Journal of the Korean Magnetic Resonance Society **2001**, *5*, 29-36

Application of multi dimensional NMR experiments to VBS RNAs of Yeast *Saccaromyces cerevisiae* virus

Hae-Kap Cheong, Jun-Seok Yoo, Chaejoon Cheong*

Magnetic Resonance Team, Korea Basic Science Institute, Taejon 305-333, Korea
Received March 29, 2001

Abstract : The structures of two VBS (viral binding site) RNAs, SL1 and SL2, of Yeast *Saccharomyces cerevisiae* virus have been studied by 2D and 3D NMR experiments. VBSs play a crucial role in viral particle binding to the plus strand and packaging of the RNA. The secondary structures of the two VBS RNAs share a common feature of the stem-internal loop-stem-hairpin loop structure although the size of the internal loops of SL1 and SL2 differs. 2D experiments were sufficient for full assignments of SL1. However, isotope labeling of the sample and multidimensional experiments were required for 28-nucleotide-long SL2 due to the spectral overlap. Several 3D HCCH experiments have accomplished full assignment of SL2 RNA.

INTRODUCTION

RNA plays important roles in biological processes, including propagation of genetic information and enzymatic catalysis.[1] Two-dimensional ^{1}H-NMR spectroscopy provides a wealth of structural information on RNA molecules. As the first step in these structural studies, complete spectral assignments must be made. Assigning base and sugar H1' protons is often straightforward in helical regions.[2] However, assignments are difficult in the presence of mismatches and in single-stranded regions.[3] The combination of ^{13}C-^{1}H- and ^{1}H-^{1}H-correlated experiments and NOE experiments enables the identification of each individual spin system. Identifying sequential base-H1' and base-sugar crosspeaks in NOESY spectra is important because it provides the most useful information for the resonance assignments of RNAs. However, spectral overlap in these regions often limits the size of RNAs which can be assigned. Application of multidimensional experiments on labeled RNA sample can provide a way to overcome the size limitation to a great extent.

The yeast double-stranded RNA virus, *Saccharomyces cerevisiae* virus (ScV) has a single essential genomic RNA, L1, of 4.6 kb.[4] ScV-M1, the satellite virus of ScV-L1 also has a dsRNA of about 1.9 kb.[4] The Cap-Pol fusion protein of the *Saccharomyces cerevisiae*

*To whom : cheong@comp.kbsi.re.kr

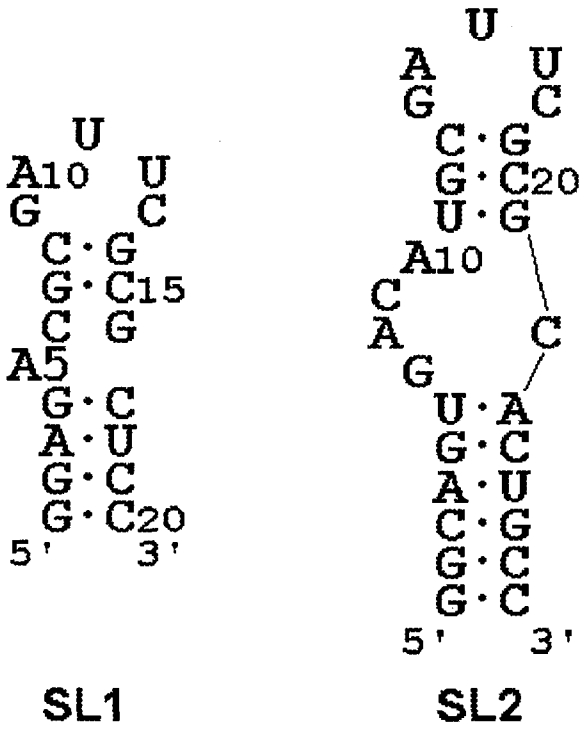

Fig. 1. Sequences and secondary structures of ScV VBSs, SL1 and SL2 RNA.

virus L1 (ScV-L1) recognizes a short sequence of the viral plus strand RNA.[4] This viral binding site (VBS) is responsible for viral particle binding to the plus strand and packaging of the RNA. This region, which is a minimum of 20 bases long, consists of a hairpin with a five-nucleotide loop and a stem interrupted by an internal bulge-loop.[5,6] ScV-L1 has a single VBS, while ScV M1 has two VBS's, SL1 and SL2.[6] It has been suggested that the structure, not the sequence, of the stem, the sequences of the loop, and the bulged A residue are important for the binding of the Cap-Pol fusion protein.[6,7] ScV-M1, which is present in some strains and is separately packaged in ScV particles provided by the L1 virus, encodes only two proteins, a secreted protein toxin (the killer toxin) and the immunity to that toxin.

In this study, we present the NMR studies on the VBS RNAs, SL1 and doubly labeled SL2 using 2D and 3D experiments (See Fig. 1 for the sequences and secondary structures). Complete assignments of the 20 nucleotide RNA, SL1, and 28 nucleotide RNA, SL2, have been accomplished by applying multidimensional NMR techniques.

MATERIALS AND METHOD

Preparation of RNA samples

SL1 (20 nucleotides) and SL2 (28 nucleotides) VBS oligonucleotides were enzymatically synthesized *in vitro* using T7 RNA polymerase and chemically synthesized DNA templates.[8,9] The transcription products were purified using 20% polyacrylamide gel electrophoresis in denaturing condition. The purified oligonucleotides were dialyzed extensively against 10 mM sodium phosphate buffer, pH 6.5, 0.01 mM EDTA. Following dialysis, the sample was lyophilized to dryness. For experiments involving the nonexchangeable protons, the sample was exchanged several times with 99.9% D_2O and then resuspended in 400 μl of 99.96% D_2O. For experiments involving the exchangeable protons, the sample was resuspended in 400 μl of 90% H_2O/10% D_2O.

To prepare the labeled RNA, the labeled nucleotide monophosphates (NMPs) were first enzymatically phosphorylated to the nucleotide triphosphates (NTPs).[10] The completeness of the reaction was monitored by HPLC. After the phosphorylation, the NTPs were ethanol precipitated and directly used in *in vitro* transcription. The transcription yields were comparable to the commercially available NTPs.

NMR spectroscopy

All experiments were carried out using a Bruker DMX600 NMR spectrometer operating at 600.13 MHz proton frequency. NMR data were processed and displayed using the program Xwin-NMR (Bruker). One-dimensional (1D) exchangeable proton spectra were obtained using the jump-return method.[11]

NOESY spectra of the sample in H_2O were acquired at 5, 10, 20 °C. Water suppression was achieved with the jump-return pulse having a 50 us delay between the observe pulses. Approximately 512 FIDs of 2K complex points were collected. The spectral width in both dimensions was 12,000 Hz. NOESY spectra taken in D_2O were acquired with mixing times of 50, 100, 150, 200, and 400 ms primarily at 20 °C, and at 10 °C and 30 °C to resolve overlapped resonances.

Double-quantum filtered COSY (DQF-COSY) spectra were acquired with a spectral width of 4,800 Hz, using the standard pulse sequence.[12] A high-resolution, phosphorus-decoupled, DQF-COSY spectrum was also acquired with a spectral width of 1,800 Hz in both dimensions, 2K complex points in t_2, and 512 points (FIDs) in t_1. TOCSY spectra were acquired with DIPSI-2 sequence for mixing.[13]

Natural abundance heteronuclear ^1H-^{13}C multiple quantum coherence (HMQC) spectra were acquired using the standard pulse sequence.[14] The spectral widths were 4,800 Hz (8 ppm) in ^1H dimension and 15,000 Hz (100 ppm) in ^{13}C dimension. Constant time heteronuclear single quantum coherence (HSQC) spectra were acquired for the labeled RNA in order to remove the carbon-carbon splitting in ^{13}C dimension. Proton-detected ^1H-^{31}P heteronuclear correlation spectra (HETCOR) were acquired as previously proposed.[15] The

spectral widths were 1,800 Hz (3 ppm) in ^{1}H dimension and 580 Hz (2.4 ppm) in ^{31}P dimension.

HCCH-COSY, HCCH-RELAY, HCCH-TOCSY spectra were acquired for the labeled RNA in order to correlate the sugar resonances in 2D and/or 3D modes.[16] The delays for the HCCH experiments were optimized for $^{1}J_{HC}$=160 Hz and $^{1}J_{CC}$=45 Hz. A DIPSI-2 mixing scheme was applied during the ^{13}C TOCSY spin lock period of 27 ms. 2D and 3D NOESY-HSQC or NOESY-HMQC spectra were acquired for the purpose of internuclear spin resonance assignments.

RESULTS AND DISSCUSSION

Assignment of exchangeable protons

The imino and amino proton spectra were assigned by standard methods using one- and two-dimensional NMR spectroscopy in H_2O. SL1 RNA contains seven guanosine and three uridine residues; 1D imino proton spectrum shows 5 sharp resonances. The resonances have been assigned to the imino protons of the stem from the 2D NOESY experiments. The imino proton of 5′-terminal guanosine G1 is not observed due to the fraying effect. G16 next to the bulge A5 does not show imino proton resonance at 1 °C or higher. SL2 RNA contains nine guanosine and five uridine residues; 2D ^{1}H-^{15}N HMQC spectrum contains 7 crosspeaks in the imino region. The secondary structure of SL2 RNA could not be deduced from 2D NOESY spectrum because of the overlapped G5 and U25 imino protons. In the NOESY-HMQC experiment we could observe the NOE crosspeaks between G5 and U25 which were well resolved in nitrogen dimension. In contrast to the secondary structure proposed previously by the biochemical data, U11 and G21 do not form a base pair whereas G7 and C22 do. 4 nucleotides and 1 nucleotide at the 5' and 3'-side, respectively, constitute the internal loop of the SL2 molecule.

Assignment of non-exchangeable protons

Pyrimidine H5 and H6 resonances were identified by their strong cross-peaks in the DQF-COSY spectrum. Cytosines were distinguished from uridines by the chemical shifts of their C5 carbons in the heteronuclear correlated spectroscopy. This information provided a starting point for the assignment of non-exchangeable protons followed by the standard procedure based on the sequential NOE connectivities and throughbond correlations.

H8/H6-H1' connectivities of SL1 are continuous from G1 to A10 and from U11 to C20. No cross peak was observed between A10 H1' and U11 H6 in the 400 ms NOESY. H1'-H2' region of the 50 and 100 ms NOESY and the DQF-COSY spectra were used to assign the H2' resonances. Alternate H8/H6-H2' connectivities confirmed the sequential assignments and resolved some ambiguities in the H8/H6-H1' connectivities. H8/H6-H2' connectivities are continuous for the entire SL1 molecule, except for the A10-U11 step. The remaining sugar protons were assigned by identifying the sugar spin systems using high-

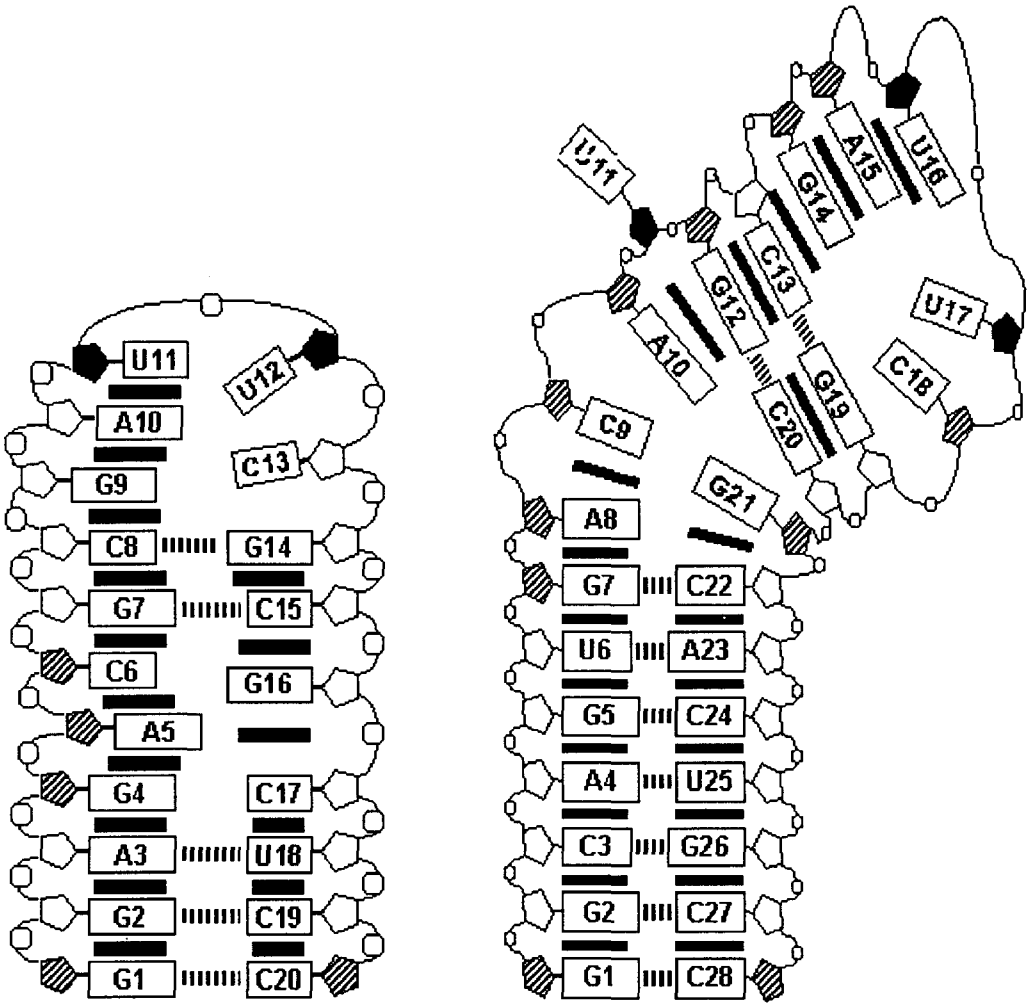

Fig. 2. Schematic summary of the NMR data for SL1 (left) and SL2 (right) RNA. Base pairing is represented by hatched bars and base stacking is represented by solid bars. Sugars are represented by pentagons. Sugar pucker: open pentagon for C3'-endo, filled for C2'-endo, and striped for an equilibrium between C2'- and C3'-endo. Phosphate groups are represented by small circles.

resolution, ^{31}P-decoupled DQF-COSY. The assignment was confirmed by the sequential H8/H6-H3' NOE connectivities and the inter- and intra-H1' sugar proton NOEs. Adenosine H2 resonances were identified by the chemical shifts of the bound carbons in the heteronuclear correlated spectroscopy.

^{1}H-^{31}HETCOR can be used in determining the backbone sequence. The cross peaks in

the spectrum were identified by the previously determined H3' and H5'/H5" assignments. Strong sequential H3'(i)-^{31}P($i+1$) crosspeaks and weak intranucleotide H5'/H5"-^{31}P crosspeaks have been observed. These scalar connectivities unambiguously confirmed the nucleotide sequence, especially that in the loop region where structural assumptions cannot be made.

DQF-COSY spectrum of SL1 shows 10 crosspeaks, which is expected for a single conformation. However, only 13 crosspeaks have been shown in DQF-COSY spectrum of SL2 that has 14 pyrimidine. Although SL2 RNA adopts a single conformation, one pyrimidine nucleotide must be in the flexible or exchanging conformation. The H5-H6 crosspeak which does not show in DQF-COSY exhibits a very broad peak shape in 400 ms NOESY spectrum. The base-sugar sequential connectivity of SL2 continues through the whole molecule except the hairpin loop and internal loop. The cross peak between A15 H1' and U16 H6 was very weak in SL2. Unexpectedly, A10 H8 exhibited an NOE to G12 H1' suggesting that U11 was bulged out and A10 and G12 bases were stacked together.

Figure 2 shows the schematic summary of the NMR data for SL1 and SL2 VBS RNAs.

Multinuclear and multidimensional NMR of SL2 RNA

As the molecular weight increases, the number of resonance peaks increases and spectral overlap becomes one of the major problems of the assignment. The use of ^{13}C and ^{15}N labeled sample and multidimensional NMR spectroscopy can solve the peak overlap problem.

HMQC or HSQC were used to separate some overlapped resonances by the ^{13}C and ^{15}N chemical shifts. HMQC was mainly used for natural abundance experiments and constant time HSQC was used for ^{13}C labeled sample to eliminate carbon-carbon coupling. In ^{1}H-^{13}C HMQC/HSQC spectrum, uridine C5 is 5-10 ppm downfield shifted from cytidine C5 and adenine C2 is about 10 ppm downfield shifted from purine C8 and pyrimidine C6. Carbon chemical shifts of sugars are more dispersed than those of protons; C1' resonates in the range of 85-95 ppm, C2'/C3' in70-79 ppm, C4' in80-87 ppm and C5' in 60-69 ppm. For the exchangeable protons, ^{1}H-^{15}N correlation experiments can distinguish G imino (~150 ppm) from U imino (~160 ppm) resonances.

In the NMR studies of RNA, the sugar resonance assignment is the most challenging work because the sugar proton resonances overlap in the narrow region between 4 and 5 ppm. HCCH experiments that use the proton-carbon-carbon-proton magnetization transfer played a crucial role in sugar resonance assignment of SL2. Figure 2A shows an example of HCCH experiments on the uniformly ^{13}C labeled SL2 RNA. The coupling constants of ^{1}H-^{13}C (~180 Hz) and ^{13}C-^{13}C (~40 Hz) are much larger than the ^{1}H-^{1}H coupling constant (<10 Hz) and are not affected by the sugar puckering conformation. Figure 3B shows a strip of HCCH-COSY, HCCH-RELAY and HCCH-TOCSY experiments. In HCCH-COSY spectrum, the H1'-H2' correlations of SL2 are shown; they were not observed in DQF-

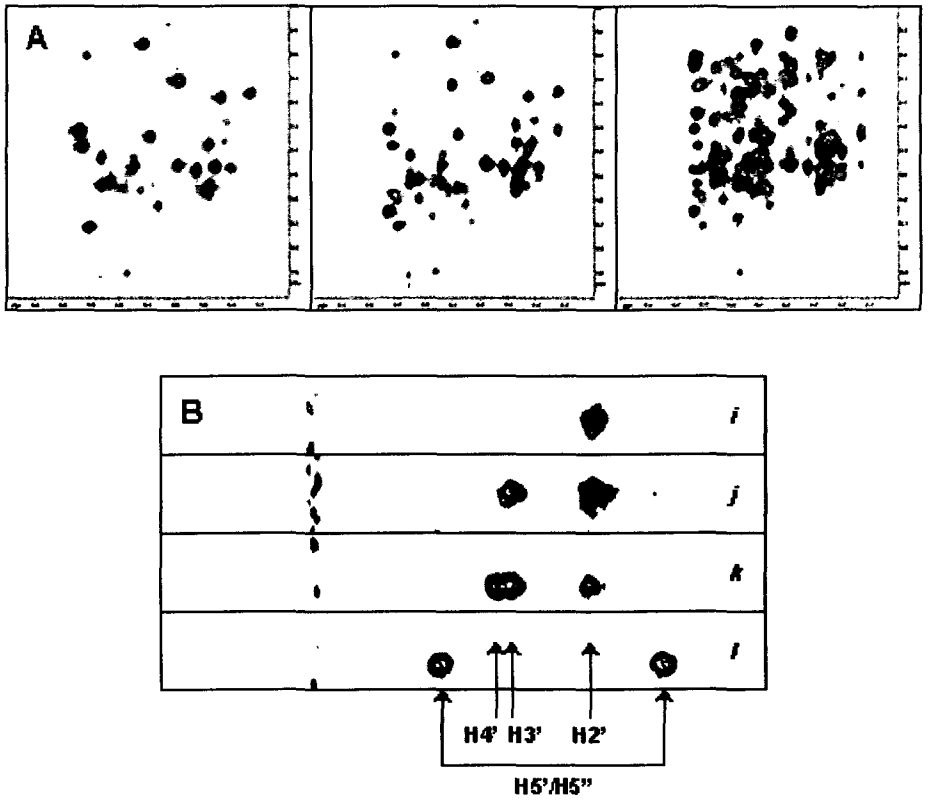

Fig. 3. (A) H1' (ω_2) to H2' through H5'/H5" (ω_1) region of HCCH-COSY (*left*), HCCH-RELAY (*middle*), and HCCH-TOCSY (*right*) spectra on the uniformly ^{13}C labeled SL2 RNA. (B) Strip plots of H1' region for C24 of HCCH-COSY (*i*) , HCCH-RELAY (*j*), HCCH-TOCSY methylene filtered (*k*), and HCCH-TOCSY methylene selected (*l*) spectra.

COSY or TOCSY because of small coupling constant of H1'-H2'. In HCCH-RELAY spectrum, the H1'-H3' as well as H1'-H2' correlations were observed. In HCCH-TOCSY, the methylene filtered experiment showed the H1'-H4' correlation of SL2 as well as H1'-H2' and H1'-H3' correlations, whereas the methylene selected experiment showed the H1'-H5' and H1'-H5" correlations of SL2.

3D HCCH experiments on SL2 made it possible to identify complex sugar spin systems. The large (~30 ppm) carbon chemical shift dispersion greatly simplified the assignment of sugar resonances. From the chemical shift of the carbon, the sugar resonances could be easily assigned. The structural modeling of SL1 and SL2 RNAs is in progress and the results with the biological implication will be published elsewhere.

Acknowledgement

This work was financially supported by the National Research Laboratory program of the Ministry of Science and Technology, the Republic of Korea.

REFERENCES

1. R. E. Gesteland and J. F. Atkins, *The RNA World*, Cold Spring Harbor Laboratory Press, New York , 1993.
2. G. Varani and I. Tinoco, JR, *Quart. Rev. Biophys.*, **24**, 479 (1991).
3. H. K. Cheong, C. Cheong, Y. S. Lee, B.L. Seong, and B.S. Choi, *Nucleic Acids Res.*, **27**, 1392 (1999).
4. R. B. Wickner, *Microbiol. Rev.*, **60**, 250 (1996).
5. T. Fujimura, R. Esteban, L. M. Esteban, and R. B. Wickner, *Cell*, **62**, 819 (1990).
6. Y. Shen and J. A. Bruenn, *Virology*, **195**, 481 (1993).
7. W. Yao, K. Adelman, and J. A. Bruenn, *J. Virol.*, **71**, 2157 (1997).
8. J. F. Milligan, D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck, *Nucleic Acids Res.*, **15**, 8783 (1987).
9. J. R. Wyatt, M. Chastain, and J. D. Puglisi, *Biotechniques*, **11**, 764 (1991).
10. C. W. J. Smith, *RNA:Protein Interactions*, Oxford University Press, Oxford, 1998.
11. P. J. Hore, *J. Magn. Reson.*, **55**, 283 (1983).
12. N. Müller, R. R. Ernst, and K. Wüthrich, *J. Am. Chem. Soc.*, **108**, 6482 (1986).
13. M. Rance, *J. Magn. Reson.*, **74**, 557 (1987).
14. G. Varani, and I. Tinoco, Jr., *J. Am. Chem. Soc.*, **113**, 9349 (1991).
15. V. Sklenar, H. Miyashiro, G. Zon, H. T. Miles, and A. Bax, *FEBS Lett.*, **208**, 94 (1986).
16. E. P. Nikonowicz and A. Pardi, *J. Mol. Biol.*, **232**, 1141 (1993).