FEBS Letters 585 (2011) 1049–1053

journal homepage: www.FEBSLetters.org

The solution structure of coronaviral stem-loop 2 (SL2) reveals a canonical CUYG tetraloop fold

Chul Won Lee [a], Lichun Li [b,1], David P. Giedroc [a,*]

[a] Department of Chemistry, Indiana University, Bloomington, IN 47405, USA
[b] Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA

ARTICLE INFO

Article history:
Received 18 November 2010
Revised 8 February 2011
Accepted 1 March 2011
Available online 4 March 2011

Edited by Christian Griesinger

Keywords:
Nuclear magnetic resonance
Solution structure
RNA tetraloop
SARS
Coronavirus
Mouse hepatitis virus

ABSTRACT

The transcription and replication of the severe acute respiratory syndrome (SARS) coronavirus (SARS-CoV) is regulated by specific viral genome sequences within 5′- and 3′-untranslated regions (5′-UTR and 3′-UTR). Here we report the solution structure of 5′-UTR derived stem-loop 2 (SL2) of SARS-CoV determined by NMR spectroscopy. The highly conserved pentaloop of SL2 is stacked on 5-bp stem and adopts a canonical CUYG tetraloop fold with the 3′ nucleotide (U51) flipped out of the stack. The significance of this structure in the context of a previous mutagenesis analysis of SL2 function in replication of the related group 2 coronavirus, mouse hepatitis virus, is discussed.

1. Introduction

Severe acute respiratory syndrome (SARS) is a disease caused by the SARS-associated coronavirus (SARS-CoV) comprised of a single-stranded, positive-sense RNA genome of ≈30 kb in length. For all CoVs the 5′ two-thirds of the genome encode non-structural proteins involved in proteolytic processing of the gene1 polyprotein, virus genome replication and subgenomic RNA (sgRNA) synthesis, and the 3′ one-third of the genome encodes structural and accessory proteins (Fig. 1A).

Coronaviruses express seven to nine sgRNAs during replication, each containing a common 5′ leader sequence and 3′-untranslated region (UTR) that harbor important structural elements involved in replication and/or translation [1–5]. Although the mechanism of CoV transcription and replication remains poorly understood, discontinuous transcription during minus-strand synthesis is the currently accepted model. A nested set of subgenome-sized co-terminal negative-sense RNAs are transcribed from positive-sense genomic RNA by the viral transcriptase/replicase complex (TRC), which then serve as templates for subgenomic mRNA (sg mRNA) synthesis. The 3′- end of the ≈70-nt leader within the 5′-UTR contains a short (6- to 8-nucleotides) sequence, the transcriptional regulatory sequence (TRS-L), which also is present in the genome just 5′ to each structural gene (TRS-B) [6]. Molecular genetic studies are consistent with a leader-body joining model which the complement to TRS-B on newly synthesized minus strands base-pairs with TRS-L to regulate the synthesis of sgRNAs by template switching [7–10].

Secondary structural models predict that the 5′ region of the 5′-UTR folds into three major stem-loops, SL1, SL2, and SL4b [11,12]. SL3, which harbors the TRS-L (5′-CUCAAAC) is only predicted to be stable at 37 °C for OC43 and SARS-CoV [11] (Fig. 1B). Mutations in the helical stem of SL1 or the loop of SL2 have pronounced effects on mouse hepatitis virus (MHV) replication, largely manifest as a defect in sgRNA transcription [5,11,13]. Although the sequences and predicted secondary structures of MHV and SARS-CoV 5′ UTRs are significantly different, the SARS-CoV SL1, SL2, and SL4 can functionally replace their MHV counterparts in the MHV genome and produce viable chimeric viruses [14].

Excepting the TRS, SL2 is the most highly conserved sequences in the 5′-UTRs of CoVs [11] and is characterized pentaloop (C47-U48-U49-G50-U51 in SARS-CoV) stacked on a 5-bp stem (Fig. 1C and D), with some CoV sequences containing an additional U 3′ to U51 [11]. Here we report the structure of SL2 of SARS-CoV determined by NMR spectroscopy. SL2 adopts a tetraloop fold stacked on a helical stem. Tetraloops have been grouped by their

* Corresponding author. Address: Department of Chemistry, Indiana University, 212 S. Hawthorne Drive, Bloomington, IN 47405, USA. Fax: +1 812 856 5710.
E-mail address: giedroc@indiana.edu (D.P. Giedroc).
[1] Present address: Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, IN 47405, USA.

0014-5793/$36.00
doi:10.1016/j.febslet.2011.03.002

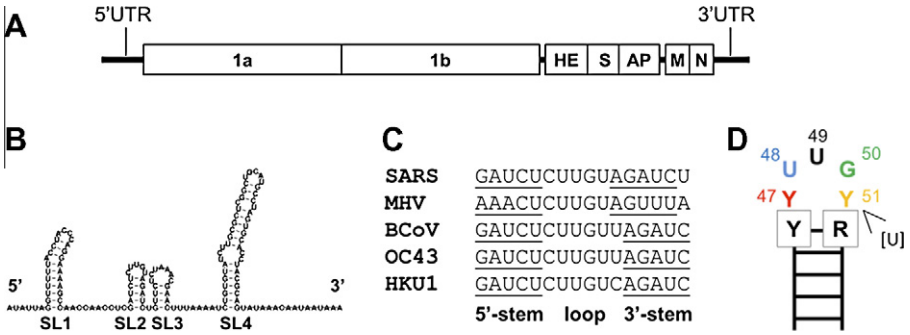

Fig. 1. Coronavirus genome and SL2 secondary structure. (A) ORFs of coronavirus. Gene 1 (1a and 1b), hemagglutinin-esterase (HE), spike glycoprotein (S), accessory proteins (AP), membrane protein (M), nucleocapsid (N), and 5′- and 3′-UTR. (B) Predicted secondary structure of 5′-UTR of SARS-CoV. Stem-loops (SL1, SL2, SL3, and SL4) are indicated. (C) Primary sequences of the SL2 loop region from five coronaviruses. Stem regions are underlined. (D) Schematic representation of the sequence conservation of SL2 of all coronaviruses.

sequence and conserved structures into five types: (i) GNRA, (ii) UNCG, (iii) ANYA, (iv) (U/A)GNN, (v) CUYG. Recently, they have been further subclassified according to specific deviations from the standard tetraloop motif, *e.g.,* a 3–2 switch, deletion, insertion, and strand clips [15]. SL2 adopts the CUYG-like, insertion-type tetraloop structure which features a C47–G50 Watson–Crick (WC) base pair with the conserved 3′ nucleotide, U51 flipped out of the stack.

2. Materials and methods

2.1. Sample preparation

Unlabeled and ^{13}C, ^{15}N-[U]-labeled wild-type (WT) RNA were prepared as described previously [5]. For NMR, SL2 was dissolved in 10 mM potassium phosphate, pH 6.0 in 10% D_2O/90% H_2O or 100% D_2O. All RNAs were monomeric under these conditions verified non-denaturing polyacrylamide gel electrophoresis.

2.2. NMR spectroscopy

NMR experiments were acquired on a Varian Inova 500 or 600 MHz spectrometer at 283 and 298 K [5]. NMR data were processes and analyzed with NMRPipe [16], Sparky [17] and NMRView [18]. Several mixing times (τ_m = 60, 250, and 280 ms) in 2D-NOESY experiments were tested to confirm the absence of significant spin diffusion. A 2D ^1H-NOESY spectrum (τ_m = 200 ms) in 10% D_2O/90% H_2O was acquired to obtain imino proton resonance assignments, while 2D ^1H-NOESY (τ_m = 250 ms) and 2D ^1H-TOCSY experiments in D_2O were performed to obtain non-exchangeable proton resonance assignments and NOE restraints using standard methodologies [19].

2.3. Structure calculation and analysis

NOE peak assignment and initial NOE constraints were obtained with CYANA [20] and CANDID [21]. All NOE constraints were manually confirmed during the CYANA calculations. Hydrogen bonding constraints were introduced for all base pairs and artificial torsion angle restraints derived from the high-resolution crystal structures of A-form double-helical RNA were used to impose better convergence of the ensemble [22].

The initial 100 structures were calculated by a simulated annealing protocol with Xplor-NIH [23] and were further refined using a conformational database potential [24] and planarity restraints for the helical stem region. Iterative refinement and editing of the distance restraints based on the NOESY spectra to remove incorrect and ambiguous assignments reduced the number of restraints. Force constants were 0.2–30 kcal mol^{-1} Å^{-2} for NOE restraints and 10–100 kcal mol^{-1} rad^{-2} for dihedral angle restraints in the refinement calculations. The final 27 structures with the lowest energy were chosen for analysis using the programs Xplor-NIH and 3DNA [25] and are deposited in the PDB (accession code 2L6I). NOEs in the loop region (U46–A52) of the SL2 RNA were confirmed by back-calculation of the NOE intensity using Xplor-NIH (see Table S1 and Fig. S1). Chemical shifts of the SL2 RNA are deposited in the BMRB (accession code 17309). Figures were prepared using the program PyMOL [26].

3. Results and discussion

3.1. Solution structure of SL2

Coronavirus SL2 used in this study is SARS-CoV SL2 containing a conserved 5′-CUUGU pentaloop, which differs from the MHV SL2 only in the identity of two of the five bp in the stem (Fig. 1C). The SL2 construct used for NMR contains a non-native 3′ A to stabilize the base of the stem. In the initial CYANA-derived structures, C47 was found to stack on U46 with G50 stacked on the A52 (Fig. 2A) and U51 flipped out from the stem (see also [5]). G50 adopted a high-*anti* glycosidic bond angle. These structural characteristics are found in the CUGG tetraloop structure containing a base pair between C_i (C47) and G_{i+3} (G50) [22], a finding also consistent with the recovery of second-site C47A-G50U MHV viruses from G50U MHV stocks after multiple passages [5]. We therefore added hydrogen bonding constraints between C47 and G50 in the final refinement step, although the imino proton associated with this base pair could not be detected experimentally.

The NMR structure of SL2 is fully consistent with our previous studies of SL2 [5]. The bundle of structures is well converged with 0.47 Å RMSD for all heavy atoms (Table 1). The stem adopts an A-form helix containing five WC base pairs with the 3′ terminal nucleotide A57 disordered (Fig. 2B). The pentaloop is quite well defined and stabilized by base pairing and intra- and inter-nucleotide interactions (Fig. 2C). U48 base lies in the minor groove of the stem, with the orientation of this base not fully converged (Fig. 2B) but likely stabilized by hydrophobic contacts between H5 and H6 edge of the U48 base and the sugar ring of C47 (Fig. 3A). U49 stacks on C47 in the C47-G50 base pair and thus caps the helical stem and the O2 of U49 and H42 proton of C47 are in close proximity (Fig. 3A). U48, U49 and U51 each adopt a C2′-*endo* ribose conformation in the SL2 structure, consistent with the strong H1′–H2′ cross peaks in an ^1H–^1H TOCSY spectrum which reports on large ^3J(H1′,H2′) vicinal coupling (Table S2 and Fig. S2). In contrast, C47 and G50 adopt at least some C3′-*endo* ribose pucker consistent with their weaker H1′–H2′ cross peaks (Fig. S2) as might

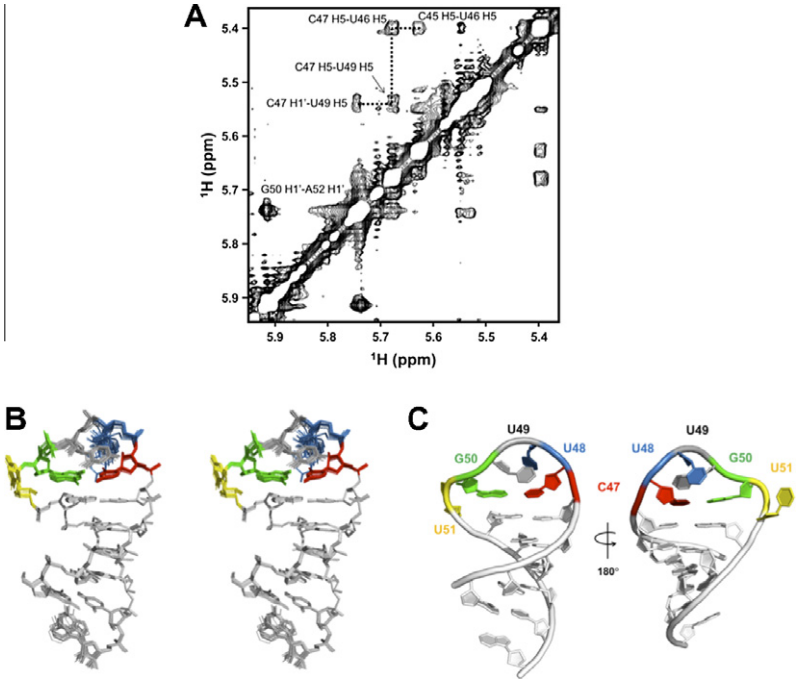

Fig. 2. NOEs and solution structure of SL2. (A) Key NOEs in loop region observed in a ^{1}H–^{1}H NOESY spectrum acquired in D_2O that establish interresidue interactions (C45-U46-C47-U49, G50-A52). (B) Stereo pair of the lowest energy 27 structures superposed on heavy atoms. The loop residues are colored in red (C47), blue (U48), grey (U49), green (G50), and yellow (U51). (C) Ribbon representation of the SL2 structure using in the same coloring scheme as in (B). The left and right models are rotated 180° relative to one another about a vertical axis. Residues in the pentaloop are labeled according the SARS-CoV/MHV nucleotide sequences.

Table 1
NMR restraints and structural statistics.

NMR constraints	
Total NOE distance restraints	213
Intra-residue (*i*, *i*)	104
Sequential (*i*, *i* + 1)	90
Medium-range ($2 \leqslant \lvert i-j \rvert \leqslant 4$)	12
Long-range ($\lvert i-j \rvert \geqslant 5$)	7
Hydrogen bonds	42
Total dihedral angle restraints	95
Structural statistics (*27 structures*)	
Violations	
Number of distance restraint >0.3 Å	0
Number of dihedral angle restraint >5°	0
Rms deviation from experiments	
Distance (Å)	0.052 ± 0.001
Dihedral angle (°)	0.100 ± 0.080
Rms deviation from idealized geometry	
Bonds (Å)	0.0047 ± 0.0001
Angles (°)	0.9860 ± 0.0210
Impropers (°)	0.5840 ± 0.0119
Average pairwise RMS deviations (Å)	
Backbone heavy atoms	0.33 ± 0.13
All heavy atoms	0.47 ± 0.18

be anticipated on the basis of C47–G50 base pair. The pentaloop is clearly more dynamic than the helical stem region, but this was not systematically investigated further. U51 is flipped out of the stack between G50 and A52. There are no inter-residue interactions involving U51, thus revealing that U51 is solvent exposed and likely mobile in solution; this is consistent with the sharp linewidths of the H5 and H6 protons [5].

3.2. SL2 adopts a CUYG-like tetraloop structure

The consensus pentaloop sequence of CoV SL2 is 5′ yYUUGY(U)$_n$r, (n = 0 or 1) [5] (Fig. 1D) and is therefore consistent with either a U-turn-like structure containing a UNR triloop stacked on the stem as in the VS ribozyme (Fig. 3B) or a 5′-gCUYGc tetraloop, the prototype member of a more diverse CNGG(N)$_n$ family of tetraloops (Fig. 3C). The structure of CoV SL2 reveals that the loop structure of SL2 adopts a CNGG(N)$_n$ tetraloop topology [5]. Fig. 3A and C show the structures of the loop of SL2 and Smaug recognition element (SRE), respectively, the latter of which is a member of CNGG(N)$_n$ tetraloop family. Both pentaloops stack on the stem closing U-A base pair. The first and fourth residues (C47 and G50 in SL2 and C10 and G13 in SRE) in the loop form a base pair in which the fourth residue adopts a high-*anti* χ angle (G50 in SL2 = ∼−80° and G13 in SRE = ∼−60°) [22]. The second base (U48 in SL2 and U11 in SRE) lies in the minor groove and is stabilized by hydrophobic interactions. The third residue (U49 in SL2 and G12 in SRE) stacks on the loop base pairing interaction between the first and fourth residues on the opposite of the molecule. The fifth residue (U51 in SL2 and C14 in SRE) is flipped out from the stack. One difference between these two structures is the identity of the third loop residue in SL2 vs. SRE. The identity of this nucleotide is functionally unimportant in MHV since all U49 substitution mutants of SL2 are viable [5], a finding compatible with the structure.

3.3. Structure-function correlations

We previously reported that the MHV SL2 loop is rather highly functionally tolerant of base substitutions [5]. In fact, when a more stable SARS-CoV SL2 stem sequence replaces the native MHV SL2 stem containing multiple A-U base pairs at the base of the stem, both originally characterized lethal U48C and G50C mutations in an all-MHV context were found to be viable [5]. We therefore previously suggested that SL2 plays generic structural role in stabilizing a higher-order structure within the 5′-UTR or a 5′-UTR–3′-UTR complex that is important specifically for sgRNA synthesis. Structural and functional data suggest that the identity

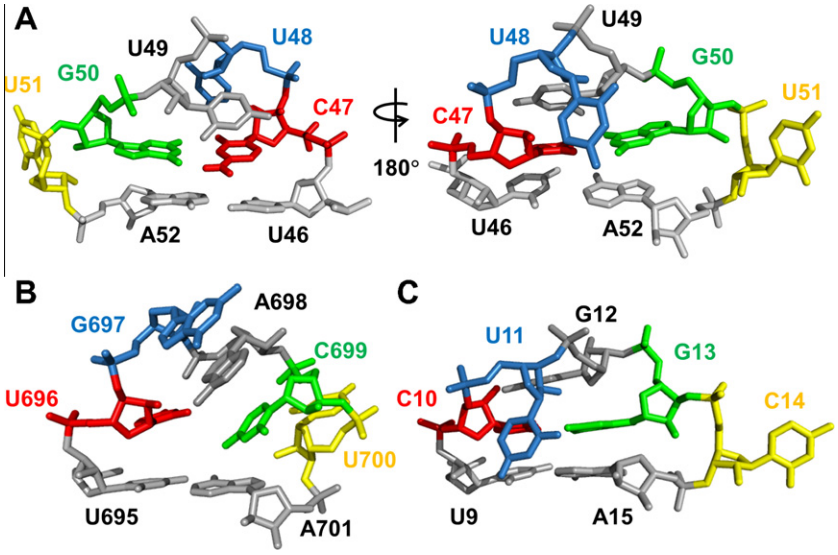

Fig. 3. Comparison of pentaloop structures of SARS-CoV SL2 (A) (this work; PDB ID, 2L6I), VS ribozyme stem-loop (B) (PDB ID, 1TBK) [31], and Smaug recognition element stem-loop (C) (PDB ID, 2ES5) [22]. In panel (A), the left and right models use the same coloring scheme as Fig. 2B and are rotated 180° relative to one another about a vertical axis. Panels (B) and (C) use a coloring code that is analogous to that shown for SL2 in panel (A) to facilitate comparison of the pentaloop structures.

of U51 is unimportant but may facilitate the folding of SL2 rather than specifically mediating a long-range RNA–RNA or RNA–protein interaction required for replication [5]. Interestingly, in all recovered ΔU51 MHV viruses, U51 was added back in; furthermore, extrahelical U51-like residues are often conserved in stable tetraloops, including the 5′-CNGG and 5′-YNMG-like tetraloop structures. These findings suggest that U51 plays a critical role in stabilizing the loop structure required for virus viability.

A base pair between C47 and G50 in SL2 is consistent with the fact that all G50 substitution mutants were found to be lethal in MHV; in contrast, corresponding C47 substitutions appeared to have comparatively little negative impact on virus titer [5]. Fig. 4 shows predicted secondary structures and free energy differences (ΔG) between selected SL2 C47 and U51 mutants relative to wild-type SL2 [27]. As can be seen, all C47 mutations potentially extend the helical stem by forming an additional base pair with U51, creating a capping YYR triloop which can be stabilizing [28]. The C47U mutant may incorporate a canonical U47–G50 Wobble pair with a wild-type-like tetraloop fold or a non-canonical U–U base pair (U47–U51) closing a YYR triloop as found in 16S rRNA [29,30]. In addition, a U51G mutation is predicted to even more stabilizing (Fig. 4). Taken together, these predictions partially explain why C47 mutations in an all MHV context did not strongly

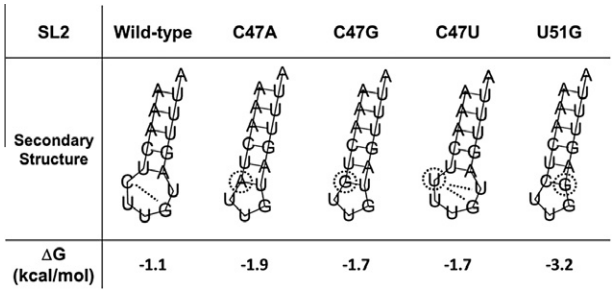

Fig. 4. Prediction of secondary structures and folding free energies (ΔG) of MHV SL2 substitution mutants relative to wild-type SL2. Base pairs in the loop are indicated with a *dotted* line. Mutated residues are indicated with a *dotted* circle. The free energies shown (37 °C) were calculated using Vienna RNAfold and a constant free energy increment for the 5′-UUG triloop in each mutant RNA. The stabilities of wild-type SL2 have not yet been experimentally determined.

negatively impact virus viability, but were absolutely co-dependent on the presence of U51 in the loop. On the other hand, our SL2 structure provides no clear structural rationale as to why U48C and U48A mutants were lethal in MHV [5]; one strong possibility is that these mutations induce misfolding in the 5′ leader region, facilitated by the weaker SL2 helical stem in MHV relative to SARS-CoV (Fig. 1C–D). Additional structural studies of the entire CoV 5′-UTR will be required to substantiate this proposal.

Acknowledgments

This work was supported by Grant AI067416 from the National Institutes of Health. We thank Dr. Xiangming Kong, Texas A&M University, for help in acquiring some of the NMR data.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2011.03.002.

References

[1] Raman, S., Bouma, P., Williams, G.D. and Brian, D.A. (2003) Stem-loop III in the 5′ untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J. Virol. 77, 6720–6730.
[2] Raman, S. and Brian, D.A. (2005) Stem-loop IV in the 5′ untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J. Virol. 79, 12434–12446.
[3] Goebel, S.J., Miller, T.B., Bennett, C.J., Bernard, K.A. and Masters, P.S. (2007) A hypervariable region within the 3′ cis-acting element of the murine coronavirus genome is nonessential for RNA synthesis but affects pathogenesis. J. Virol. 81, 1274–1287.
[4] Zust, R., Miller, T.B., Goebel, S.J., Thiel, V. and Masters, P.S. (2008) Genetic interactions between an essential 3′ cis-acting RNA pseudoknot, replicase gene products, and the extreme 3′ end of the mouse coronavirus genome. J. Virol. 82, 1214–1228.
[5] Liu, P., Li, L., Keane, S.C., Yang, D., Leibowitz, J.L. and Giedroc, D.P. (2009) Mouse hepatitis virus stem-loop 2 adopts a uYNMG(U)a-like tetraloop structure that is highly functionally tolerant of base substitutions. J. Virol. 83, 12084–12093.
[6] Budzilowicz, C.J., Wilczynski, S.P. and Weiss, S.R. (1985) Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3′ end of the viral mRNA leader sequence. J. Virol. 53, 834–840.
[7] van Marle, G., Dobbe, J.C., Gultyaev, A.P., Luytjes, W., Spaan, W.J. and Snijder, E.J. (1999) Arterivirus discontinuous mRNA transcription is guided by base

pairing between sense and antisense transcription-regulating sequences. Proc. Natl. Acad. Sci. USA 96, 12056–12061.
[8] Pasternak, A.O., van den Born, E., Spaan, W.J. and Snijder, E.J. (2003) The stability of the duplex between sense and antisense transcription-regulating sequences is a crucial factor in arterivirus subgenomic mRNA synthesis. J. Virol. 77, 1175–1183.
[9] Zuniga, S., Sola, I., Alonso, S. and Enjuanes, L. (2004) Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J. Virol. 78, 980–994.
[10] Yount, B., Roberts, R.S., Lindesmith, L. and Baric, R.S. (2006) Rewiring the severe acute respiratory syndrome coronavirus (SARS-CoV) transcription circuit: engineering a recombination-resistant genome. Proc. Natl. Acad. Sci. USA 103, 12546–12551.
[11] Liu, P., Li, L., Millership, J.J., Kang, H., Leibowitz, J.L. and Giedroc, D.P. (2007) A U-turn motif-containing stem-loop in the coronavirus 5′ untranslated region plays a functional role in replication. RNA 13, 763–780.
[12] Chen, S.C. and Olsthoorn, R.C. (2010) Group-specific structural features of the 5′-proximal sequences of coronavirus genomic RNAs. Virology 401, 29–41.
[13] Li, L., Kang, H., Liu, P., Makkinje, N., Williamson, S.T., Leibowitz, J.L. and Giedroc, D.P. (2008) Structural lability in stem-loop 1 drives a 5′ UTR-3′ UTR interaction in coronavirus replication. J. Mol. Biol. 377, 790–803.
[14] Kang, H., Feng, M., Schroeder, M.E., Giedroc, D.P. and Leibowitz, J.L. (2006) Putative cis-acting stem-loops in the 5′ untranslated region of the severe acute respiratory syndrome coronavirus can substitute for their mouse hepatitis virus counterparts. J. Virol. 80, 10600–10614.
[15] Hsiao, C., Mohan, S., Hershkovitz, E., Tannenbaum, A. and Williams, L.D. (2006) Single nucleotide RNA choreography. Nucleic Acids Res. 34, 1481–1491.
[16] Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293.
[17] Goddard, T.D., Kneller, D.G. SPARKY 3, University of California, San Francisco.
[18] Johnson, B.A. and Blevins, R.A. (1994) NMR view: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR 5, 603–614.
[19] Furtig, B., Richter, C., Wohnert, J. and Schwalbe, H. (2003) NMR spectroscopy of RNA. Chembiochem 4, 936–962.
[20] Guntert, P. (2004) Automated NMR structure calculation with CYANA. Methods Mol. Biol. 278, 353–378.
[21] Herrmann, T., Guntert, P. and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J. Mol. Biol. 319, 209–227.
[22] Oberstrass, F.C., Lee, A., Stefl, R., Janis, M., Chanfreau, G. and Allain, F.H. (2006) Shape-specific recognition in the structure of the Vts1p SAM domain with RNA. Nat. Struct. Mol. Biol. 13, 160–167.
[23] Schwieters, C.D., Kuszewski, J.J., Tjandra, N. and Clore, G.M. (2003) The Xplor-NIH NMR molecular structure determination package. J. Magn. Reson. 160, 65–73.
[24] Clore, G.M. and Kuszewski, J. (2003) Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. J. Am. Chem. Soc. 125, 1518–1525.
[25] Lu, X.J. and Olson, W.K. (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. Nat. Protoc. 3, 1213–1227.
[26] The PyMOL Molecular Graphics System, Version1.2r3pre, Schrödinger, LLC.
[27] Hofacker, I.L. (2003) Vienna RNA secondary structure server. Nucleic Acids Res. 31, 3429–3431.
[28] Thulasi, P., Pandya, L.K. and Znosko, B.M. (2010) Thermodynamic characterization of RNA triloops. Biochemistry 49, 9058–9062.
[29] Wimberly, B.T., Brodersen, D.E., Clemons Jr., W.M., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T. and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. Nature 407, 327–339.
[30] Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T. and Ramakrishnan, V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. Nature 407, 340–348.
[31] Campbell, D.O. and Legault, P. (2005) Nuclear magnetic resonance structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping. Biochemistry 44, 4157–4170.