Nucleic Acids Research Advance Access published December 28, 2012

Nucleic Acids Research, 2012, 1–13
doi:10.1093/nar/gks1302

Candidate RNA structures for domain 3 of the foot-and-mouth-disease virus internal ribosome entry site

Segun Jung and Tamar Schlick*

Department of Chemistry, Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA

Received August 31, 2012; Revised November 13, 2012; Accepted November 14, 2012

ABSTRACT

The foot-and-mouth-disease virus (FMDV) utilizes non-canonical translation initiation for viral protein synthesis, by forming a specific RNA structure called internal ribosome entry site (IRES). Domain 3 in FMDV IRES is phylogenetically conserved and highly structured; it contains four-way junctions where intramolecular RNA–RNA interactions serve as a scaffold for the RNA to fold for efficient IRES activity. Although the 3D structure of domain 3 is crucial to exploring and deciphering the initiation mechanism of translation, little is known. Here, we employ a combination of various modeling approaches to propose candidate tertiary structures for the apical region of domain 3, thought to be crucial for IRES function. We begin by modeling junction topology candidates and build atomic 3D models consistent with available experimental data. We then investigate each of the four candidate 3D structures by molecular dynamics simulations to determine the most energetically favorable configurations and to analyze specific tertiary interactions. Only one model emerges as viable containing not only the specific binding site for the GNRA tetraloop but also helical arrangements which enhance the stability of domain 3. These collective findings, together with available experimental data, suggest a plausible theoretical tertiary structure of the apical region in FMDV IRES domain 3.

INTRODUCTION

The foot-and-mouth-disease virus (FMDV) belonging to the picornavirus family is the contagious agent of foot-and-mouth-disease, a severe plague for animal farming. The viral replication of FMDV begins with a translation initiation by forming a specific RNA structure called internal ribosome entry site (IRES). FMDV IRES consists of ~450 nt and can fold in multiple stem–loops organized in five domains (Figure 1A). These domains can host binding proteins such as eukaryotic initiation factors (eIFs) and IRES transacting factors (ITAFs), which play crucial roles in IRES-directed translation (1–7). Among these, the third domain is the largest and contains structural elements critical for IRES activity (8).

Domain 3 consists of basal and apical regions (Figure 1A). The basal region consists of a long internal loop, and the apical region contains multiple four-way junctions. Recent biochemical data have suggested that it is the apical region that contributes significantly to the structural organization and stability of domain 3, as well as to the critical function of IRES activity (8,9, 10,11).

Specifically, the apical region of domain 3 includes two conserved motifs, GNRA and RAAA motif (8). The GNRA (N is any nucleotides; R is A or G) tetraloop motif is common in folded RNA (12); the loop–helix interactions combine base pairing and stacking to define a tertiary contact that stabilizes the global fold of an RNA molecule. In the IRES domain 3 (Figure 1), the GNRA motif is situated at the apex of a stem–loop motif (13,14). Biochemical studies demonstrate that this motif is critical for IRES function (8,15). RNA probing experimental data further show that the RAAA motif also contributes to enhance IRES activity via RNA–RNA long-range tertiary contacts, but only in the presence of GUAA tetraloop–receptor long-range interactions (9).

Deciphering the contribution of domain 3 to IRES-driven translation has been challenging. Based on the potential capacity for inter and intramolecular RNA–RNA interactions, it has been proposed that this domain stabilizes the entire IRES element (9,16). More recently, structural analysis based on SHAPE probing and microarray data confirmed domain 3's role in the organization of other domains (10,11).

The GNRA motif in helix H_5, along with its potential distal binding region in helix H_4 for intramolecular RNA–RNA interactions, is located in the four-way RNA

*To whom correspondence should be addressed. Tel: +1 212 998 3116; Fax: +1 212 995 4152; Email: schlick@nyu.edu

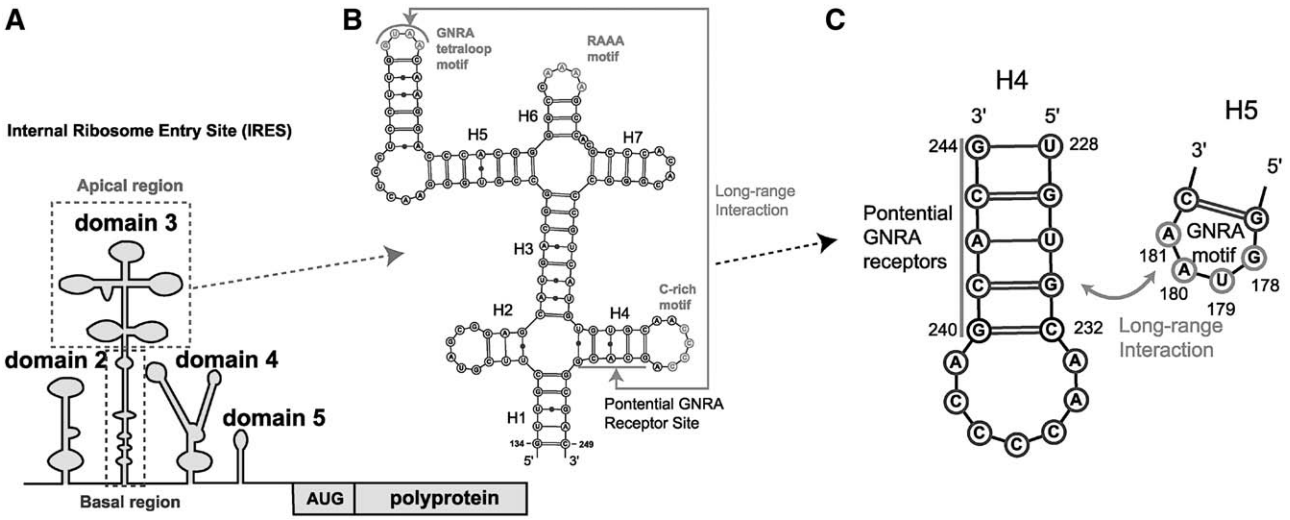

Figure 1. Global organization of FMDV IRES and a secondary structure of truncated domain 3 including conserved RNA motifs for RNA–RNA long-range interactions. (**A**) Schematic representation of the viral genome organization including four subdomains of FMDV IRES. (**B**) Secondary structure (deduced from RNA structure probing experiment) of the truncated FMDV IRES domain 3 which consists of a pair of four-way junctions and is a self-folding region containing conserved GNRA and RAAA motif at the apical region for RNA–RNA long-range interactions. (**C**) Potential long-range interactions between helices H_4 and H_5. $G_{240}CACG_{244}$ residue in helix H_4 is a potential receptor site of the $G_{178}UAA_{181}$ tetraloop.

junctions of FMDV IRES domain 3 (Figure 1). RNA junctions in general provide a hub for different double-stranded helical arms to come together (17). Thus, junctions occur in many RNAs, including, for example, the hepatitis C virus IRES for the translation initiation (18). Because the global conformation of a RNA is thought to be largely determined by topological constraints encoded at the secondary structure level (19), an understanding of the 3D structural aspects of RNA junctions in IRES's domain 3 is essential to decipher the mechanism of IRES-driven translation.

Among recurrent structural elements (or 'motifs') common to RNA junctions, coaxial stacking is prominent. Coaxial stacking between two continuous helices stabilized by base stacking in a shared single strand (16) is a major determinant of three or higher-order junctions of RNA. Recent studies of RNA junctions have identified structural patterns in coaxial stacking that define different RNA junction family types (20–22). These classifications also link the RNA junction family type to nucleotide length in a single strand; fewer nucleotides in the single strand between two helices increase the probability of forming coaxial stacking. Although other factors such as protein binding can alter these noted patterns of coaxial stacking arrangements, the above correlation holds in general particularly for self-folding RNA molecules, including transfer RNA (tRNA) (23), P4–P6 domain of the *Tetrahymena* group I ribozyme (24), hepatitis C virus IRES (18) and FMDV IRES domain 3 (8). The coaxial stacking motif often cooperates with other tertiary motifs including A-minor and loop–helix interaction to enhance the stability of RNAs.

Currently, computational programs cannot predict multiple RNA junction structures well, though there are many useful 3D prediction programs as recently surveyed (25,26). Very recently, our RNA junction structure prediction program Junction-Explorer, based on data mining and bioinformatics, was shown to predict the topology of individual RNA junction domains with ~70% or more prediction accuracy (27). Here, to construct plausible structures for the two consecutive four-way junctions in the apical region of IRES domain 3, we devise a divide-and-conquer approach that combines various effective computational techniques.

We began with the IRES secondary structure determined by RNA probing (13,28). Considering RNA–RNA long-range interactions involving the GNRA motif, we partitioned RNA into subsystems and then modeled each RNA junction topology on the basis of knowledge from four-way RNA junction classification coupled with Junction-Explorer. Further analysis produced four viable candidates for 3D models constructed using MC-Sym (29).

Subjecting these four candidate models to molecular dynamics (MD) simulations allowed identification of the most energetically favorable and stable conformational states in the presence of the GNRA tetraloop–receptor long-range interactions. The dynamics data also suggested specific tertiary interactions and helical rearrangements. Only one model emerged as viable, revealing not only the specific binding site for the GNRA tetraloop but also helical junction arrangements that enhance the stability of domain 3 further. We propose this structure, compatible with available experimental data, as a feasible tertiary structure for the apical region in FMDV IRES domain 3.

MATERIALS AND METHODS

RNA target structure

Domain 3 of FMDV IRES is a self-folding RNA that is 214 nt long. We consider the sequence of the FMDV C-S8

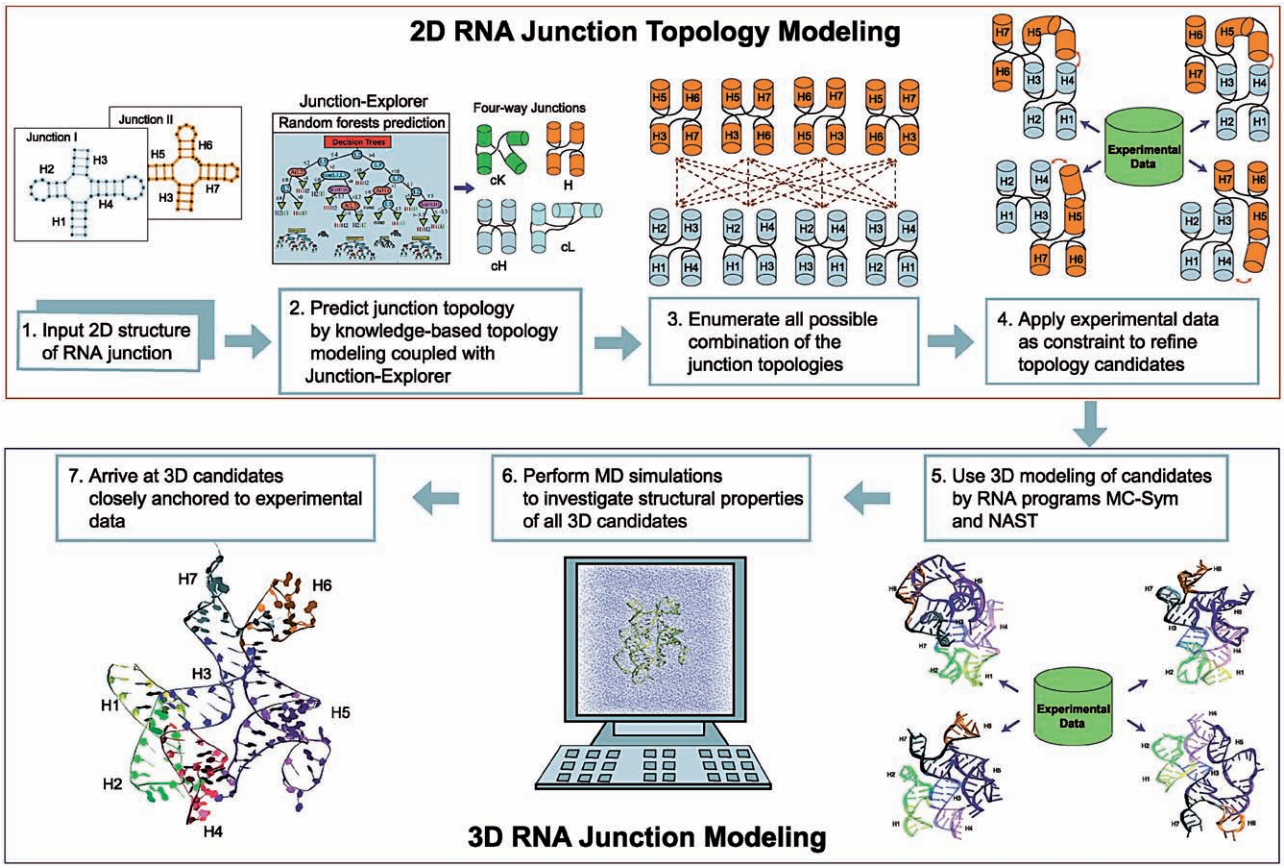

Figure 2. Computational procedure for modeling multiple four-way RNA junction structures. (1) Multiple four-way RNA junction structures are separated into individual RNA junction as input. (2) Each junction is analyzed for 2D helical arrangements based on coaxial stacking and junction family type in conjunction with Junction-Explorer. (3) These topologies are processed to enumerate all possible junction combinations. (4) Available experimental data are applied as constraints to refine the topology candidates. (5) 3D models based on the topology candidates are developed using computational programs. (6) MD simulations are performed to arrive at (7) potential 3D structure.

IRES and model the apical and basal region separately; the apical region contains two consecutive four-way junction structures, which consists of 116 nt (G_{134} to C_{249}) and the basal region is a long internal loop containing 98 nt (G_{86} to U_{133} and C_{249} to C_{299}).

RNA sequence alignments and conservation analysis

To assess the significance of the structural key elements involved in long-range RNA–RNA interactions in domain 3, we perform sequence alignments of many IRES sequences and analyze the four-way junctions, focusing on the sequence conservation of the GNRA loop and its binding receptors. Three hundred eighteen FMDV IRES sequences are collected from the GenBank database (30) using the standard Nucleotide Blast webserver with a query sequence of the FMDV C-S8 IRES. Incomplete and identical sequences we removed, and the remaining 318 sequences were aligned using the ClustalW program (31). See Supplementary Tables S2 and S3 for the list of FMDV IRES sequences and aligned sequences.

We use sequence logos to analyze patterns in aligned RNA sequences. The RNA sequence logos consist of stack of four letters (measured in 2 bits)—A, U, G and C—at each position in a sequence. While the overall height of the stack indicates a degree of sequence conservation, the height of each letter within the stack shows a relative frequency at each position. The logos are generated using the RNALogo webserver (32). See Supplementary Figure S1 for the RNA sequence logos of the four-way junctions.

Multiple RNA junction topology modeling for the apical region

To tackle multiple consecutive four-way RNA junctions, we use a divide-and-conquer approach by partitioning the large complex. Each four-way junction is analyzed with regards to the loop size of single strands between helices; this analysis is coupled to the Junction-Explorer program to help determine coaxial stacking patterns and helical arrangements. Junction-Explorer is based on the random forests data mining algorithm (33) and uses various geometric and energetic parameters as 'feature vectors' (which contain information on free energies, loop sizes between junctions and adenine content) for training. Using the predicted topology for each four-way junction, we search for all possible combinations of the multiple four-way junctions to produce combined structures. These potential topologies for the secondary structures

are then refined further by incorporating experimental data as constraints (Figure 2).

3D modeling of multiple RNA junction structures in the apical region

Using state-of-the-art 3D modeling programs, we build RNA 3D models of FMDV IRES domain 3 combined with experimental data. We primarily use MC-Sym, which utilizes a fragment-based library to obtain all possible structures of RNA junctions (29). To complement the modeling results, we explore conformational space using NAST, a knowledge-based coarse-grained simulation tool (34); these two programs have been shown to perform well in predicting native RNA structures (26). We hypothesize that fewer nucleotides between helices should naturally restrict the orientational flexibility at some degree yielding coaxially stacked helices. Thus, we first model Junctions I and II following the 5′- to 3′-direction without constraints for coaxial stacking arrangement. This yields thousands of structures for each junction. Because helical elements in RNA junctions tend to form coplanar arrangements (18), we ranked the predicted structures for coplanarity and collected the best 1000 structures. These junctions are then assembled by imposing a distance constraint for potential long-range interactions from experimental data (see 'Modeling atomic junction structures for the apical region' in 'Results' section for more details).

Using NAST, the Nucleic Acid Simulation Tool, we performed a coarse-grained MD simulation for 40 ns (10×10^6 time step) with one tertiary contact between A_{180} and C_{232}/G_{240}. For the 10 000 coarse-grained templates generated, we filter the templates by a potential energy with the cutoff energy of 1000 kJ.

Molecular dynamics simulations for the apical region

Each system was solvated with the explicit TIP3P water model in a water box of dimension 10 Å on each side. Simulations were performed using the Amber Parmbsc0 force field (35,36) with sodium ions to neutralize the system charge.

We minimize the system in two steps, first over the water and ion molecules holding domain 3 fixed and, second, with all constraints removed. The minimization was performed using the Powell conjugate gradient algorithm. The initial equilibration was achieved over 60 ps at constant temperature (300 K) and pressure (1 atm). Pressure was maintained at 1 atm using the Langevin piston method, with a piston period of 100 fs, damping constant of 50 fs and piston temperature of 300 K. Temperature coupling was enforced by velocity reassignment every 2 ps. Both minimization and equilibration are performed using the NAMD program (37).

For the production run, we simulated a conventional MD trajectory for 100 ns with the Parmbsc0 force field using the NAMD package. The system was simulated at constant temperature (300 K) and volume using weakly coupled Langevin dynamics of non-hydrogen atoms, with a damping coefficient of $c = 10\,\mathrm{ps}^{-1}$ with a 2 fs time step maintaining bonds to all hydrogen atoms rigid. Non-bonded interactions are truncated at 12 and 14 Å for van der Waals and electrostatic forces, respectively. Periodic boundary conditions are applied, and the particle mesh Ewald method is used to calculate electrostatic interactions.

All simulations using the NAMD package were run on IBM Blue Gene/L supercomputer at the Computational Center for Nanotechnology Innovations based in Rensselaer Polytechnic Institute, NY.

Stability of the simulated structures for the apical region

Our simulated structures contain 116 nt, of which 82 involve base pairs. To justify the stability of simulations maintaining secondary structure in trajectories, we computed average distance of base-paired residues in helices (5 bp in H_1, 3 bp in H_2, 8 bp in H_3, 5 bp in H_4, 13 bp in H_5, 3 bp in H_6 and 4 bp in H_7) and measured RMS deviations (RMSDs) considering all residues and only base pairs with reference to the starting structure as we hypothesize unpaired residues contribute to increase RMSD.

The average distances of all helices in our simulations are <3 Å and most of them are <2.5 Å (Supplementary Figure S2A). In addition, RMSD values (considering all except hydrogen atoms) are for the entire system (116 residues) including the paired 82 residues indicating overall stability (Supplementary Figure S2B).

Very recently, problems in χ torsion angles of MD simulations for RNA systems have been reported (namely, ladder-like structures that loose the helical twist of A-form RNA conformation, especially in long RNA MD simulations) (38,39), and improved force fields have been introduced (40,41). We have carefully monitored these potential problems but have not observed in our dynamics data.

Entire sequence modeling including the basal region

We model the basal region containing 98 nt (G_{86} to U_{133}, C_{249} to C_{299}) based on 2D information of FMDV C-S8 IRES domain 3 using MC-Sym. Evaluating all the 717 structures based on RMSD and clustering analysis yields four candidate models that were chosen from the first four large clusters containing at least 10 structures (Supplementary Figures S3–S4). Since overall shapes of these four candidates were similar, we chose a representative model from the largest cluster (Supplementary Figure S4A) to build a complete 3D model of domain 3. Structures of the apical and basal region were merged using a python library of modeRNA (42). Both minimization and equilibration were performed on the entire domain 3 following the protocol in the 'Molecular dynamics simulation for the apical region' section above.

RESULTS

Sequence conservation analysis of the apical region

Sequence similarity provides evidence for structural conservation and hence essential biological function.

Sequence logos of the aligned 318 sequences of FMDV IRES systems suggest that the apical region is largely conserved (Supplementary Figure S1), implying that its 2D structure is constrained under an evolutionary pressure to carry an important biological function for non-canonical IRES-mediated translation initiation. In particular, conservation of the potential binding receptors (G_{229} to C_{232} and G_{240} to C_{243}) of the GNRA loop is near perfect (316 out of 318 sequences); the sequence logos of H_4 marked in the red box (Supplementary Figure S1) indicate that the entire hairpin including the binding nucleotides are conserved almost fully. We also observe that the GUAA sequence appears most frequently with 233 instances (73.2%) followed by GUGA (17%), GCAA (7%), GCGA (2.5%) and GAGA (0.3%) (Supplementary Table S1).

Multiple four-way junction topology prediction at secondary structure level for the apical region

Initial guess of each four-way junction topology

We partition domain 3's 2D structure into two four-way junctions and list all possible junction topologies. We denote the two four-way junctions as Junctions I and II following the 5′- to 3′-direction (Figure 3). As loop size dictates orientation and flexibility of helices in RNA junctions (21,22), we build candidate topological models accordingly. Because very few nucleotides are present between helices in both junctions, we consider two coaxial stacking patterns, parallel to each other with a possible crossing at the point of single-strand exchange.

For Junction I, two types of pairwise coaxial stacking patterns are likely (because no nucleotides are present between helices). Helix H_1 can coaxially stack with either H_2 or H_4 (Figure 3A). This results in two coaxial stacking: H_1H_2 with H_3H_4 or H_1H_4 with H_2H_3, as shown in the figure.

Similarly, for Junction II we consider H_3H_5 with H_6H_7 or H_3H_7 with H_5H_6 (Figure 3B). However, we speculate that the latter pattern is more likely due to the presence of 2 nt in a single-strand loop between H_6 and H_7, whereas no nucleotides are in other single strands between coaxially stacked helices H_3H_7, H_5H_6 and H_3H_5 (see enlarged view in the middle of Figure 3); a strong preference for coaxial stacking has been observed with a smaller loop size (21,22,43). The Junction-Explorer program also predicts a pair of coaxial stacking formation for both four-way junctions, parallel to each other. On the basis of these combined models, we arrive at four candidate helical arrangements for each junction (Figure 3) that correspond to *H* and *cH* family types containing two coaxially stacked helices (see Supplementary Figure S5 for the nine major junction family types) based on our four-way junction classification study (21). Note that only three four-way junction families *H*, *cH* and *cL* contain two coaxially stacked helices and are distinguished by the angle between the two stacked helices with roughly 0°, 180° and 90°, respectively. To achieve a particular configuration for each family, different lengths of single strands between stacked helices are required. While the shape of family *H* and *cH* can be achieved with a relatively short single strand, family *cL* requires a long single strand. Because the junctions in domain 3 contain 2 nt in single-stranded regions at most, we do not consider the *cL* family as a candidate. Next, we consider these combinations of configurations compatible with experiment.

Full junction topology prediction

Considering the nine major junction family types in four-way RNA junctions (21), the number of ways to pair two four-way junctions is (9 family types × 2 different helical arrangements)^2 = 324 when no other information is considered. Using the two possible family types for each junction (Figure 3A and B), the number of likely conformations becomes (2 family types × 2 different helical arrangements)^2 = 16 (Figure 4B). We further consider the potential RNA–RNA long-range interactions between

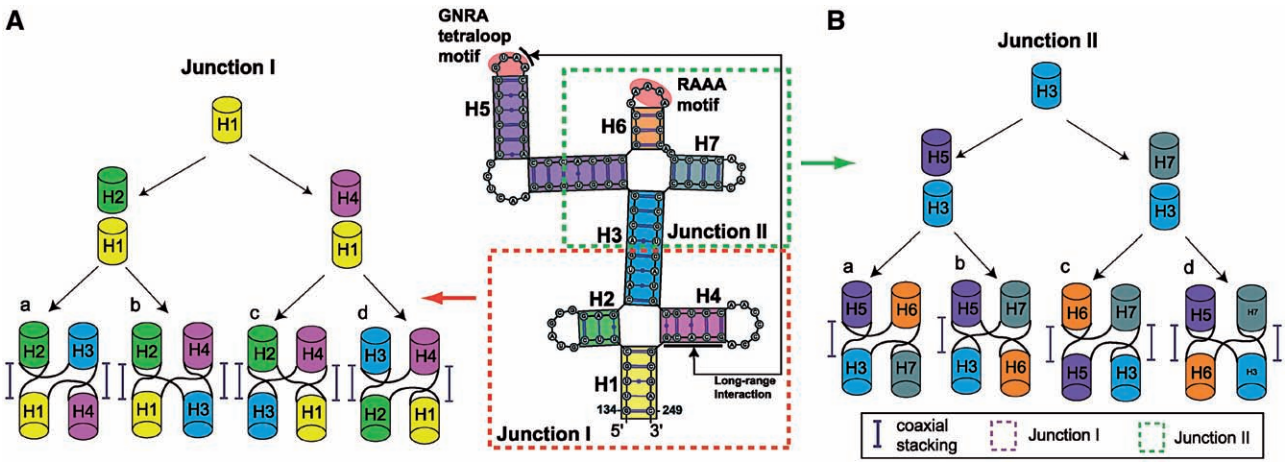

Figure 3. Piecing the possible helical arrangements for the two four-way junctions in domain 3 of IRES. (**A**) Possible junction topologies of junction I with two coaxial stacking following the 5′- to 3′-direction: (a) H_1H_4 and H_2H_3 without crossing in the single-stranded region, (b) H_1H_4 and H_2H_3 with crossing, (c) H_1H_2 and H_3H_4 with crossing and (d) H_1H_2 and H_3H_4 without crossing. (**B**) Predicted junction topologies of junction II with one or two coaxial stacking following the 5′- to 3′-direction: (a) H_3H_5 and H_6H_7 without crossing, (b) H_3H_5 and H_7H_6 with crossing, (c) H_5H_6 and H_7H_3 without crossing and (d) H_6H_5 and H_7H_3 with crossing.

the GNRA motif and its distal receptor region from experiment (9) to eliminate some of these 16.

Given that the GNRA tetraloop and its potential receptors are located in helices H_4 and H_5 (Figure 4A), we can eliminate some helical arrangements. To make the long-range interactions possible, both H4 and H5 are required, in positions either parallel or perpendicular with respect to H3. Thus, the two configurations *c* and *h* in Figure 4B can be eliminated because the bridging helix H_3 between Junction I and II is diagonal to H_4 and H_5 [see the helical arrangements in third row (for *c*) and fourth column (for *h*) in Supplementary Figure S6]; these two models are not eligible to make tertiary contacts between H_4 and H_5 due to either the orientation of these two helices that are opposite one another [(*c*, *e*), (*c*, *f*), (*c*, *h*), (*a*, *h*) and (*b*, *h*) in Supplementary Figure S6] or some steric clashes [(*c*, *g*) and (*d*, *h*) in Supplementary Figure S6]. In addition, the five pairs of helical arrangements (*a*, *e*), (*a*, *f*), (*b*, *e*), (*b*, *f*) and (*d*, *g*) in Figure 4B can also be excluded because helices H_3, H_4 and H_5 are aligned in the same direction (either to parallel or perpendicular). Four viable models remain (Figure 4C); the junctions in the final topology models correspond to either *H* or *cH* family types, and thus *H*/*H* or *H*/*cH* combinations are possible overall. See Supplementary Figure S6 for all 16 combinations elaborated from Figure 4B.

Modeling atomic junction structures for the apical region

Mutational analysis proposed a non-specific receptor site, $G_{240}CACG_{244}$ in H_4 of Junction I, for the $G_{178}UAA_{181}$ tetraloop of H_5 in Junction II (9). The two adenosines in the GUAA tetraloop prefer to interact with a pair of C/G base pairs or alternatively a combination of C/G and G/C base pairs. In the potential receptor site, we identify a combination of C_{232}/G_{240} and G_{231}/C_{241} base pairs that was reported as receptors of GUAA loop by an *in vitro* selection experiment (44). Note that the C_{232}/G_{240} pair is highly conserved in 130 FMDV sequences while the G_{231}/C_{241} pair is invariant (45) and thus probably significant for correct RNA folding.

Because the energetics of tertiary interactions have not yet been considered, at this stage we model the RNA–RNA long-range interactions by imposing a loose distance constraint of 10 Å between helices H_4 and H_5, specifically between A_{180} and $C_{232}G_{240}$ using C1′ atoms.

Sampling these constrained models using MC-Sym reduces the number of models to 267: in Junction I, 160 of these contain stacked helices of H_1H_2 with H_3H_4 while the remaining 107 structures contain stacking of H_1H_4 with H_2H_3. These numbers may reflect the preference of RNA's helical arrangements in Junction I. Until now, the coaxial stacking pattern—H_1H_2 and H_3H_4—appears to dominate the possibilities when long-range interactions are considered. After evaluating all the 267 structures by structural similarity based on RMSD, clustering analysis and visual inspection (Supplementary Figures S7 and S8), we arrive at the consensus with initial junction topology models (Figure 5). Some variations in H_6 and H_7 may occur due to flexibility introduced by two unpaired nucleotides in the single-stranded region.

Similarly, when we consider another invariant—G_{229}/C_{243} base pair—near the junction core for tertiary interactions, we obtain 52 viable structures; about one-fifth of the structures targeting the C_{232}/G_{240} base pair. We speculate that the relatively small number of sampled structures targeting near the junction core may explain unfavorable potential binding receptor. In fact, the representative 3D structures show that helical arrangements in Junction II are rather distorted than structured (Supplementary Figures S9–S11).

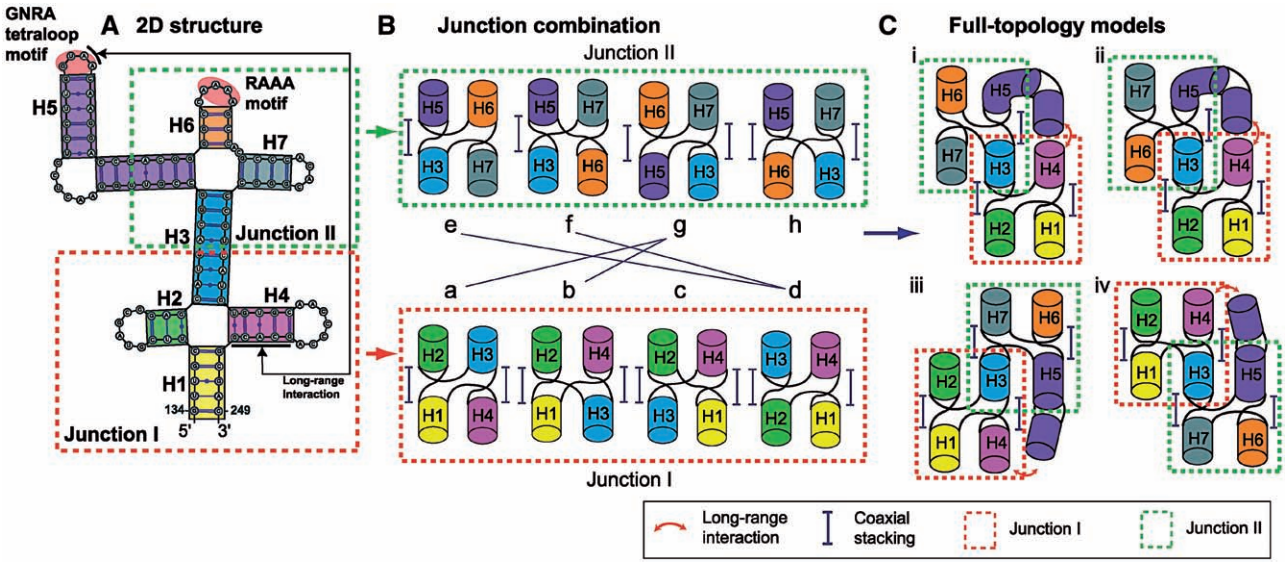

Figure 4. Four candidate models derived from combinations of two four-way Junctions I and II of Figure 3. (**A**) Secondary structure of domain 3 in FMDV IRES. (**B**) Combinations of the two four-way junctions considered. To accommodate the long-range interactions between helices H_4 and H_5, helix H_3 must be either parallel or perpendicular to the helices H_4 and H_5 in space. Two arrangements, *c* and *h*, of junctions I and II do not satisfy the long-range interactions and are thus eliminated. (**C**) Four complete junction topology models where Junctions I (dotted red box) and II (dotted green box) are stitched via helix H_3 considering the GNRA tetraloop long-range interaction between H_4 and H_5.

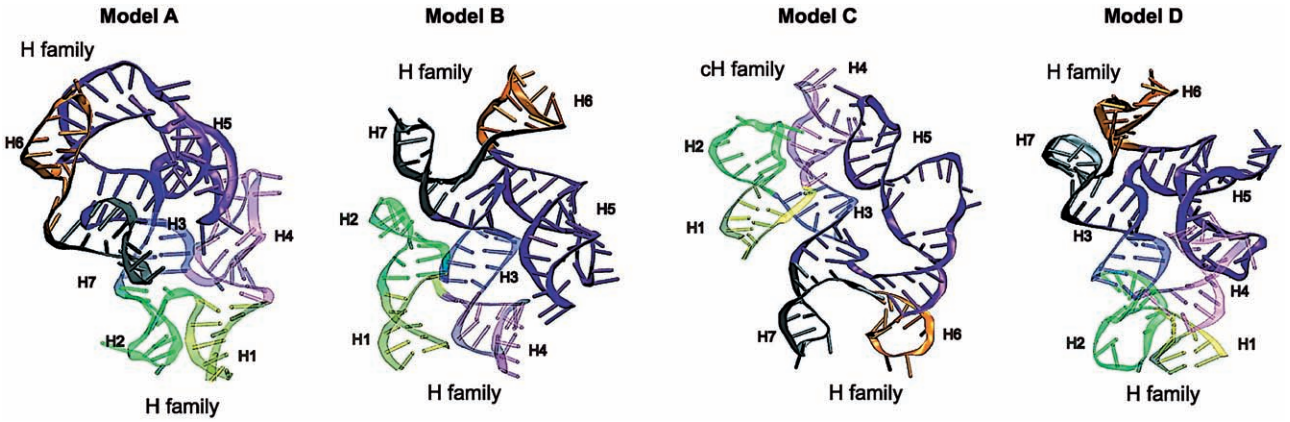

Figure 5. Four candidate 3D models of the FMDV IRES domain 3. Models A, B and C correspond to our junction topology models (Figure 4C), whereas Model D is new. All structures have two coaxial stacking both in Junctions I and II. A combination of *H*/*H* or *H*/*cH*, but not *cH*/*cH* family types is observed in the junctions.

We also arrive at a new candidate model, a modified combination of the junction topologies in Figure 4B(d) and (g) predicted by MC-Sym [see Supplementary Figure S6(d, g)*]. Although excluded while modeling junction topologies, this model may be possible due to a versatile nature of RNA molecules.

The four resulting candidate 3D models are shown in Figure 5. A, B and C correspond to the junction topology models in Figure 4C(i), (iii) and (iv), respectively, while D is a new 3D model. Note that we select four 3D models from six clusters (Supplementary Figure S9) considering similarity of overall helical arrangement. All structures have two coaxial stacking in Junctions I and II. Interestingly, the 3D model corresponding to the topology model in Figure 4C (ii) is not predicted by MC-Sym; this model, different from the three other topology models, has a crossing at the point of single-strand exchange in Junction II. We speculate that this particular helical arrangement in Junction II makes it difficult to satisfy the distance constraint criteria in 3D space.

We also explored different RNA conformations by simulating models over 40 ns by one-bead coarse-grained MD simulations using NAST. These simulations yield three representative conformations where helical arrangements are identical in Junction I, but some variations in Junction II (data not shown). Overall, these simulations lead further support to the models in Figure 5B and C.

Assessment of structural properties of the apical region using MD simulations

We use MD to supplement the structural studies above and to further explore the feasibility of our structural candidates. Despite algorithmic approximations as well as force field imperfections, MD is widely used to provide further insights into atomic-level interactions and energetic aspects that are not readily revealed from other techniques (46). Hence, we perform 100 ns MD simulations for all four candidate structures (Models A–D) in Figure 5; specifically to investigate structure stability and potential long-range interactions suggested by experimental data.

Long-range interactions including a novel tertiary contact revealed by MD simulation

Experimental data have proposed intramolecular long-range interactions in domain 3 of FMDV IRES (9,13). Although this tertiary contact is required for efficient IRES activity, the specific binding receptor of GNRA tetraloop is yet unknown. To explore this, we track for each trajectory the distances between the GUAA hairpin in H_5 and each of potential target receptors in H_4, specifically between $A_{180}A_{181}$ and $G_{240}CACG_{244}$ (including their complementary residues). Only for Model C we detected two receptor candidates—G_{231}/C_{241} and C_{232}/G_{240} base pairs—interacting with $A_{180}A_{181}$ residues in GUAA tetraloop (Figure 6). The trajectory for Model C shows that the two adenosines retain a distance <3 Å. In contrast, only the first adenosine A_{180} of Models A, B and D retain a distance <4 Å during the initial 12, 15 and 26 ns, respectively. In Model C, the average distance between C_{232}/G_{240} pair and A_{180} is 2.1 ± 0.59 Å while C_{231}/G_{241} pair and A_{181} is 2.0 ± 0.20 Å. These findings suggest that the C_{232}/G_{240} and C_{231}/G_{241} pairs may be the target receptors of A_{180} and A_{181} residues, respectively.

To further explore the tertiary interaction of Model C, we consider the Leontis/Westhof nomenclature (47) and analyze the three edges—Watson–Crick, Hoogsteen and Sugar edge—for potential hydrogen bonding interactions. The measured minimum distances between the Sugar edge of each C_{232}/G_{240} and C_{231}/G_{241} base pair with three edges of each A_{180} and A_{181} over the 100 ns time course in Figure 7 show tightly formed hydrogen bonding interactions for the Sugar edge/Watson–Crick between the C_{232}/G_{240} pair and A_{180} and Sugar edge/Hoogsteen edge tertiary interactions between G_{231}/C_{241} pair and A_{181}. In addition, we observe tertiary contacts between U_{179} and A_{234} residues via *trans* Watson–Crick/Watson–Crick edge interactions at ~22 ns. These long-range interactions occur sequentially: at ~7, 20 and 22 ns, involving A_{180}, A_{181} and U_{179}, respectively (Figure 7A). These cooperative long-range interactions help stabilize the IRES domain 3.

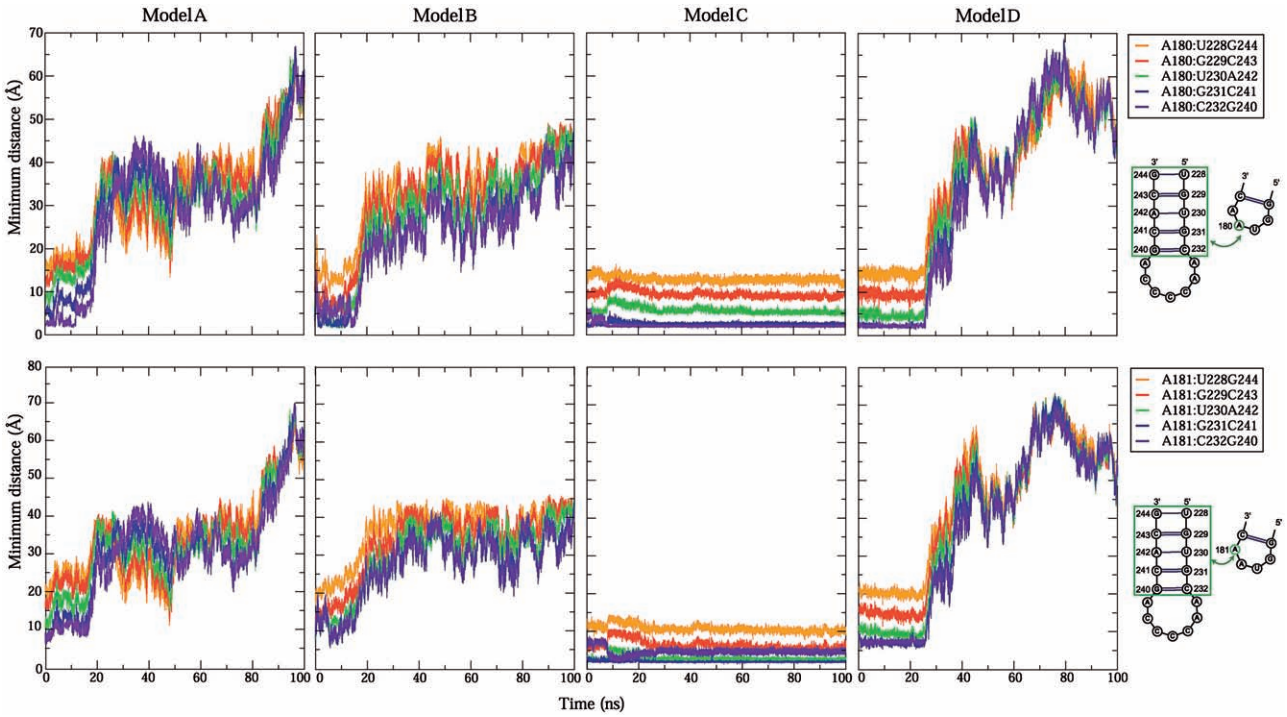

Figure 6. RNA–RNA long-range interactions identified by distance measures of atoms between two adenosine—A_{180} and A_{181}—in GUAA tetraloop and its potential receptors during the MD trajectories.

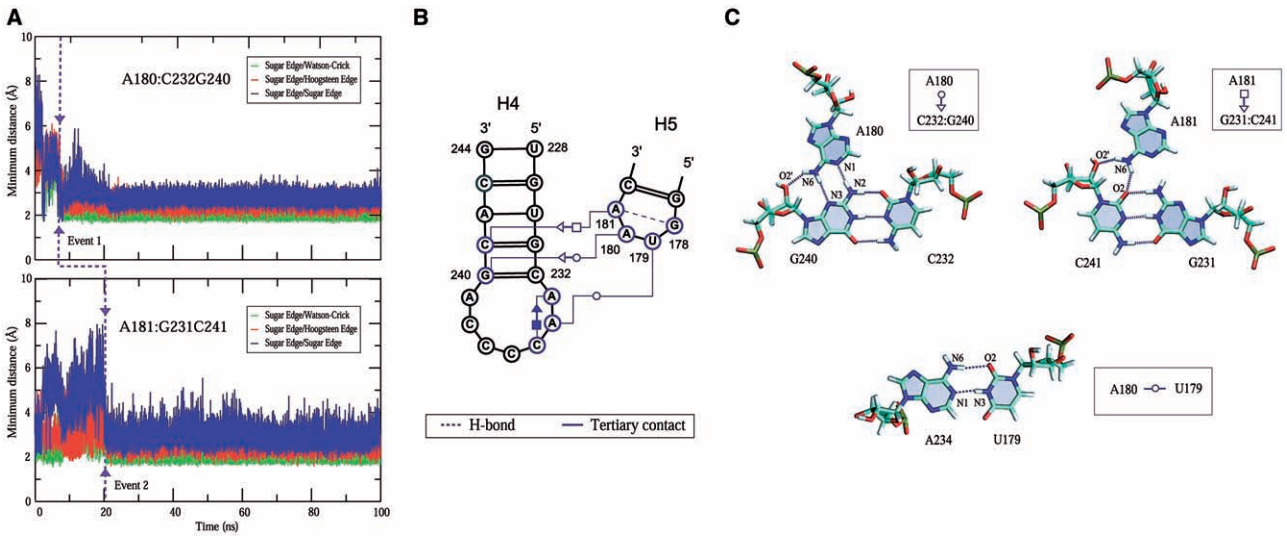

Figure 7. Intramolecular RNA–RNA long-range interactions involving GUAA hairpin loop during the Model C MD simulation. (**A**) A minimum distance of atoms for the GUAA tetraloop–receptor long-range interactions. (**B**) Long-range interactions in a time-averaged secondary structure obtained from the dynamics data. (**C**) Atomic details of these three tertiary contacts involving U_{179}, A_{180} and A_{181} residues in GUAA loop and their binding receptors.

The corresponding time-averaged secondary structure from the 100 ns dynamics data underscores these three long-range interactions involving the GUAA tetraloop (Figure 7B). The $A_{180}A_{181}$ residues in the GUAA tetraloop form hydrogen bonds via non-canonical base pairing interactions with the C_{232}/G_{240} and G_{231}/C_{241} base pairs, respectively; specifically, *trans* Sugar edge/Watson–Crick edge where N_1 and N_6 atoms of A_{180} interact with N_2, N_3 and O_2' atoms of G_{240} and Sugar edge/Hoogsteen edge interaction where N_6 atom of A_{181} forms hydrogen bonds with O_2 and O_2' atoms of C_{241} (Figure 7C). The U_{179} and A_{234} residues interact via *trans* Watson–Crick/Watson–Crick edge interactions involving N_1 and N_6 atoms of A_{234} with N_3 and O_2

atoms of U_{179}. As demonstrated by *in vitro* selection experiment (44), this U_{179}:A_{234} tertiary contact promotes the loop–helix long-range interactions.

How might these long-range interactions affect the structural organization of domain 3?

Because the single-stranded region is more dynamic and flexible than double-stranded helices, we speculate that the five hairpins and one long internal loop present in the system may contribute to the overall exhibit of the structure. Thus, the tertiary contacts in domain 3 of IRES may restrict these fluctuations and therefore help recruit ribosomes for viral protein synthesis. Below we further analyze dynamics data for all four models to discern the contributions of the long-range interactions to structural stability as well as organization based on root-mean-square (RMS) fluctuation and radius of gyration (Rg).

In the RMS fluctuation plot (Figure 8A), we observe six peaks that correspond to hairpins and an internal loop. Among them, two highest peaks are from hairpins located in the helices H_4 and H_5. Interestingly, these helices involve in the long-range interactions and have been emphasized for its important role in IRES activity.

Overall, the RMS fluctuations of all four models follow a similar trend, albeit at different scales. Overall, Models A, B and D fluctuate widely with the values from ~2.5 to ~21 Å, whereas Model C ranges between ~2 and ~7.5 Å which is about a four-fold decrease. Notably, the GUAA tetraloop in H_4 fluctuates between 12 and 18 Å for Models A, B and D, whereas Model C experiences only ~2.5 Å deviation; this underscores the potential stabilizing role of the tertiary contacts. The long-range interactions appear to stabilize not only adjacent stem–loops but also the entire structure of IRES domain 3.

From the combined data above, involving bioinformatic, experimental and MD modeling data, we propose a theoretically feasible tertiary structure for the apical region in FMDV IRES domain 3 (Figure 9). Here the non-canonical long-range interactions occur between the GUAA tetraloop in helix H_5 and the distal region in helix H_4. The overall configuration is highly structured; each four-way junction contains two coaxial stacks parallel to each other and Junction I has a crossing at the point of strand exchange. Junctions I and II are classified as family type *cH* and *H*, respectively, according to the nomenclature in (21). The three helices H_3, H_4 and H_7 are coaxially stacked together. Each of these two four-way junctions is nearly planar and these two planes are perpendicular to each other.

Modeling of domain 3 including the basal region

The basal region contains a long internal loop formed by 98 nt (G_{86} to U_{133} and C_{249} to C_{299}). Sampling this region using MC-Sym produces 717 models from which four representative structures were selected from the four largest clusters (Supplementary Figures S3–S4). The overall shape and orientation of the helical axis in these four structures are relatively similar, with differences explained by the flexibility of bending in unpaired bases. Thus, our candidate model for the basal region is chosen from the largest cluster (Supplementary Figure S4A), and then the apical region model (Figure 9) was combined to it to complete a 3D model of the entire domain 3; the minimum distance between the basal (C_{250}) and apical (G_{194}) regions is 23.5 Å; the orientation of the basal region (turned away from the apical region) suggests that the former region is unlikely to be involved in RNA folding of the junctions of the latter region (Figure 10).

DISCUSSION

Picornavirus IRES elements are considered as efficient regulatory RNAs which make possible initiation of translation for viral RNAs. FMDV requires RNA-binding proteins such as translation initiation factors (eIFs) and ITAFs that can affect IRES activity; for instance, domains 2, 4 and 5 provide binding sites for cellular proteins including PTB, eIF4G, eIF3 and eIF4B (48).

The FMDV IRES domain 3, often denoted as a central domain, consists of two structural elements—a long internal loop in the 'basal' region and four-way junctions

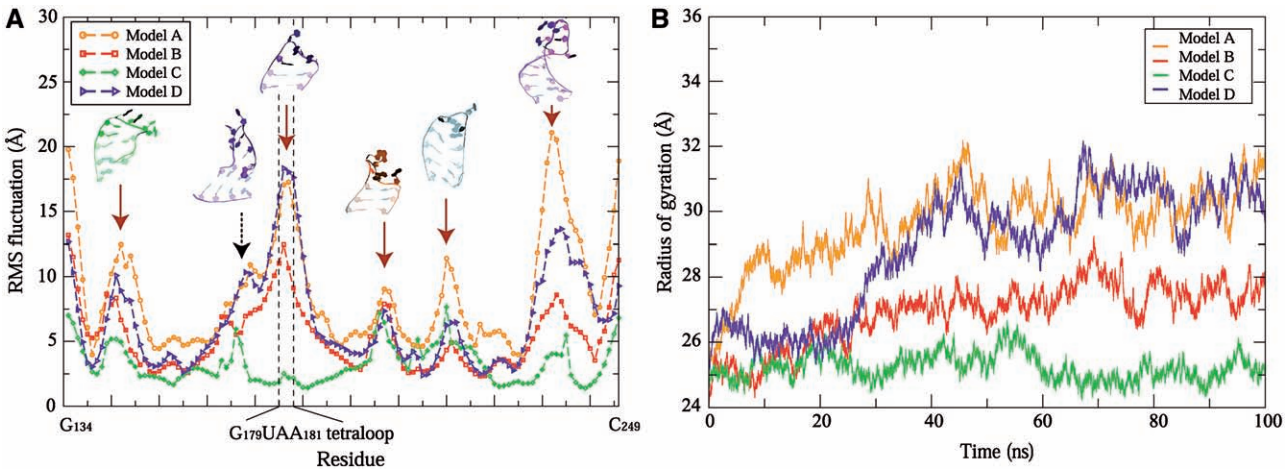

Figure 8. RMS fluctuations (**A**) and Rg (**B**) measures for four candidate 3D models. In the RMS fluctuations, high peaks (dotted black arrow for internal loop and solid brown arrows for hairpins) correspond to unpaired regions shown as solid color in the structures.

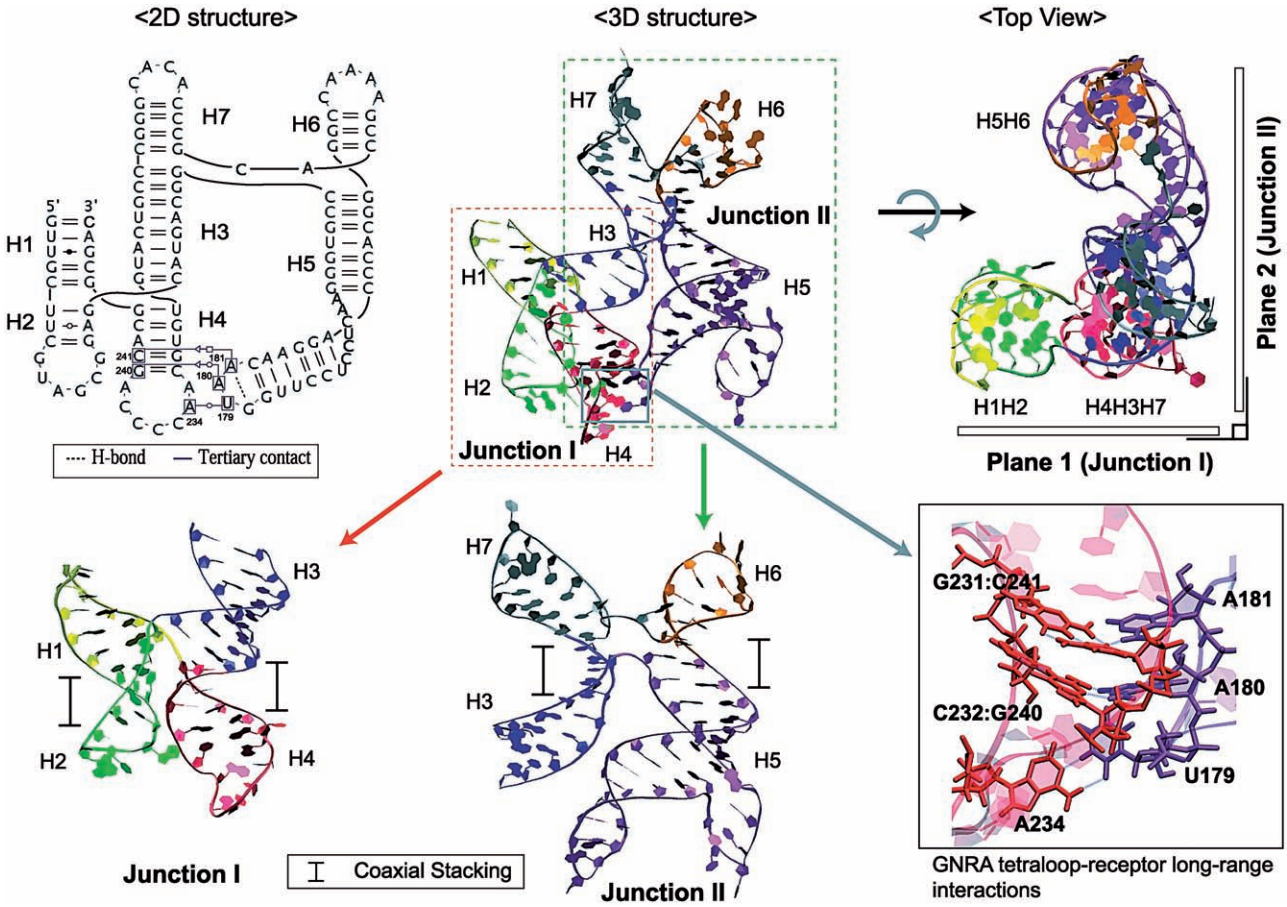

Figure 9. Time-averaged tertiary structure of domain 3 taken from the 100 ns dynamics data (top middle), where the long-range interactions occur between helices H_4 and H_5 (details shown at bottom right). Both Junctions I and II contain two coaxial stacking, parallel to each other and Junction I with a crossing in the single-stranded region (bottom-left and -middle for Junctions I and II, respectively). Both junctions are planar locally and are arranged in a perpendicular orientation to each other globally (top right); note that the three helices H_4, H_3 and H_7 are coaxially stacked all together.

in the 'apical' region; each of which is ~50% of the entire sequence. Sequence logos of the 318 aligned FMDV IRES sequences show that the apical region of domain 3 is highly conserved (Supplementary Figure S1); in particular, hairpin loop H_4 that contains the potential binding receptors of the GNRA tetraloop is nearly perfectly conserved. The GUAA loop in H_5 is also strongly preferred in FMDV IRES systems. This analysis suggests that these structural elements provide an important role in maintaining the functional 3D structure of FMDV IRES domain 3.

It was determined by biochemical experiments that the apical region is a self-folding structural element due to the intramolecular RNA–RNA interactions involving the crucial GNRA motif (13,49,50). This region has thus been suggested to contribute significantly to the structural organization and stability of domain 3, and to the critical function of IRES activity (8–11). IRES-mediated translation initiation is closely linked to structural organization in domain 3, specifically the apical region formed by two four-way junctions enabling the RNA–RNA intramolecular interactions. Thus, we focused on the apical region of domain 3 to decipher the spatial arrangement of the RNA fold that is a prerequisite essential step to understand the initiation mechanism of translation.

Grounded in our recent RNA four-way junction classification study and the Junction-Explorer program (21,27), we have constructed possible junction topologies for the apical region in domain 3 where a pair of coaxial stacks are arranged parallel to each other in the presence and absence of a crossing at the point of strand exchange (Figure 9). Utilizing only the information for the helical arrangements—*H* and *cH* family types, we built 16 candidate topologies (Figure 4) and these were reduced to four after applying constraints from experimental data regarding the GNRA tetraloop–receptor long-range interactions. Our next step in modeling was employing MC-Sym to explore conformational space using experimental data. The combined data from junction topology and 3D modeling produced four representative structures where three of the four confirmed the junction topology models. We speculate that the excluded model does not satisfy geometric criteria due to steric clashes. Using MD simulations, we attempted to identify geometrically accessible binding receptors to the GNRA tetraloop for long-range interactions. Among the five residues in the potential receptor site, the bases near the hairpin loop emerged viable over one near the junction core in helix H_4; they also have great potential to form long-range interactions with the GNRA tetraloop in H_5.

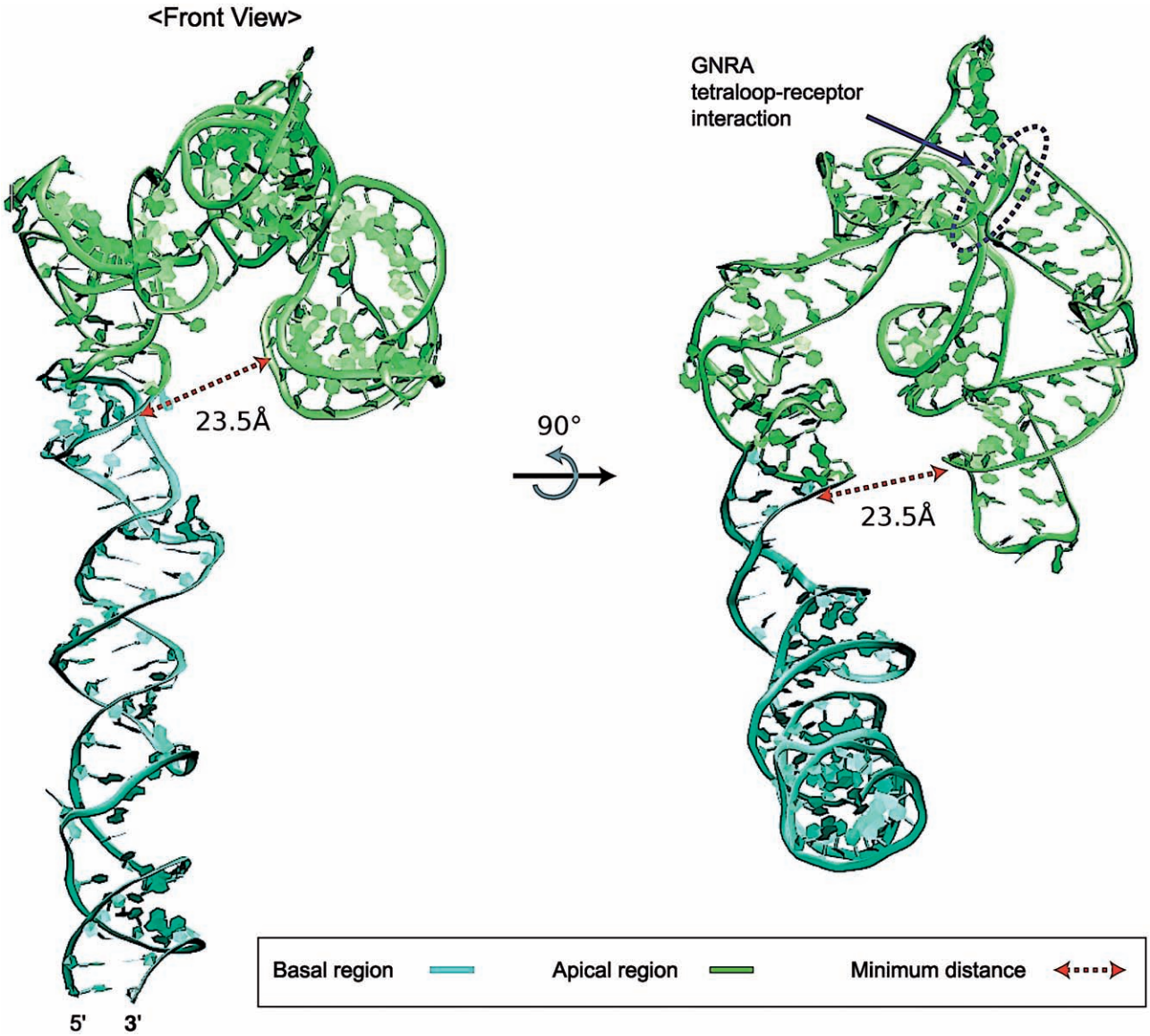

Figure 10. 3D model of the entire sequence in domain 3. Domain 3 consists of basal and apical regions; the corresponding structural elements are an internal loop and four-way junctions, respectively. The minimum distance between the two regions is 23.5 Å that the basal region is not likely involved in RNA folding of the junctions.

Specifically, the MD simulations revealed a GUAA tetraloop-binding site in addition to a novel tertiary interaction in Model C (Figure 5). The two adenosines A_{180} and A_{181} form hydrogen bonds with the receptors C_{230}/G_{242} and G_{231}/C_{241} base pair, respectively. The dynamics data also suggest a U:A tertiary contact which enhances the structural stability (U_{179} in GUAA tetraloop interacts with A_{234} in a hairpin loop of H_4), an interaction also observed by an *in vitro* selection experiment (44). Interestingly, these tertiary interactions form sequentially.

A previous study suggested that RNA–RNA long-range interactions involving an RAAA motif occur in the presence of GNRA tetraloop long-range interactions as well as Mg^{2+} ions (9). We have not observed this RAAA motif, but speculate that the distant contacts associated with the RAAA motif might occur when the current RNA system ($G_{134}\ldots C_{249}$ residues) is extended to include 25 more residues ($U_{121}\ldots G_{133}$, $U_{250}\ldots A_{261}$). Such an extended system has greater potential for the long-range interactions (Supplementary Supplementary Figure S12). Due to current limitations of the force field for treating divalent ions (51–54), we have not attempted to include magnesium ions in our system.

In picornavirus, types 1 and 2 of IRES species exist. FMDV IRES belongs to type 2, whereas poliovirus IRES belongs to type 1. Although the overall contents of sequence and 2D structure are different, these two IRES systems share the GNRA motif. NMR data of stem–loop in domain IV revealed the L-shaped conformations are required in order to provide a protein-binding site. However, little is known how the L shape is achieved and maintained. In FMDV IRES domain 3, helix H_5 contains similar 2D structure of loop B which includes the GNRA motif. In our dynamics simulations, we observe

that H_5 forms an L-shape configuration in the presence of long-range RNA–RNA interactions. The overall shape of H_5 agrees well with the NMR data (Supplementary Figure S13). However, the shape of H_5 in the absence of long-range interactions is variable, with potential diverse phases (S-shaped or U-shaped). Thus, we speculate that long-range interactions involving the GNRA motif have a role in stabilizing the L-shaped loop B in poliovirus IRES.

Based on the above modeling of the apical, self-folding region of IRES domain 3 containing four-way junctions and the experimental data discussed above (13,49,50) combined with our extended modeling of the entire sequence of FMDV IRES domain 3, we hypothesize that the influence of the basal region on structural stability and organization of the junctions is not primary. This is because the basal region is set apart from the junction domains in the apical region with a minimum distance of 23.5 Å (Figure 10).

Although our combined modeling strategy involves many proven approaches and is closely anchored to available experimental data, it is not possible to rule out other plausible overall 3D structures. Further studies using the candidate models for long-time MD studies or with advanced sampling techniques and investigation of potential receptors for RAAA motif may be useful. Yet, the overall 3D configuration reached in Figure 9 and the suggested long-range interactions in the central domain of FMDV IRES provide insights into the potential role of the long-range interactions for structural stability and organization of IRES domain 3 and thus may help in further analysis of the structure, mechanism, and function of viral RNAs. Ultimately, structures may lead to the development of antiviral drugs that inhibit IRES activity and thus virus multiplication.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–3, Supplementary Figures 1–14 and Supplementary PDB file.

ACKNOWLEDGEMENTS

The authors thank Dr Christian Laing for discussions related to the virus IRES structure, Dr Marc Parisien for help with MC-Sym and Mr Elmetwaly for technical assistance. The computations were made possible by support of the IBM Blue Gene/L supercomputer at the Computational Center for Nanotechnology Innovations based in Rensselaer Polytechnic Institute, NY. Molecular images were generated using the VMD program (55) and Gromacs tool packages (56) were used to analyze the dynamics data.

FUNDING

The National Science Foundation [EMT award CCF-0727001 to T.S.]; National Institutes of Health [grants GM081410 and GM100469 to T.S.]; MacCracken Fellowship (to S.J.). Funding for open access charge: National Science Foundation, National Institutes of Health.

Conflict of interest statement. None declared.

REFERENCES

1. Kolupaeva,V.G., Hellen,C.U. and Shatsky,I.N. (1996) Structural analysis of the interaction of the pyrimidine tract-binding protein with the internal ribosomal entry site of encephalomyocarditis virus and foot-and-mouth disease virus RNAs. *RNA*, **2**, 1199–1212.
2. Lopez de Quinto,S. and Martinez-Salas,E. (2000) Interaction of the eIF4G initiation factor with the aphthovirus IRES is essential for internal translation initiation in vivo. *RNA*, **6**, 1380–1392.
3. Pilipenko,E.V., Pestova,T.V., Kolupaeva,V.G., Khitrina,E.V., Poperechnaya,A.N., Agol,V.I. and Hellen,C.U. (2000) A cell cycle-dependent protein serves as a template-specific translation initiation factor. *Genes Dev.*, **14**, 2028–2045.
4. Lopez de Quinto,S., Lafuente,E. and Martinez-Salas,E. (2001) IRES interaction with translation initiation factors: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII. *RNA*, **7**, 1213–1226.
5. Stassinopoulos,I.A. and Belsham,G.J. (2001) A novel protein-RNA binding assay: functional interactions of the foot-and-mouth disease virus internal ribosome entry site with cellular proteins. *RNA*, **7**, 114–122.
6. Andreev,D.E., Fernandez-Miragall,O., Ramajo,J., Dmitriev,S.E., Terenin,I.M., Martinez-Salas,E. and Shatsky,I.N. (2007) Differential factor requirement to assemble translation initiation complexes at the alternative start codons of foot-and-mouth disease virus RNA. *RNA*, **13**, 1366–1374.
7. Pacheco,A., Lopez de Quinto,S., Ramajo,J., Fernandez,N. and Martinez-Salas,E. (2009) A novel role for Gemin5 in mRNA translation. *Nucleic Acids Res.*, **37**, 582–590.
8. Lopez de Quinto,S. and Martinez-Salas,E. (1997) Conserved structural motifs located in distal loops of aphthovirus internal ribosome entry site domain 3 are required for internal initiation of translation. *J. Virol.*, **71**, 4171–4175.
9. Fernandez-Miragall,O., Ramos,R., Ramajo,J. and Martinez-Salas,E. (2006) Evidence of reciprocal tertiary interactions between conserved motifs involved in organizing RNA structure essential for internal initiation of translation. *RNA*, **12**, 223–234.
10. Fernandez,N., Fernandez-Miragall,O., Ramajo,J., Garcia-Sacristan,A., Bellora,N., Eyras,E., Briones,C. and Martinez-Salas,E. (2011) Structural basis for the biological relevance of the invariant apical stem in IRES-mediated translation. *Nucleic Acids Res.*, **39**, 8572–8585.
11. Fernandez,N., Garcia-Sacristan,A., Ramajo,J., Briones,C. and Martinez-Salas,E. (2011) Structural analysis provides insights into the modular organization of picornavirus IRES. *Virology*, **409**, 251–261.
12. Correll,C.C. and Swinger,K. (2003) Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 A resolution. *RNA*, **9**, 355–363.
13. Fernandez-Miragall,O. and Martinez-Salas,E. (2003) Structural organization of a viral IRES depends on the integrity of the GNRA motif. *RNA*, **9**, 1333–1344.
14. Du,Z., Ulyanov,N.B., Yu,J., Andino,R. and James,T.L. (2004) NMR structures of loop B RNAs from the stem-loop IV domain of the enterovirus internal ribosome entry site: a single C to U substitution drastically changes the shape and flexibility of RNA. *Biochemistry*, **43**, 5757–5771.
15. Robertson,M.E., Seamons,R.A. and Belsham,G.J. (1999) A selection system for functional internal ribosome entry site (IRES) elements: analysis of the requirement for a conserved GNRA tetraloop in the encephalomyocarditis virus IRES. *RNA*, **5**, 1167–1179.
16. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.

17. Lilley,D.M., Clegg,R.M., Diekmann,S., Seeman,N.C., Von Kitzing,E. and Hagerman,P.J. (1995) A nomenclature of junctions and branchpoints in nucleic acids. *Nucleic Acids Res.*, **23**, 3363–3364.
18. Kieft,J.S., Zhou,K., Grech,A., Jubin,R. and Doudna,J.A. (2002) Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. *Nat. Struct. Biol.*, **9**, 370–374.
19. Bailor,M.H., Sun,X. and Al-Hashimi,H.M. (2010) Topology links RNA secondary structure with global conformation, dynamics, and adaptation. *Science*, **327**, 202–206.
20. Laing,C., Jung,S., Iqbal,A. and Schlick,T. (2009) Tertiary motifs revealed in analyses of higher-order RNA junctions. *J. Mol. Biol.*, **393**, 67–82.
21. Laing,C. and Schlick,T. (2009) Analysis of four-way junctions in RNA structures. *J. Mol. Biol.*, **390**, 547–559.
22. Lescoute,A. and Westhof,E. (2006) Topology of three-way junctions in folded RNAs. *RNA*, **12**, 83–93.
23. Kim,S.H., Sussman,J.L., Suddath,F.L., Quigley,G.J., McPherson,A., Wang,A.H., Seeman,N.C. and Rich,A. (1974) The general structure of transfer RNA molecules. *Proc. Natl Acad. Sci. USA*, **71**, 4970–4974.
24. Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.
25. Laing,C. and Schlick,T. (2011) Computational approaches to RNA structure prediction, analysis, and design. *Curr. Opin. Struct. Biol.*, **21**, 306–318.
26. Laing,C. and Schlick,T. (2010) Computational approaches to 3D modeling of RNA. *J. Phys. Condens. Matter*, **22**, 283101.
27. Laing,C., Wen,D., Wang,J.T. and Schlick,T. (2011) Predicting coaxial helical stacking in RNA junctions. *Nucleic Acids Res.*, **40**, 487–498.
28. Pilipenko,E.V., Blinov,V.M., Chernov,B.K., Dmitrieva,T.M. and Agol,V.I. (1989) Conservation of the secondary structure elements of the 5′-untranslated region of cardio- and aphthovirus RNAs. *Nucleic Acids Res.*, **17**, 5701–5711.
29. Parisien,M. and Major,F. (2008) The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature*, **452**, 51–55.
30. Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J., Ostell,J. and Sayers,E.W. (2011) GenBank. *Nucleic Acids Res.*, **39**, D32–D37.
31. Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
32. Chang,T.H., Horng,J.T. and Huang,H.D. (2008) RNALogo: a new approach to display structural RNA alignment. *Nucleic Acids Res.*, **36**, W91–W96.
33. Breiman,L. (2001) Random forests. *Mach. Learn.*, **45**, 5–32.
34. Jonikas,M.A., Radmer,R.J., Laederach,A., Das,R., Pearlman,S., Herschlag,D. and Altman,R.B. (2009) Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA*, **15**, 189–199.
35. Perez,A., Marchan,I., Svozil,D., Sponer,J., Cheatham,T.E. 3rd, Laughton,C.A. and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.*, **92**, 3817–3829.
36. Cornell,W.D., Cieplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 5179–5197.
37. Phillips,J.C., Braun,R., Wang,W., Gumbart,J., Tajkhorshid,E., Villa,E., Chipot,C., Skeel,R.D., Kale,L. and Schulten,K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**, 1781–1802.
38. Mlynsky,V., Banas,P., Hollas,D., Reblova,K., Walter,N.G., Sponer,J. and Otyepka,M. (2010) Extensive molecular dynamics simulations showing that canonical G8 and protonated A38H+ forms are most consistent with crystal structures of hairpin ribozyme. *J. Phys. Chem. B*, **114**, 6642–6652.
39. Reblova,K., Sponer,J. and Lankas,F. (2012) Structure and mechanical properties of the ribosomal L1 stalk three-way junction. *Nucleic Acids Res.*, **40**, 6290–6303.
40. Zgarbova,M., Otyepka,M., Sponer,J., Mladek,A., Banas,P., Cheatham,T.E. 3rd and Jurecka,P. (2011) Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.*, **7**, 2886–2902.
41. Yildirim,I., Kennedy,S.D., Stern,H.A., Hart,J.M., Kierzek,R. and Turner,D.H. (2012) Revision of AMBER torsional parameters for RNA improves free energy predictions for tetramer duplexes with GC and iGiC base pairs. *J. Chem. Theory Comput.*, **8**, 172–181.
42. Rother,M., Rother,K., Puton,T. and Bujnicki,J.M. (2011) ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucleic Acids Res.*, **39**, 4007–4022.
43. Kim,S.H. and Cech,T.R. (1987) Three-dimensional model of the active site of the self-splicing rRNA precursor of Tetrahymena. *Proc. Natl Acad. Sci. USA*, **84**, 8788–8792.
44. Costa,M. and Michel,F. (1997) Rules for RNA recognition of GNRA tetraloops deduced by in vitro selection: comparison with in vivo evolution. *EMBO J.*, **16**, 3289–3302.
45. Burks,J.M., Zwieb,C., Muller,F., Wower,I.K. and Wower,J. (2011) In silico analysis of IRES RNAs of foot-and-mouth disease virus and related picornaviruses. *Arch. Virol.*, **156**, 1737–1747.
46. Schlick,T., Collepardo-Guevara,R., Halvorsen,L.A., Jung,S. and Xiao,X. (2011) Biomolecularmodeling and simulation: a field coming of age. *Q. Rev. Biophys.*, **44**, 191–228.
47. Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.
48. Martinez-Salas,E. (2008) The impact of RNA structure on picornavirus IRES activity. *Trends Microbiol.*, **16**, 230–237.
49. Ramos,R. and Martinez-Salas,E. (1999) Long-range RNA interactions between structural domains of the aphthovirus internal ribosome entry site (IRES). *RNA*, **5**, 1374–1383.
50. Serrano,P., Gomez,J. and Martinez-Salas,E. (2007) Characterization of a cyanobacterial RNase P ribozyme recognition motif in the IRES of foot-and-mouth disease virus reveals a unique structural element. *RNA*, **13**, 849–859.
51. McDowell,S.E., Spackova,N., Sponer,J. and Walter,N.G. (2007) Molecular dynamics simulations of RNA: an in silico single molecule approach. *Biopolymers*, **85**, 169–184.
52. Sponer,J., Burda,J.M.S., Leszczynski,J. and Hobza,P. (1998) Interaction between the guanine-cytosine Watson-Crick DNA base pair and hydrated group IIa (Mg2+, Ca2+, Sr2+, Ba2+) and group IIb (Zn2+, Cd2+, Hg2+) metal cations. *J. Phys. Chem. A*, **102**, 5951–5957.
53. Reblova,K., Lankas,F., Razga,F., Krasovska,M.V., Koca,J. and Sponer,J. (2006) Structure, dynamics, and elasticity of free 16 s rRNA helix 44 studied by molecular dynamics simulations. *Biopolymers*, **82**, 504–520.
54. Razga,F., Koca,J., Sponer,J. and Leontis,N.B. (2005) Hinge-like motions in RNA kink-turns: the role of the second a-minor motif and nominally unpaired bases. *Biophys. J.*, **88**, 3466–3485.
55. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 27–38.
56. Van Der Spoel,D., Lindahl,E., Hess,B., Groenhof,G., Mark,A.E. and Berendsen,H.J. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.*, **26**, 1701–1718.