Ion-RNA Interactions in the RNA Pseudoknot of a Ribosomal Frameshifting Site: Molecular Modeling Studies

Shu-Yun Le, Jih-H. Chen†,
N. Pattabiraman†, and Jacob V. Maizel, Jr.

Laboratory of Experimental
and Computational Biology
DBS, FCRDC, National Cancer Institute
NIH, Frederick, MD 21702
Tel: 301-846-5576; Fax: 301-846-5598
e-mail: shuyun@orleans.ncifcrf.gov

† Frederick Biomedical Supercomputing Center

SAIC-NCI/FCRDC, P.O. Box B, Frederick, MD 21702

Abstract

The three-dimensional (3-D) structure of a RNA pseudoknot that causes the efficient ribosomal frameshifting in the *gag–pro* region of mouse mammary tumor virus (MMTV) has been determined recently by nuclear magnetic resonance (NMR) studies (9). But since the structure refinement in the studies did not use metal ions and waters, it is not clear how metal ions participate in the stabilization of the pseudoknot, and what kind of ion-RNA interactions dominate in the tertiary contacts for the RNA pseudoknotting. Based on the reported structure data of the pseudoknot VPK of MMTV (9), we gradually refined the structure by restrained molecular dynamics (MD) using NMR distance restraints (17). Restrained MD simulation of the RNA pseudoknot was performed with sodium ions and water molecules. Our results are in good agreement with known NMR data and delineate the importance of the metal ion coordination in the stability of the pseudoknot. In the non-coaxially stacking pseudoknot, stem 1 (S1), stem 2 (S2), and the intervening A14 involves unconventional stacking of base pairs coordinated by Na^+ and/or bridging water molecules. A6 and G7 of loop L1 make a perfect base stacking in the major groove and are further stabilized by coordinated Na^+ ions and water molecules. The first 4-nucleotide (nt) ACUC of loop L2 form a sharp turn and the following 4-nt AAAA cross the minor groove of S1 and are steadied by interactions with the nucleotides of S1, bridging water molecules and coordinated Na^+ ions. Our studies suggest that the metal ion plays a crucial role in the RNA pseudoknotting of VPK. In the stacking interior of S1 and S2, the Na^+ ion is positioned in the major groove and interacts directly with the carbonyl group O6 of G28 and carbonyl group O4 of U13 in the wobble base pair U13:G28. The ion-RNA interactions in MMTV VPK not only stabilize the RNA pseudoknot but also modify the electrostatic properties of the nucleotides at the critical parts of the pseudoknot VPK.

Introduction

Ribosomal frameshifting is an important regulating process of translational control in retrovirus replication. A growing number of data (1–4) indicate that an efficient –1 ribosomal frameshifting requires two cis-acting mRNA elements: a 'slippery sequence' at which the frameshift takes place, and a higher-order RNA structure just downstream of the frameshift site, usually in the form of a RNA pseudoknot that contains a quasi-continuous stem stacking. A recent report (5) demonstrates that the pseudoknot VPK (Fig. 1) promotes the efficient ribosomal frameshifting in the *gag–pro* overlapping region of mouse mammary tumor virus (MMTV) mRNA. The pseudoknot VPK is composed of 34 nucleotides (nt) that are folded into two stems, S1 of 5 base-pairs (bp) and S2 of 6 bp, and two loops, L1 with two nt (A6 and G7) and L2 (A20–A27) of 8 nt (5). The stems S1 and S2 are separated by an unpaired nt (A14) at the junction.

The structural feature observed in VPK is of general importance, since similar functional pseudoknots (6, 7) have also been determined in other retroviruses, such as feline immunodeficiency virus (FIV) and simian retrovirus-serotype 1 (SRV-1). *in vitro* –1 ribosomal frameshifting studies have demonstrated that these efficient pseudoknots contain an unpaired nt and share the same structural feature in the stem junction region with the pseudoknot VPK. Although these pseudoknots are different in their primary sequences and the sizes of stems and loops, both the pseudoknots of FIV and SRV-1 can promote the same or more efficient frameshifting at the MMTV *gag–pro* shift site as the MMTV pseudoknot itself (5). Replacing the MMTV pseudoknot (5) with a Tinoco's synthetic pseudoknot (8) of coaxial stacking, however, the frameshifting efficiency is reduced to a very low level (0.5%). Up to now, a few three-dimensional (3-D) structures of RNA pseudoknots have been reported (8-11, 33). At present, nuclear magnetic resonance (NMR) studies of the pseudoknot VPK indicate that the functional pseudoknot is bent by A14 and the nt of L1 and that two A-form RNA helical stems are not coaxially stacked (9). Taken together, these experimental results strongly suggest that a non-coaxially stacked junction of S1 and S2 and an unpaired nt in the RNA pseudoknot play a crucial role in the efficient frameshifting of MMTV *gag–pro* overlap. This may be a common structural determinant of RNA pseudoknots that can promote an efficient frameshifting in the ribosomal –1 frameshift on retroviral RNAs.

Results from numbers of laboratories have shown that tRNA, hammerhead ribozyme, group I introns and other large ribozymes require metal ions in the formation of their native tertiary structures (12-15). The coordination of metal ions by direct interactions to oxygens, nitrogens and hydroxyls of nt is a common property of ion-RNA interactions. Studies on the stability of RNA pseudoknots using both biochemical and biophysical methods have demonstrated that the stability of the pseudoknot (formed by two hairpins) relative to the two hairpins is dependent on the presence of Mg^{2+} (16). Also, high concentrations of Na^{+} stabilize the pseudoknot in relation to the two hairpin structures. At low Na^{+} concentrations, the equilibrium is shifted towards the hairpin structure (16). The active structure of RNA pseudoknots is thought to pack together through tertiary contacts. Thus, ion-RNA interactions should play essential roles in RNA pseudoknotting.

It is not clear how metal ions participate in the stabilization of the RNA pseudoknot and what kind of ion-RNA interactions dominate in the tertiary contacts for the RNA pseudoknotting. The determination of the 3-D structure of the pseudoknot VPK with metal ions and water molecules is highly desirable for our better understanding of the structural role of the RNA pseudoknot in the efficient –1 ribosomal frameshifting of MMTV. We report here a molecular model of the 3-D structure of the pseudoknot VPK of MMTV with sodium ions and water molecules. We investigated the structural conformation of the pseudoknot by restrained MD and simulated annealing using NMR derived distance restraints from Tinoco's laboratory (17). Four refined 3-D structural models of VPK are consistent with the NMR data. Their root mean square deviations of the backbone atoms plus N_9 of purines or N_1 of pyrimidines among four models are from 0.43 to 1.12 $\AA$. None of the derived structures has any individual distance constraint violations of greater than 0.1 $\AA$. In the refined VPK model, the Na^{+} ion places in the stacking interior of S1 and S2 and participates the electrostatic interactions with the carbonyl group O6 of G28 and carbonyl group O4 of U13 of the wobble base pair U13:G28. The ion-RNA interactions in MMTV VPK modify the electrostatic properties of the wobble base pair U13:G28 in the unconventional stacking of two stems S1 and S2.

Materials and Methods

The Tinoco model of the pseudoknot VPK (9) was used as the starting model in the simulation. The 163 distance constraints derived from the NMR data (17) were used in the study. These distance constriants did not include other 45 constriants of hydrogen bonds of 11 base pairs in VPK. The RNA 3-D structural modeling was completed on a Silicon Graphics workstation using MidasPlus of UCSF (18), a computer graphics system of molecular interactive display and simulation. All simulations and structural analyses were performed by using AMBER 4.0 (19) on the IRIS Indigo 2, the CONVEX 3220 and the CRAY Y-MP/8128 supercomputer systems. The force field parameters used in the simulation were taken from the standard force field (20) in the AMBER 4.0 package. A time step of 0.001 ps was used during all MD simulations.

The initial model was first minimized until the energy gradient of 0.03 $kcal/mol \times \AA$ by 200 steps of steepest descent minimization and 8800 steps of conjugate gradient minimization. In the calculation, we used a non-bonded cutoff of 30 $\AA$ and a distance dependent dielectric function with dielectric constant, $\varepsilon = 4.0$. The further simulation was performed in the simulation system that was composed of the 34-nt VPK neutralized by 33 sodium ions and solvated by 738 water molecules. The Na^+ ions were initially placed along the PO_2^- bisector and at a distance of 4.0 $\AA$ from phosphorus atoms. The 738 water molecules were filled a $64.6 \times 38.1 \times 36.7$ $\AA^3$ box by TIP4P water molecules, which had been collected from a Monte Carlo liquid (21). A non-bonded cutoff of 12 $\AA$ was used for the short range non-bonded interactions in the system. The restrained MD simulated annealing was performed using the SANDER program from the AMBER package (19) with a periodic boundary condition. The default parameters of the force constants for the NMR distance restraints were used throughout restrained MD simulations.

The system was minimized by moving these water molecules and Na^+ ions in order to relax possible steric clashes present at the interface of the RNA fragment, Na^+ ions, and water molecules. The simulation was followed by 60 ps MD with fixed RNA and counterion positions and mobile water molecules. During the simulation, the temperature was first heated up to 310 K from 1.0 K by a 0.1 ps^{-1} temperature coupling constant, and MD equilibration was then carried out at a constant temperature of 310 K. This was followed by an additional 50 ps restrained MD simulation with distance restraints from the NMR spectrum data of VPK, no other constraints were applied to the system. For all restrained MD simulations, unless otherwise specified, the SHAKE algorithm was used to constrain all bonds.

After total 110 ps MD simulation, the system was minimized by 100 steps of steepest descent minimization and 8900 steps of conjugate gradient minimization with NMR distance restraints. This refined structural model was termed the model M1. Similarly, three other models, denoted by M2-M4, were also computed. M2 differs from M1 by an additional 10 ps restrained MD simulation. That is, the MD simulation was continued by annealing the system temperature to 210 K before energy minimization. M3 differs from M2 by an additional 10 ps restrained MD simulated annealing from 210 to 110 K before energy minimization. In the additional simulated annealing, we employed a temperature coupling constant of 0.08 ps^{-1}. And M4 differs from M3 by 10 ps restrained MD simulated annealing from 110 to 1 K with a temperature coupling constant of 0.06 ps^{-1} before energy minimization.

Results and Discussion

Overall Tertiary Structure of VPK

The four 3-D structure models (M1-M4) of the pseudoknot VPK are in good agreement with the NMR data. The maximum distance geometry constraint deviation from the NMR data among the four models is less than 0.10 $\mathring{A}$. A view of a superposition of the four models of the pseudoknot VPK based on the backbone atoms (P, $O_{5'}$, $C_{5'}$, $C_{4'}$, $C_{3'}$, and $O_{3'}$) and N_9 in A and G or N_1 in C and U, is depicted in Figure 1. The calculated root mean square deviations between M4 and M1, between M4 and M2, and M4 and M3 are 1.123, 0.646, and 0.430 $\AA$, respectively. The energy gradients of the four converged models are less than 0.085 kcal/mol $\times\AA$. The structure M4 has the lowest total energy among the four models and is termed the VPK model (Figure 2). The total energies of the structures M1-M3 are 16%, 12%, and 5% higher than that of structure M4, respectively. The comparison (Table I) of hydrogen-bonded base pairs in the model M4 and Tinoco model indicates that the hydrogen bond lengths of base pairs in M4 are closer to those observed in the crystal structures of yeast phenylalanyl tRNA (23) and hammerhead RNA (24) than those observed in Tinoco model. In the views of VPK model shown in Fig. 2, the double helices S1 and S2 are A-form RNA conformation, and all nucleotides in S1 and S2 have a $C_{3'}-endo$ sugar pucker, except for U34 with a $C_{2'}-endo$ sugar pucker. Their sugar pucker conformations are consistent in other three models, M1–M3. In the VPK model, the backbone torsion angles around the P–$O_{5'}(\alpha)$, $O_{5'}-C_{5'}(\beta)$, $C_{5'}-C_{4'}(\gamma)$, $C_{4'}-C_{3'}(\delta)$, $C_{3'}-O_{3'}(\epsilon)$, and $O_{3'}-P(\zeta)$, as well as angle χ (around the glycosyl bond) are within the ranges allowed for tRNA crystal structures (22). All bases in the VPK model are in the *anti* orientation, and Watson-Crick base pairs are maintained in the stems S1 and S2.

S1 and S2 are bent so that the base-pair C5:G15 is distorted and the stem S1 does not coaxially stacked on the stem S2. The base of intervening A14 makes a sandwiched stacking to both G15 and base-pair U13:G28. In the junction region, the unpaired A14 makes a distinct hinge between stems S1 and S2. The loop L1 of A6 and G7 bridges the major groove of the stem S2. A6 and G7 make a perfect base stacking and their bases are towards outside of the major groove. Although no direct interactions between L1 and S2 are observed, the loop L1 is stabilized by the coordinations of three Na^+ ions, bridging water molecules, and nt in S1 and S2 including G4, C5, U8, G9, and G10 (Table II). Among them, Na_2^+ makes a hexahedral coordination by interacting to the phosphate oxygens of G4 and C5, N_7 atoms of A6 and G7, and two water molecules, where one water molecule W112 makes a hydrogen bond to phosphate oxygens of C5. The loop L2, from A20 to A27 (ACUCAAAA) connects G19 of the helix S1 and G28 of S2. All nt in L2 are stacked very well except A27 which is unstacked to its neighboring nt. The unstacked A27 in L2 has been observed in Tinoco model (9). Among these, the first 4-nt (ACUC) make a sharp turn following S1, where A20 and C21 continue base stacking to S1, and the sugar ring of U22 reverses direction. U22 has a $C_{1'} - exo$ sugar pucker that is close to $C_{2'} - endo$ sugar pucker. The other 4 nt AAAA in L2 cross the minor groove of S1 and the last two nt A26 and A27 are actively involved in the interactions with stem S1 by hydrogen bonds with C3, C18, G4 and G17. Consequently, S1, L1, S2 and L2 form a compact RNA tertiary structure.

Structural Features in the Junction Region of S1 and S2

The four nt C5, G15, U13 and G28 do form Watson-Crick and wobble base pairs in the junction of S1 and S2. The stem-stem junction is stabilized by base stacking among U13:G28, A14 and G15, Na^+ coordinations, bridging water molecules and involved hydrogen bonds (Fig. 3). These important, simultaneous coordinations include the hexahedral coordination of Na^+_{11} to O_6 of G28 and G29, O_4 of U13, phosphate oxygen of A6, as well as bridging water molecules W275 and W794 (Table III). Among them, Na^+_{11} is directly bound to the carbonyl groups of wobble base pair U13:G28 and G29 within the major groove of the pseudoknot VPK. Furthermore, the bridging W275 makes a hydrogen bond to O_4 of U13, and W794 hydrogen-bonds to both N_7 of G28 and N_1 of A14. Thus, the base pair interactions of U13:G28 are strengthened by the Na^+_{11} coordination and bridging water molecules. The bent base pair C5:G15 is stabilized by a water molecule (W503) that bridges among the functional groups O_6 of G15 and G4, and 4-NH_2 of C5. The non-coaxial stacking between S1 and S2 is further strengthened by the coordination of Na^+_{27} with the 2'-hydroxyl of C5 and oxygen atoms of G28 and G29 and two hydrogen bonds that are formed between 2'-hydroxyl of U13 and $O_{4'}$ of A14, and between 2'-hydroxyl of A14 and $O_{4'}$ of G15. Also, a number of hydrogen bonds that are coordinated by bridging water molecules are involved in the stabilization of the non-coaxial stacking. For instance, 2-NH_2 of G15 hydrogen-bonds to $O_{4'}$ of G28 by the bridging water molecule, W104.

In summary, the stability of the distinct conformation of the center core structure is strengthened by the coordinating network of Na^+ ions and bridging water molecules.

The sodium ion positioned in the RNA major groove takes part in chemistry of functional groups of these nucleotides in the stem junction domain.

Sharp Turn in the Loop L2

To cross the minor groove of S1 with 5 bp and maintain a close contact between L2 and S1, the first 4 nt, ACUC, form a sharp turn and other 4 nt, quartet A, cross the minor groove of 5 bp and make a connection to G28. The sharp turn (Figures 4) formed by ACUC is promoted by a number of hydrogen bonds occurred among these nt (Table IV). For instance, 4-NH_2 of C23 makes a hydrogen bond to O_4 of U22, and 6-NH_2 of A24 forms a three-center hydrogen bond to both $O_{5'}$ of A20 and $O_{3'}$ of C19. In addition to these hydrogen bonds, the coordinations promoted by Na^+_{17}, Na^+_{19} and Na^+_{24} also participate in the formation of the sharp turn.

Hydrogen bonds created by bridging water molecules are also observed in this domain (Fig. 4). The typical examples of these hydrogen bonds are the interactions among the N_7 and phosphate oxygen of A20, 4-NH_2 of C21, and the water molecule, W464; and among the phosphate oxygen of A20, 4-NH_2 of C23 and W494. The water molecule W494 is also involved in the coordination promoted by Na^+_{17}. It seems that these coordinating networks of Na^+ ions, and bridging water molecules can stabilize the specific base orientation and contribute to the stability of the 3-D structure of the sharp turn. As a result, A24 and U22 are distorted, their riboses are $C_{2'}-endo$ or close to $C_{2'}-endo$ sugar pucker.

In the connection to G28, A27 is unstacked to any other bases in the VPK structure. In order to interact properly with nt in S1, its ribose has a $C_{1'}-exo$ sugar pucker. The observed interactions include hydrogen bonds between N_3 of A27 and 2-NH_2 of G4, between N_7 of A27 and 2-NH_2 of G17, between 6-NH_2 of A27 and $O_{4'}$ of C18, and so on (Table IV). Also, A26 makes interactions with stem S1 by hydrogen-bonding to the 2'-hydroxyl of C3 and C18. Clearly, these hydrogen bonds stabilize the VPK structure. This kind of interactions between the loop L2 and stem S1 have also been observed in the NMR structure of TYMV pseudoknot (11).

Effects of Na^+ on the Stability of the VPK Structure

In our VPK model, most of the Na^+ ions participate mainly in the coordination of phosphate oxygens and water molecules. These coordinations are necessary for the thermodynamic stability of the overall VPK tertiary structure. Of the total 33 Na^+ ions, six (Na^+_2, Na^+_6, Na^+_{11}, Na^+_{17} Na^+_{19} and Na^+_{24}) are mainly involved in the direct electrostatic interaction with the heavy atoms of nt. Among them, three Na^+ coordinations promote the structural formation of the sharp turn in the loop L2 (Fig. 4), and two Na^+ coordinations strengthen the stability of the loop L1 conformation (Table II). Only one Na^+ ion of them, Na^+_{11}, is moved into the major grooves of RNA helices and stabilizes the non-coxial stacking among S1, A14 and S2 (Fig. 3a). If Na^+_{11} did not move into the RNA major groove, the base-pairing U13:G28 and the stem stacking in the junction region would likely be unstable. Consequently, our model indicates that the direct electrostatic interactions of sodium ions and oxygens, nitrogens and hydroxyls of nt play a crucial role in the pseudoknotting of VPK and in the formation of the specific structural determinant. Especially, the functional carbonyl groups of the wobble base pair U13:G28 provide major groove binding sites to the sodium ion in the stem junction domain of the RNA pseudoknot. These important ion-RNA electrostatic interactions provide a reasonable explanation why the stability of RNA pseudoknots is dependent on high Na^+ concentrations.

The 3-D conformation of RNA pseudoknots is highly dependent on the presence of metal ions, such as divalent ion Mg^{2+}, and monovalent ion Na^+ (16, 25). In general, divalent metal ions are more effective than monovalent ions in neutralizing the negative charges on the RNA backbone. Moreover, it is known that highly specific interactions between ribozymes and Mg^{2+} play an essential role in RNA catalysis (15). It is important for us to know these possible specific interactions of divalent ions in the VPK pseudoknotting. In our modeling studies, we did not include Mg^{2+} ions and other divalent ions in the simulation. It is our intention to study ion-RNA interactions between the VPK pseudoknot and various monovalent and multivalent ions. As the first step of our studies, we use Na^+ ions to refine VPK 3-D structure and to study ion-RNA interactions because the force field parameters of Na^+ are more solid than those of other metal ions in the AMBER 4.0 package.

As is known, a nt residue can span a distance of about 5.9 $\AA$ in a A-form RNA helical stem (22). In crossing the major groove of 6 bp of S2, two nt A6 and G7 make a 9.8 $\AA$ bridge between C5 and U8 so that the stem S2 is remarkably bent. A6 has a $C_{1'} - exo$ sugar pucker, close to a $C_{2'} - endo$ conformation, and G7 has a $O_{4'} - endo$ conformation that is a passage in which the $C_{3'} - endo$ ribose puckering transmits to the $C_{2'} - endo$ conformation. Our modeling studies also indicate that it is feasible for only a single nt A6 in L1 to cross the major groove with 7 bp stem (if G7 is base-paired with U34, data are not shown). In order to cross the minor groove of S1 with 5 bp and keep L2 and S1 as close as possible, the first 4 nt, ACUC make a sharp turn, and other 4 nt bridge the minor groove. This suggests that a sharp turn may be a common structural property for the loop crossing the helix groove in RNA pseudoknots if the loop has enough spannable distance, as suggested by Pleij et al. (26).

Over the last several years, restrained molecular dynamics in conjunction with NMR data has been successfully applied to the structure studies of various protein and nucleic acid molecules (27-32). Our initial model was built based on the reported RNA 3-D structure of VPK (9), some uncertainty related to the loop structures may also be introduced in the modeling, since there was insufficient constraints of the loop conformation from the NMR spectrum data. The restrained MD simulation of VPK with including Na^+ and water molecules improved the 3-D model of the pseudoknot VPK. The base stacked conformation of loops L1 and L2 in the proposed model M4 is quite different from that in the original model. In the model M4, the A6 and G7 of L1 form a perfect base stacking in bridging the major groove of the stem S2. However, the A6 and G7 are not base stacked in Tinoco model. Seven out of eight nt in L2 are stacked very well in the model M4, while most of nt in L2 are not stacked in Tinoco model. The improved base-stacking in the loops of L1 and L2 increases the thermodynamic stability of the pseudoknot VPK. In the model M4, the most interactions between L1 and S2 and between L2 and S1 are coordinated by Na^+ ions and bridging water molecules. Two nt of the loop L2, A26 and A27, participate in the direct interactions with stem S1. As a result, our models are more compact than the Tinoco model. The model M4 is about 28.41 % more compact than Tinoco model. But the shortest distance between two phosphorus atoms observed in M4 is larger than that observed in the Tinoco model. In the model M4, the shortest distance between two phosphorus atoms is 5.16 $\mathring{A}$, which is observed between C23 and A24 of the loop L2. For the Tinoco model, three distances between two phosphorus atoms are shorter than 5.0 $\mathring{A}$. Among them, the shortest distance, 4.14 $\mathring{A}$ is found between C21 and C23 of the loop L2.

Structural and functional studies on the MMTV pseudoknot and its mutants involved in ribosomal frameshifting (4, 5) have indicated that the stem junction domain provides a specific structural determinant for efficient frameshifting. Among them, the unpaired A14 and G28 of S2 are crucial in promoting an efficient frameshifting. For instance, the mutant pseudoknot NPK4, in which only the base pair U13:G28 of VPK was mutated to U13:A28, did not promote the efficient frameshiftting in MMTV. In our 3-D model, functional group O_6 of G28 binds directly to a Na^+ ion and its N_7 interacts with N_1 of A14 by a bridging water molecule. The N_7 and O_6 of G28 in VPK may provide the specific binding site for the two amino groups of arginine or other positive charged groups of amino acids. Our 3-D model of the pseudoknot VPK may provide a reasonable explanation for the biological function of the important wobble base pair U13:G28 based on the unique structural feature observed in the VPK. The distinct conformation of U13:G28 in the junction domain of VPK likely provide an ideal dock for the active domain of a functional RNA-binding protein.

Our results also indicate that the intervening nt A14 located between S1 and S2, can form various interactions to the stem S1 and S2 by bridging water molecules and Na^+ coordinations. These interactions and coordinations contribute the formation of the distinct conformation of the stem stacking. Moreover, its heavy atoms, $N_{1,}$, $6 - NH_2$, and N_7 are exposed and hydrogen-bonds to water molecules. It is known that $N_{1,}$, $6 - NH_2$, and N_7 of adenine are specific binding sites for amino and carbonyl groups of asparagine and glutamine. The specific interactions among A14 and two stems, S1 and S2, create a distinct structural conformation of the RNA pseudoknot that may facilitate recognition of a RNA-binding protein to the functional RNA element. These structural features provide useful information for further studies on the specific structural elements of the pseudoknot in the efficient frameshifting, and for the explication of fundamental roles of ion-RNA interactions in the formation and stability of the active 3-D structure of RNA pseudoknots.

ACKNOWLEDGEMENTS

We thank Drs. L.X. Shen, and I. Tinoco, Jr. for sharing their results; and Drs. Hugo M. Martinez and R. Andrew Byrd for their helpful discussion and comments on the manuscript. We acknowledge the Frederick Biomedical Supercomputing Center (FBSC) for allocation of computing time and staff support. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

1. I. Brierley, N.J. Rolley, A.J. Jenner and S.C. Inglis, J. Mol. Biol. **229**, 889–902 (1991)

2. N.M. Wills, R.F. Gesteland and J.F. Atkins, Proc. Natl. Acad. Sci. USA **88**, 6991–6995 (1991)

3. T.-H. Tzeng, C. Tu and J.A. Bruenn, J. Virol. **66**, 999–1006 (1992)

4. M. Chamorro, N. Parkin and H.E. Varmus, Proc. Natl. Acad. Sci. USA **89**, 713–717 (1992)

5. X. Chen, M. Chamorro, S.L. Lee, L.X. Shen, J.V. Hines, I.Jr. Tinoco and H.E. Varmus, EMBO J. **14**, 842–852 (1995)

6. S. Morikawa and D.H.L. Bishop, Virol. **186**, 389–397 (1992)

7. E. ten Dam, I. Brierley, S. Inglis and C. Pleij, Nucl. Acids Res. **22**, 2304–2310 (1994)

8. J.D. Puglisi, J.R. Wyatt and I.Jr. Tinoco, Nature **331**, 283–286 (1988)

9. L.X. Shen and I.Jr. Tinoco, J. Mol. Biol. **247**, 963–978 (1995)

10. H. Kang, J.V. Hines and I.Jr. Tinoco, J. Mol. Biol. **259,** 135–147 (1996)

11. M.H. Kolk, M. vander Graaf, S.S. Wijmenga, C.W.A. Pleij, H.A. Heus and C.W. Hilbers, Science **280,** 434–438 (1998)

12. P.E. Cole, S.K. Yang and D.M. Crothers, Biochemistry **11,** 4358–4368 (1972)

13. W.G. Scott, J.B. Murray, J.R.P. Arnold, B.L. Stoddard and A. Klug, Science **274,** 2065–2069 (1996)

14. A.M. Pyle, Science **261,** 709–714 (1993)

15. J.H. Cate, R.L. Hanna and J.A. Doudna, Nature Struc. Biol. **4,** 553–558 (1997)

16. J.R. Wyatt, J.D. Puglisi and I.Jr. Tinoco, J. Mol. Biol. **214**, 455–470 (1990)

17. L.X. Shen, Ph.D. Thesis, Dept. of Chemistry, University of California, Berkeley, California, USA (1994)

18. T.E. Ferrin, C.C. Huang, L.E. Jarvis and R. Langridge, J. Mol. Graphics **6**, 13-27

(1988); MidasPlus User's Manual, Computer Graphics Laboratory, School of Pharmacy, University of California, San Francisco (1992).

19. D.A. Pearlman, D.A. Case, J.C. Caldwell, G.L. Seibel, U.C. Singh, P. Weiner and P.A. Kollman, AMBER 4.0, University of California, San Francisco (1991)

20. S. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta and P. Weiner, J. Am. Chem. Soc. **106**, 765–784 (1984)

21. W. Jorgensen, J. Chandrasekhar, J. Madura, M. Impey and R. Klein, J. Chem. Phys. **79**, 926–935 (1983)

22. W. Saenger, Principles of Nucleic Acid Structure, Springer-Verlag, New York (1984)

23. E. Westhof and M. Sundaralingam, Biochemistry **25,** 4868–4878 (1986)

24. W.G. Scott, J.T. Finch and A. Klug, Cell **81,** 991–1002 (1995)

25. T.C. Gluick, N.M. Wills, R.F. Gesteland and D.E. Draper, Biochemistry **36,** 16173–16186 (1997)

26. C. Pleij, K. Rietveld and L. Bosch, Nucl. Acids Res. **13**, 1717–1731 (1985)

27. H.A. Heus, S.S. Wijmenga, H. Hoppe and C.W. Hibers, J. Mol. Biol. **271,** 147–158 (1997)

28. R.P. Ojha, M.M. Dhingra, M.H. Sarma, Y.P. Myer, R.F. Setlik, M. Shibata, A.L. Kazim, R.L. Ornstein, R. Rein, C.J. Turner and R.H. Sarma, J. Biomol. Struct. Dyn. **15,** 185–215 (1997)

29. R. Michalczyk and I.M. Russu, FEBS Letter **331,** 217–222 (1993)

30. M.A. Young, J. Srinivasan, I. Goljer, S. Kumar, D.L. Beveridge and P.H. Bolton, Methods Enzymol. **261,** 121–144 (1995)

31. M.R. Kiran and M. Bansal, J. Biomol. Struct. Dyn. **15,** 333–345 (1997)

32. P.S. Eis, J.A. Smith, J.M. Rydzewski, D.A. Case, D.L. Boger, and W.J. Chazin, J. Mol. Biol. **272,** 237–252 (1997)

33. Z. Du, D.P. Giedroc and D.W. Hoffman, Biochemistry **35,** 4187–4198 (1996)

Legends to Figures

Figure 1. A superposition of four models (M1–M4) of the 3-D structure of the RNA pseudoknot VPK, without displaying hydrogen atoms, water molecules and sodium ions. The 2-D model of VPK (5) is shown in the right. The stem S1 is composed of 5 base pairs of 1-GGCGC-5/15-GCGCC-19, and the stem S2 is composed of 6 base pairs of 8-UGGGCU-13/28-GGCCCA-33. Stems S1 and S2 are separated by an unpaired nt A14 at the junction. The loop L1 consists of nt A6 and G7, and the loop L2 consists of eight nt, 20-ACUCAAAA-27.

Figure 2. A ribbon representation of the proposed model (M4) of the pseudoknot VPK. Hydrogen atoms, water molecules and sodium ions are not shown in the plot. In the figure, stems S1 (cyanic) and S2 (magenta) are bent and twisted so that S1 and S2 are not coaxially stacked. The base pair C5:G15 is distorted, and the intervening A14 (yellow) between S1 and S2 makes a base stacking between G15 and U13:G28. A6 and G7 in the loop L1 (yellow) make a perfect base stacking in crossing the major groove of S2. The loop L2 (white) from nt A20 to A27 crosses the minor groove of S1. The PDB coordinate file of VPK (M4) is available from the first author by request.

Figure 3. Ball and stick model of the junction domain of S1 and S2 in VPK. The stem S1 is shown in cyanic and stem S2 is magenta. The loop L1 is green, L2 is white and A14 is orange color. Na^+ ions and water molecules are represented by red and white balls, respectively. The nt C5 is base paired with G15 and U13 is base paired with G28. (a) Coordinations among Na^+_{11}, A6, U13, G28 and G29, among Na^+_{27}, C5, G28, and G29, among Na^+_{28}, C5, A6 and G7, and among Na^+_2, C5, G4, A6, and G7. (b) Hydrogen bonds formed by bridging water molecules and nucleotides. These coordinations and hydrogen bonds are shown by yellow, broken lines. Only partial interactions are displayed in the figures (a) and (b), because of the limitation of the program MIDAS (16 distances between specified atoms can be displayed at most in the modeling session).

Figure 4. Ball and stick model of the sharp turn domain in L2 (white) and S1 (cyanic). The displayed structural domain includes two base pairs of S1 (G1:C19 and G2:C18) and five nt (20-ACUCA-24) of the loop L2, in which A20 connects to C19. Since the limitation of the program MIDAS, only partial coordinations promoted by sodium ions Na^+_{17}, Na^+_{19}, and Na^+_{24} and bridging water molecules are shown by yellow broken lines. Also, two hydrogen bonds between W494 and 4-NH_2 of C23, and between W494 and the phosphate oxygen of A20 are not shown in the figure.

Legends to Table

Table I. Comparisons of hydrogen-bonded base pairs of the model M4 and Tinoco model of the pseudoknot VPK. In the comparison, the mean of hydrogen bond lengths observed in the crystal structures of yeast phenylalanyl tRNA (23) and hammerhead RNA (24) are used for a useful yardstick.

Table II. Important coordinations of sodium ions and bridging waters (labeled by letter W) in the loop L1 and its neighborhood of the VPK (model M4).

Table III. Important coordinations of sodium ions and bridging waters in the junction region of the VPK (M4).

Table IV. Hydrogen bonds observed in the loop L2 and observed between L2 and the stem S1 of the VPK (M4).