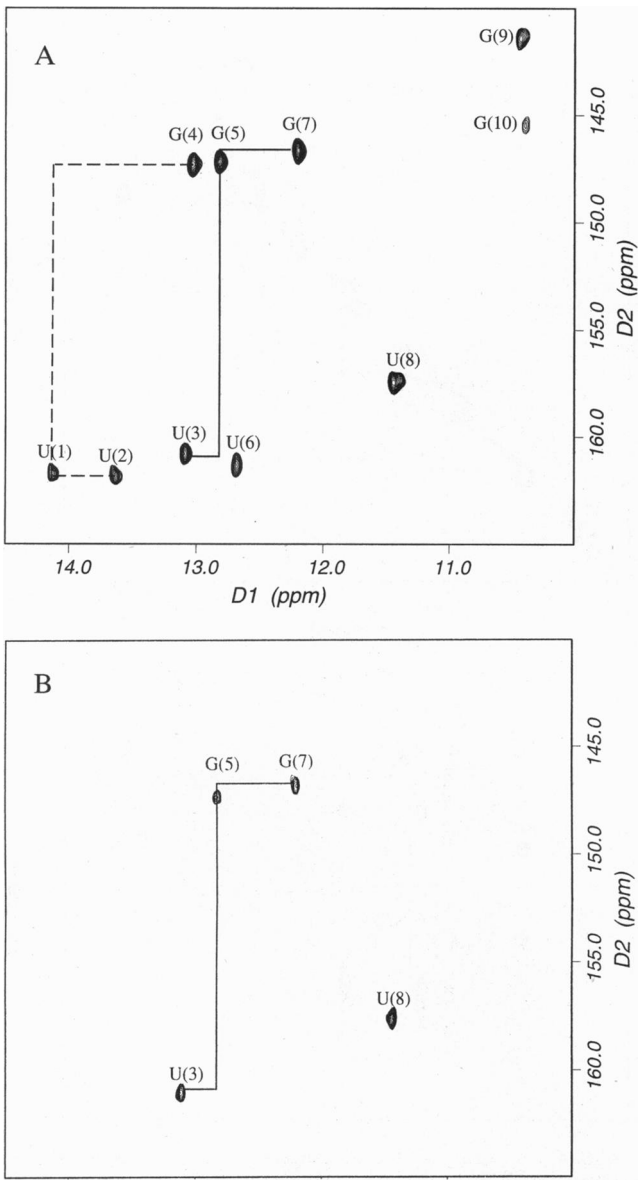

FIG. 5. Imino regions of the 2D $\{^{15}N\}^1H$ HMQC spectrum of 5CESL with uniformly labeled (*A*) and 5′ unlabeled and 3′-labeled (*B*) 5CESL. Only the resonances connected by the solid line in the water NOESY show up in *B*. Sample concentration was 0.3 mM for the uniformly labeled molecule and 0.2 mM for the segmentally labeled molecule. Spectra were acquired at 1°C with 1024 complex points in t2 and 64 t1 increments, a spectral width of 10,000 Hz in t2 and 2000 Hz in t1. The proton dimension offset was set at the water frequency (4.75 ppm). The nitrogen offset was set at 153 ppm and the ^{15}N-^1H one-bond coupling constant was assumed to be 95 Hz. For the uniformly labeled sample, 112 scans were performed per t1 value; for the segmentally labeled sample, there were 160 scans per t1 value. Prior to Fourier transformation, the time domain data were apodized with a sinebell squared window function in t2 (90°, 512 points) and in t1 (90°, 64 points). The spectrum was zero-filled in t1 to give a final 1K × 1K real data matrix.

Segmental labeling, combined with spectral editing techniques, can be expected to simplify the NMR spectrum of larger RNAs. Although the RNA studied in this paper is <40 nucleotides, the procedure described can be adapted for segmental labeling of much longer RNA molecules. By dividing a complicated spectrum into simpler subspectra, longer and more complicated RNAs that cannot be studied because of severe spectral overlap will be accessible to NMR studies, assuming that the correlation time remains short enough to allow data acquisition. (Given that RNA molecules can form domains that may move with considerable independence, it is difficult to predict the molecular size at which this strategy will fail because of a long correlation time.) Thus, a potentially promising application of segmental labeling is to focus on an interesting portion of a larger RNA molecule. It may be common to find that the secondary structure of an oligonucleotide changes when the context changes (10). Therefore, caution always has to be exercised when a domain is isolated for structural studies. Using the segmental labeling approach, it may be possible to observe interesting domains of large RNA molecules in their native structural environment. Identification of a specific chain segment with specific hydrogen-bonded imino protons leaves in general little ambiguity regarding the RNA secondary structure, in contrast to the currently more widely accessible footprinting methods for studying the same problem.

We thank Dr. A. Sitlani for reading the manuscript. J.X. would like to thank G. Sun for assistance in DNA synthesis, D. Jeruzalmi for providing some of the T7 RNA polymerase, D. P. Zimmer for help in preparation of labeled nucleotides, and J. P. Marino for NMR training. This work was supported by National Institutes of Health Grant GM 21966 to D.M.C.

1. Batey, R. T., Inada, M., Kujawinski, E., Puglisi, J. D. & Williamson, J. R. (1992) *Nucleic Acids Res.* **20,** 4515–4523.
2. Nikonowicz, E. P., Sirr, A., Legault, P., Jucker, F. M., Baer, L. M. & Pardi, A. (1992) *Nucleic Acids Res.* **20,** 4507–4513.
3. Clore, G. M. & Gronenborn, A. M. (1991) *Prog. Nucl. Magn. Reson. Spectrosc.* **23,** 43–92.
4. Nikonowicz, E. P. & Pardi, A. (1992) *Nature (London)* **355,** 184–186.
5. Jaeger, J. A., ScantaLucia, J. J. & Tinoco, I. J. (1993) *Annu. Rev. Biochem.* **62,** 255–287.
6. Jaeger, J. A., Turner, D. H. & Zuker, M. (1989) *Methods Enzymol.* **183,** 262–288.
7. Milligan, J. F. & Uhlenbeck, O. C. (1989) *Methods Enzymol.* **180,** 51–62.
8. Inoue, H., Hayase, Y., Iwai, S. & Ohtsuka, E. (1987) *FEBS Lett.* **215,** 327–330.
9. Hayase, Y., Inoue, H. & Ohtsuka, E. (1990) *Biochemistry* **29,** 8793–8797.
10. Bhattachyra, A. & Blackburn, E. H. (1994) *EMBO J.* **13,** 5721–5731.
11. Moore, M. J. & Sharp, P. A. (1992) *Science* **256,** 992–997.
12. Zuker, M. (1989) *Science* **244,** 48–52.
13. Jaeger, J. A., Turner, D. H. & Zuker, M. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7706–7710.
14. Nilsen, T. W. (1993) *Annu. Rev. Microbiol.* **47,** 413–440.
15. Bruzik, J. P., van Doren, K., Hirsh, D. & Steitz, J. A. (1988) *Nature (London)* **335,** 559–562.
16. Agabian, N. (1990) *Cell* **61,** 1157–1160.
17. LeCuyer, K. & Crothers, D. M. (1993) *Biochemistry* **32,** 5301–5311.
18. Grodberg, J. & Dunn, J. J. (1988) *J. Bacteriol.* **170,** 1245–1253.
19. Plateau, P. & Gueron, M. (1982) *J. Am. Chem. Soc.* **104,** 7310–7311.
20. Sklenar, V. & Bax, A. (1987) *J. Magn. Reson.* **74,** 469–479.
21. States, D. J., Haberkorn, R. A. & Ruben, D. J. (1982) *J. Magn. Reson.* **48,** 286–292.
22. Kumar, A., Ernst, R. R. & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.* **95,** 1–6.
23. Szewczak, A. A., Kellogg, G. W. & Moore, P. B. (1993) *FEBS Lett.* **327,** 261–264.
24. Romaniuk, R. J. & Uhlenbeck, O. C. (1983) *Methods Enzymol.* **100,** 52–59.
25. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15,** 8783–8798.
26. Marion, D., Ikura, M. & Bax, A. (1989) *J. Magn. Reson.* **84,** 425–430.