THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Bioinformatics and Molecular Dynamics Simulation Study of L1 Stalk Non-Canonical rRNA Elements: Kink-Turns, Loops, and Tetraloops

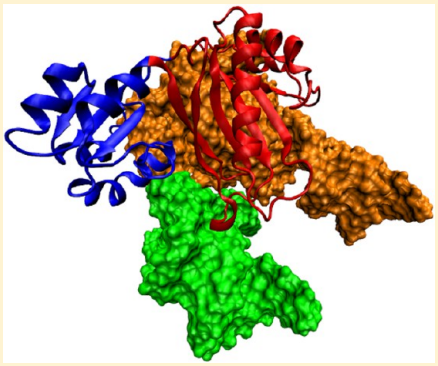
Miroslav Krepl,[†] Kamila Réblová,[‡] Jaroslav Koča,[‡] and Jiří Šponer*[,†,‡]

[†]Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic

[‡]CEITEC − Central European Institute of Technology, Campus Bohunice, Kamenice 5, 625 00 Brno, Czech Republic

Supporting Information

ABSTRACT: The L1 stalk is a prominent mobile element of the large ribosomal subunit. We explore the structure and dynamics of its non-canonical rRNA elements, which include two kink-turns, an internal loop, and a tetraloop. We use bioinformatics to identify the L1 stalk RNA conservation patterns and carry out over 11.5 μs of MD simulations for a set of systems ranging from isolated RNA building blocks up to complexes of L1 stalk rRNA with the L1 protein and tRNA fragment. We show that the L1 stalk tetraloop has an unusual GNNA or UNNG conservation pattern deviating from major GNRA and YNMG RNA tetraloop families. We suggest that this deviation is related to a highly conserved tertiary contact within the L1 stalk. The available X-ray structures contain only UCCG tetraloops which in addition differ in orientation (*anti* vs *syn*) of the guanine. Our analysis suggests that the *anti* orientation might be a mis-refinement, although even the *anti* interaction would be compatible with the sequence pattern and observed tertiary interaction. Alternatively, the *anti* conformation may be a real substate whose population could be pH-dependent, since the guanine *syn* orientation requires protonation of cytosine in the tertiary contact. In absence of structural data, we use molecular modeling to explore the GCCA tetraloop that is dominant in bacteria and suggest that the GCCA tetraloop is structurally similar to the YNMG tetraloop. Kink-turn Kt-77 is unusual due to its 11-nucleotide bulge. The simulations indicate that the long bulge is a stalk-specific eight-nucleotide insertion into consensual kink-turn only subtly modifying its structural dynamics. We discuss a possible evolutionary role of helix H78 and a mechanism of L1 stalk interaction with tRNA. We also assess the simulation methodology. The simulations provide a good description of the studied systems with the latest bsc0χ_{OL3} force field showing improved performance. Still, even bsc0χ_{OL3} is unable to fully stabilize an essential sugar-edge H-bond between the bulge and non-canonical stem of the kink-turn. Inclusion of Mg^{2+} ions may deteriorate the simulations. On the other hand, monovalent ions can in simulations readily occupy experimental Mg^{2+} binding sites.



INTRODUCTION

The L1 stalk is a prominent protuberance of the large ribosomal subunit, composed of rRNA helices 76−78 and L1 protein. Many experiments, such as FRET data,[1−8] revealed large-scale movements in this region, suggesting the L1 stalk is the most dynamical segment of the ribosome.[9] Its movement has been implicated in release of deacylated tRNA from the ribosome. In relation to the large ribosomal subunit, the L1 stalk may adopt three distinct conformations: open (no contact with tRNA, non-ratcheted state), half-closed (tRNA in E/E site, non-ratcheted state), and closed (tRNA in P/E site, ratcheted state).[4,10] The L1 stalk is in direct contact with deacylated tRNA in the half-closed and closed conformations. This contact is maintained through the entire translocation step.

The atomistic structure of isolated L1 stalk was first obtained in 2003.[11] The L1 protein from *Sulfolobus acidocaldarius* was crystallized with the 23S rRNA fragment from *Thermus thermophilus* (*T.t.* in the following) (PDB: 1MZP). The 23S rRNA fragment was synthetically created as a minimalistic version of the L1 stalk RNA to measure its binding energy with L1 protein.[12] While this structure was later revealed to be different from wild-type *T.t.* L1 stalk (due to the missing portion of H78 helix and subsequent changes in protein−RNA interface), it shows an accurate picture of the remaining L1 stalk RNA interactions.

Atomistic resolution structures of L1 stalk in the ribosome are difficult to obtain, as large fluctuations in this region preclude its visualization at higher resolution. Nevertheless, a 3.6 Å resolution atomistic structure of L1 stalk in the context of the entire ribosome was obtained in 2009[13] (PDB: 2WRJ; Figure 1) when the *T.t.* ribosome was trapped in posttranslocational state, with the L1 stalk bent inward into intersubunit space to contact tRNA. Finally, the isolated structure of the L1 stalk from *T.t.* was resolved in 2012 at 2.0 Å resolution (PDB: 3U4M).[14] It largely validated the 2WRJ ribosomal structure from 2009 with few differences that we will specifically discuss where relevant. The remaining interactions are virtually identical in both structures, and we will discuss them as "bacterial

Received: February 10, 2013
Revised: March 24, 2013

   dx.doi.org/10.1021/jp401482m | *J. Phys. Chem. B* XXXX, XXX, XXX−XXX

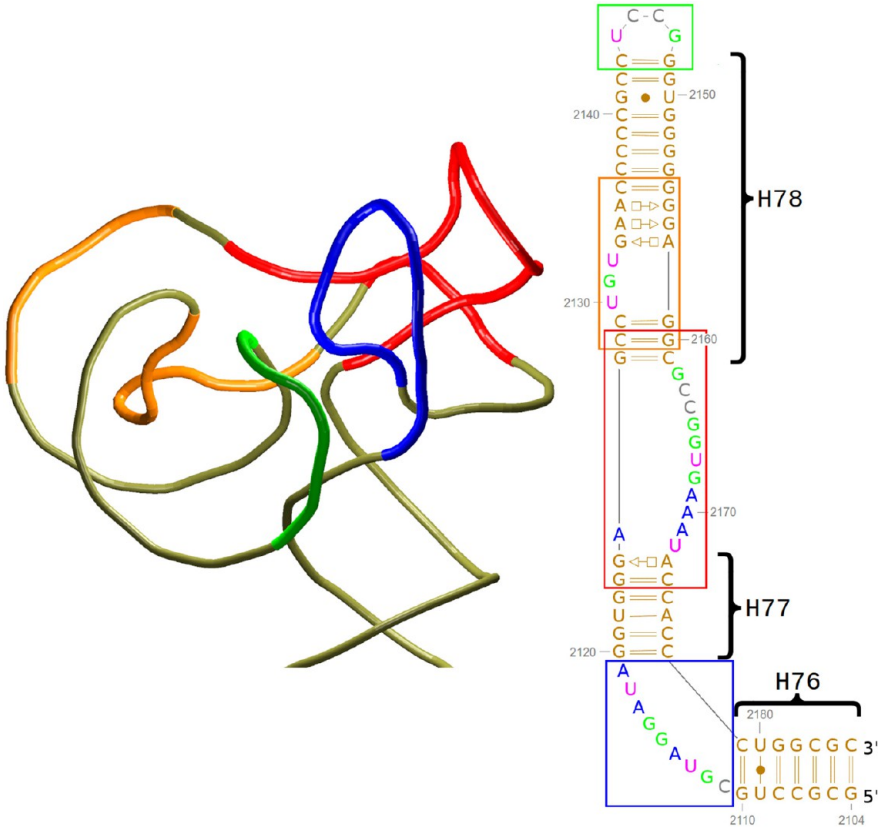


Figure 1. Tertiary and secondary structure of the *T.t.* L1 stalk. The non-canonical motifs are highlighted. Blue is L2111/2119 (nts 2111−2119), red is kink-turn Kt-77 (nts 2124−2128 and 2160−2174), orange is kink-turn Kt-78 (nts 2128−2136 and 2155−2160), and green is UCCG (nts 2143-2148) tetraloop. Standard base pair classification is used.[37] Curly brackets indicate the ribosomal helices.

structure" interactions. As we mainly used the 2WRJ structure as a source for our simulations, quality of the initial geometry was of some concern to us. However, comparison with the higher-resolution 3U4M structure indicates that it was sufficient for such a task.

Full-atom structures of eukaryotic ribosome became available in the course of our work. In these structures, the L1 stalk is visible at 4.0 Å (near-atomic) resolution[15] but becomes disordered at 3.0 Å resolution.[16] This precludes atomistic simulations of the eukaryotic L1 stalk. The eukaryotic L1 stalk (or at least the one from yeast)[15] contains significant deletion compared with the bacterial structure.[13] Almost the entire H78 helix (including kink-turn Kt-78) is missing. The tetraloop normally capping the H78 is then in the immediate vicinity of the kink-turn Kt-77 canonical stem.[17] This would make the yeast L1 stalk[15] virtually identical to the above-noted Nikulin et al. 2003 structure[11] (Figure 2). Simulations based on the 2003 structure may provide useful insights into the behavior of eukaryotic L1 rRNA. Earlier biochemical experiments also suggested that L1 protein from bacteria recognizes the eukaryotic 28s rRNA L1 stalk and binds to it with a high rate.[18]

Non-Canonical Segments in L1 Stalk RNA. We focus our study on non-canonical segments of the L1 stalk RNA,[13,14] namely, the UCCG tetraloop (nts 2143−2148), L2111/2119 (nts 2111−2119), kink-turn Kt-77 (nts 2124−2128 and 2160−2174) and kink-turn Kt-78 (nts 2128−2136 and 2155−2160). Kt-78 along with the entire H78 is missing in the 2003 structure.[11] Kink-turns are recurrent RNA 3D motifs.[17,19] They facilitate RNA bending, are involved in protein-assisted RNA folding,[17] and may act as a flexible anisotropic molecular hinge.[20−23] The tetraloops are the most common RNA hairpin loops. They cap the canonical RNA helices with four nucleotides (N_{L1}−N_{L4}) and facilitate the backbone inversion.[24−27] The most abundant families of tetraloops are YNMG (with UNCG subfamily) and GNRA (where Y stands for pyrimidine, N for any nucleotide, M for adenine or cytosine, and R for purine).[26] For the remaining non-canonical segment, we chose the name "L2111/2119". There has been some confusion in nomenclature of the two L1 stalk kink-turns; specifically Kt-77 is called Kt-77/78 in ref 19, while Kt-78 is not mentioned. In ref 28, Kt-78 is called Kt-77/78 and Kt-77 is not mentioned. Cf. Figure 1 for nomenclature used in this study. A short description of kink-turns and tetraloops that is essential for understanding of the present simulations can be found in the Supporting Information.

The Tetraloop. The UCCG tetraloop is capping the end of helix H78. There is a major difference in the UCCG TL between the 2WRJ and 3U4M structures. Specifically, the base of nucleotide G2147 (G_{L4}) is in *anti* and *syn* conformation, respectively. In the canonical UNCG tetraloop,[24] G_{L4} is always in *syn*[29−32] which would indicate that the 2WRJ structure is incorrectly refined. However, the situation is complicated by the fact that the L1 tetraloop has specific contacts with the L2111/2119 (see below) which is uncommon for the UNCG tetraloops. In the 2WRJ structure, only two UNCG signature features are present, namely, the U_{L1}(O2′)/G_{L4}(O6) and 7BPh C_{L3}(N4)/U_{L2}(*pro*-R_p) H-bonds. These interactions are nevertheless the most essential for the UNCG tetraloop stability.[32−35] In the 3U4M structure, the G_{L4} nucleotide is in the native *syn* conformation and all signature features of the UNCG tetraloop are present including the tWS G_{L4}/U_{L1} base

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

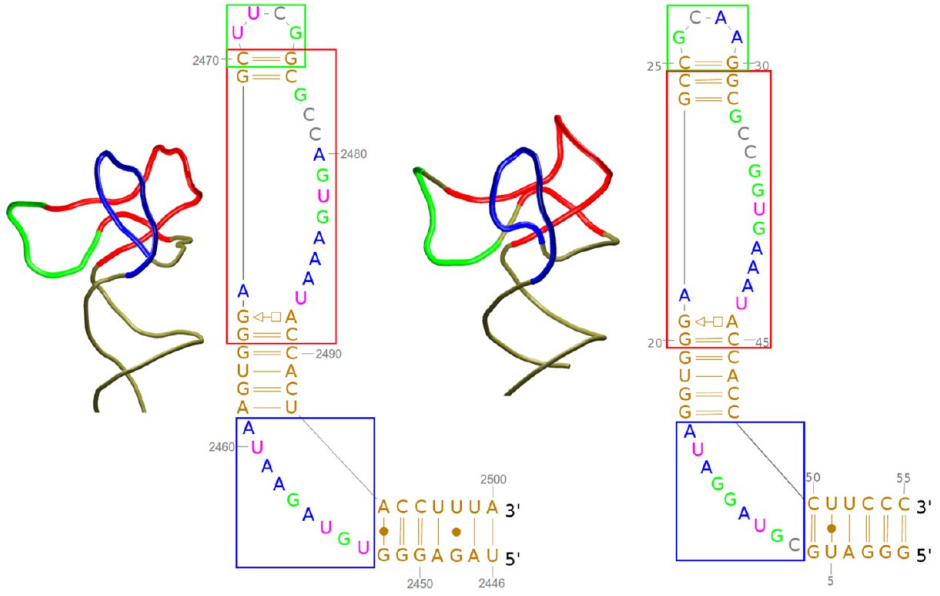


Figure 2. (left) Model of putative tertiary and secondary structure of the yeast L1 stalk based on its 3D 4 Å resolution structure.[15] (right) Tertiary and secondary structure of the Nikulin et al. structure.[11] Non-canonical motifs are highlighted. Blue is L2111/2119, red is kink-turn Kt-77, and green is the tetraloop. We used the original experimental nucleotide numbering for both schemes.

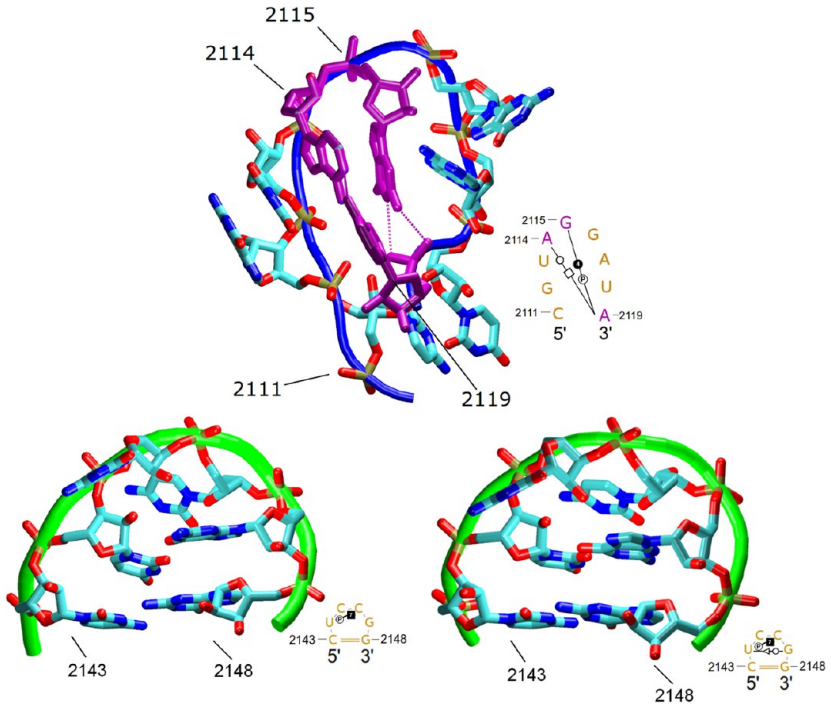


Figure 3. (top) L2111/2119 tertiary and secondary structure. *t*WH A2114/A2119 base pair and G2115/A2119 4BPh interactions are in purple. The backbone is in blue. (bottom) Tertiary and secondary structure of the UCCG tetraloop of H78. The 2WRJ structure (left) has G2147 base in *anti* conformation, while the 3U4M structure (right) is in *syn*. Standard nomenclature for base pairs and base-phosphate (BPh) interactions is used.[37,38]

pair (Figure 3). The L1 stalk UNCG tetraloop is not conserved in bacteria, as there is, e.g., a GCCA tetraloop in *E. coli* ribosome.[36] However, what appears to be conserved is the interaction between this tetraloop and L2111/2119 (see below) which is consistent with all the above-described tetraloop variants. We thus conducted a phylogenetic study of this tetraloop in all three phylogenetic domains and used molecular modeling to simulate the *E. coli* GCCA tetraloop.

The L2111/2119 Loop. The most distinct structural features of L2111/2119 are base stacks of nucleotides 2112−2115, *t*WH A2114/A2119 base pair,[37] G2115/A2119 4BPh interaction (base−phosphate interaction type 4),[38] and C2111/U2118 stacking (Figure 3). L2111/2119 interacts both with the UCCG tetraloop and the Kt-77 bulge (see below).

The Kink-Turns. Kink-turn (see Figure 4 and the Supporting Information) Kt-77 has an A-minor I interaction[39,40] and N(O2′)/A(N1)[41,42] essential interaction typical of other K-turns.[28] Significant deviation from consensus kink-turn is lack of involvement of A2126 in the *t*HS A/G base pair (Figure 4). The kink-turn further possesses an unusual 11-nucleotide-long

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

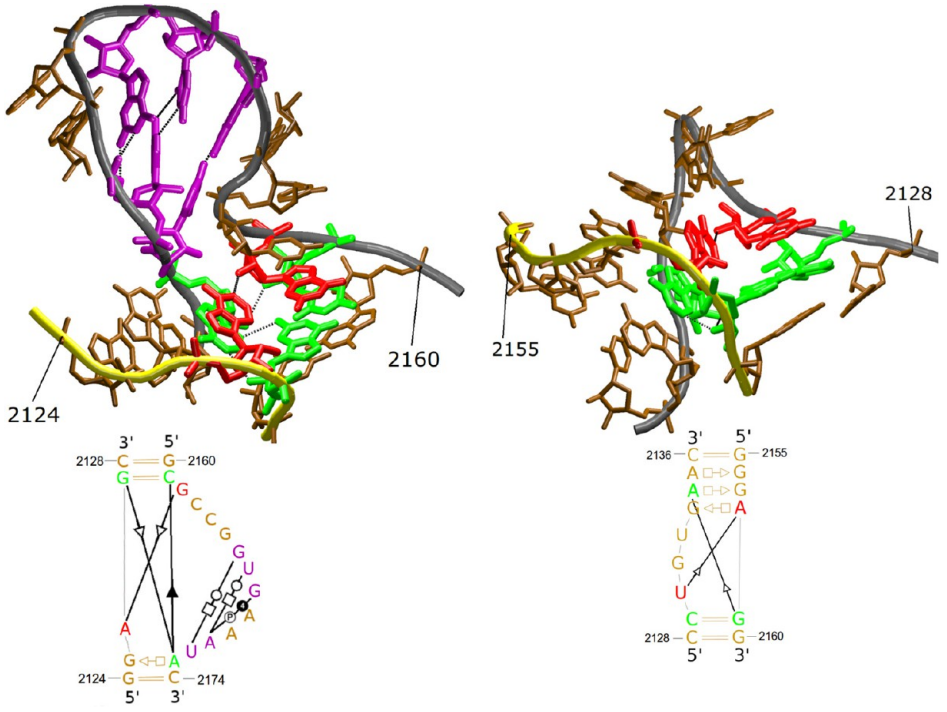


Figure 4. (left) Kt-77 tertiary and secondary structure. The shorter chain backbone path (nts 2124−2128) is in yellow, and the longer chain (nts 2160−2174) is in gray. A-minor I and essential interactions are in green and in red, respectively. Nucleotides involved in internal interactions of the bulge—*t*WH G2166/U2172, *t*WH U2167/A2171, and G2168/A2171 4BPh interactions—are in purple. The remaining nucleotides are in brown. (right) Kt-78; the longer chain (nts 2128−2136) is in gray, and the shorter chain (nts 2155−2160) is in yellow. A-minor 0 and the essential interaction are in green and red, respectively.

bulge (nts 2162−2172) instead of the usual bulge of three nucleotides. Kt-77, although somewhat differing from consensual kink-turns,[17] is highly conserved in all domains of life.[43] The kink-turn and its adjourning helices were previously identified as the most important site for rRNA−L1 protein recognition.[44] Bases of its signature kink-turn interactions (A-minor interaction and essential interaction) and three bases from its bulge (specifically, the ones that are involved in internal bulge interactions) show a high degree of conservation among bacteria. There is severe reduction in the binding affinity of L1 protein when nucleotides comprising signature interactions of Kt-77 were mutated.[44] The L1 protein also acts as a translational suppressor of its own mRNA operon.[45,46] The L11/L1 mRNA has at its 5′-end a sequential and structural homologue of Kt-77.[47,48] The Kt-77 bulge (nts 2162−2172) interacts with L2111/2119 and serves as a binding site of tRNA.[49] Its internal structure includes *t*WH G2166/U2172 and *t*WH U2167/A2171 base pairs complemented by G2168/A2171 4BPh and 2163−2168 stacking interaction.

The essentially consensus Kt-78 (see Figure 4 and the Supporting Information) has A-minor 0 interaction, essential interaction, and three-nucleotide bulge. It is located in the direct vicinity of Kt-77 (their A-minor interactions are separated by only one base pair). The polarity of the GC base pair (G in the shorter strand) seen in Kt-78 acting as an acceptor of the A-minor interaction[39] usually prefers A-minor I interaction[50] and is strongly conserved in bacteria. Thus, the weaker A-minor 0 Kt-78 interaction is likely invoked by the context (Figure 4). Kt-78 facilitates a sharp bend of helix H78, which brings the tetraloop into the vicinity of L2111/2119. Kt-78 structure was resolved only recently.[13,14]

Tertiary Interactions. There are extensive interactions between Kt-77 bulge, L2111/2119, and UCCG tetraloop (Figure 5). There are *c*HS U2118/G2148 and 6BPh C2145/C2111 interactions between L2111/2119 and UCCG tetraloop. There are other contacts between phosphates and sugar rings. In the 2WRJ structure, there is a *c*WW C2111/G2147 base pair which utilizes G2147 of UCCG tetraloop in its unusual *anti* conformation. In the 3U4M structure, the G2147 base is in *syn* conformation, resulting in the *c*WH C2111/G2147 base pair. There are *t*HH G2112/A2169, *t*HH U2113/A2170, *t*WS A2119/G2168, *c*SS U2167/A2114, *t*WS A2171/G2115, and *c*WW A2117/U2172 interactions between the Kt-77 bulge and L2111/2119. Contacts between the Kt-77 bulge and L2111/2119 can be described as "base pair layers", as they form three consecutive triads and two base pairs (Figure 6). The top of this structure serves as a binding site for E-site tRNA.

L1 Protein. In bacteria, the L1 protein is a multidomain protein 228 (or 229) amino acids long. It belongs in the ribosomal protein L1 family.[51] The first domain is larger and belongs to the class of $\alpha+\beta$ proteins. The second domain is smaller and sequentially interrupts (72−159 a.a.) the first domain. It belongs to the class of α/β proteins (Figure 7).[52] Both domains have a hydrophobic core with charged residues located on the surface. The protein forms many contacts with both rRNA and tRNA. Typical contacts involve interaction of a lysine or arginine side chain with the phosphate backbone; however, several specific major/minor groove interactions are also found. Since the bacterial ribosome is in the post-translocational state,[13] the L1 stalk is in direct contact with tRNA. There are contacts between L1 protein and tRNA and also between rRNA and tRNA (base stacking interactions between tRNA and rRNA bases G2112 and A2169) (Figure 7).

dx.doi.org/10.1021/jp401482m | *J. Phys. Chem. B* XXXX, XXX, XXX−XXX

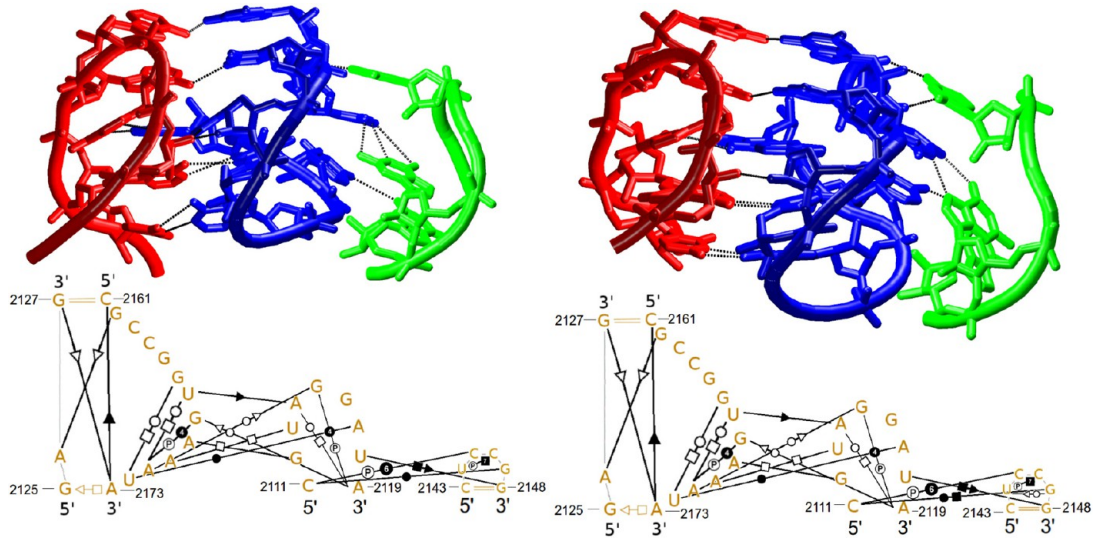


Figure 5. Secondary and tertiary structure of Kt-77 bulge, L2111/2119, and UCCG tetraloop (red, blue, and green, respectively); 2WRJ structure (left) and 3U4M structure (right).

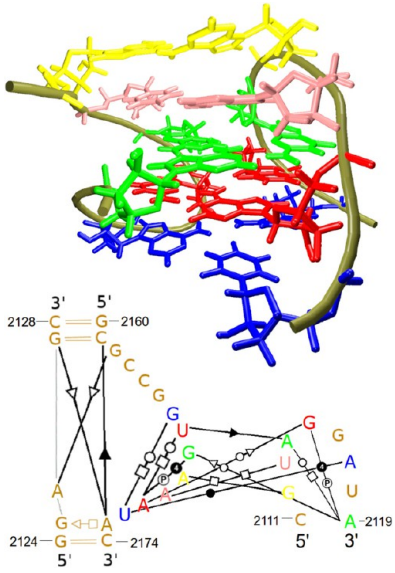


Figure 6. Interactions between Kt-77 bulge and L2111/2119. The first triad (blue) consists of nucleotides A2117, G2166, and U2172, the second triad (red) of G2115, U2167, and A2171, the third triad (green) of A2114, A2119, and G2168, the fourth base pair (pink) of U2113 and A2170, and the fifth base pair (yellow) of G2112 and A2169.

Although we have carried out many simulations including the L1 protein, in this study, we concentrate for space reasons on analysis of the RNA dynamics.

Prokaryotic vs Eukaryotic Structure. In the Nikulin et al. structure,[11] the helix H78 is missing (Figure 2). This leaves the capping GCAA tetraloop in the immediate vicinity of Kt-77 so it does not interact with L2111/2119. Otherwise, the structure is similar to the wild-type *T.t.* L1 stalk structure[13] (cf. Figures 1 and 2). We suggest that the Nikulin et al. structure is a suitable model for eukaryotic (yeast) L1 stalk based on the near-atomic resolution structure.[15] However, there are few differences between the structures. There is one missing base pair between Kt-77 and the tetraloop, and there is a different capping tetraloop (UUCG) in the yeast structure. In eukaryotic structure, there is likely also an unusual *c*WW A/G base pair at the top of helix H76. However, base pairing in the 4.0 Å eukaryotic structure of L1 stalk was merely estimated from the C1′/C1′ distances (Figure 2).

In this study, we use bioinformatics combined with extended MD simulations to analyze the conservation patterns in the individual building blocks of the L1 stalk RNA and their tertiary interactions with special emphasis given to the L1 stalk tetraloop. Further, we investigate structural differences in the available X-ray structures, discuss the possible roles of the L1 stalk RNA building blocks, and test the simulation methodology.

METHODS

We have carried out simulations for a number of systems (see below), ranging from simulations of isolated RNA motifs (such as kink-turns) through simulations of larger RNA segments up to protein RNA complexes.

The first set of simulations is based on the 3.6 Å resolution X-ray structure of *T.t.* ribosome[13] (PDB 2WRJ). 23S rRNA nucleotides 2093−2196 (original experimental numbering), L1 protein (chain C), and E-site tRNA (chain W in PDB 2WRK) were considered. In most simulations, H76 helix nucleotides without direct interaction with protein were deleted, using the 2104−2185 segment for simulations (Figure 1). Since they lacked any contact with L1 stalk, both anticodon and acceptor stem regions of tRNA were removed from the molecule in the simulations, leaving the 2−28 and 44−72 segments remaining. There is uncertainty (to our opinion) whether the first residue of L1 protein (methionine) is present. The methionine would be in the vicinity of Kt-78. Since coordinates of this residue are not available in the crystal structure, we used molecular modeling to obtain them for the purpose of three control simulations. Starting coordinates of methionine are obtained by simulated annealing in implicit solvent[53] with large positional restraint on the rest of the system.

For the second set of simulated systems, coordinates were taken from the 2003 2.65 Å resolution X-ray structure of the L1 protein/RNA complex (PDB: 1MZP).[11] Only coordinates of the rRNA and magnesium ions were used in the simulations.

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

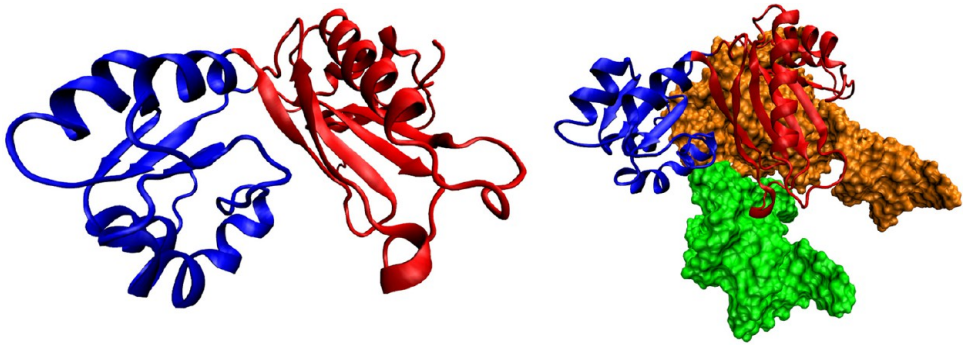


Figure 7. (left) Bacterial L1 protein.[52] Red and blue are domains I and II. (right) L1 interaction with tRNA fragment (green) and L1 RNA (orange).

The third set of simulated systems was obtained from the 2012 2.0 Å resolution structure of L1 protein/RNA complex (PDB: 3U4M).[14] All waters, monovalent cations, and anions, divalent cations, L1 stalk RNA, and L1 protein from this experimental structure were used in our simulations. This structure became available during the project.

The simulations were primarily carried out with the standard bsc0χ_{OL3}[54−57] RNA version of the Cornell et al. force field,[58] while ff99SB[59] was used for the protein. Initially, we also used older RNA force fields for some simulations (Supporting Information, Table S1). We used the TIP3P[60] and SPC/E[61] water models. We used AMBER-adapted Aqvist parameters for sodium (radius of 1.8680 Å and well depth of 0.00277 kcal/mol)[62] with the TIP3P water model and modified Dang's parameters for potassium (radius of 1.8687 Å and well depth of 0.100 kcal/mol)[63] with the SPC/E water model. Smith parameters (radius of 2.4700 Å and well depth of 0.100 kcal/mol)[64] for chloride were used in KCl excess salt simulations. We also used Mg^{2+} ions (radius of 0.7926 Å and well depth of 0.8947 kcal/mol)[62] in some simulations. In simulations of the 3U4M structure, we used protonated cytosine for nucleotide 2111.

For equilibration, we used 1000 steps of minimization followed by 100 ps of MD, with the system heated from 100 to 300 K. The atomic positions of the solute molecule were fixed by 25 kcal/mol Å^2 restraint. Then, five minimization and MD simulation series were carried out with restraints of 5, 4, 3, 2, and 1 kcal/mol Å^2 applied to all solute atoms. Then, 50 ps of restrained (0.5 kcal/mol Å^2) and unrestrained MD followed. Production simulation runs were performed with the pmemd module of AMBER 10.[65] The particle mesh Ewald (PME)[66,67] method was used for calculation of electrostatic interactions. Covalent bonds involving hydrogen atoms were constrained using the SHAKE[68] algorithm. Periodic boundary conditions, a 2 fs integration step, and 300 K temperature (Berendsen weak coupling)[69] were used. Analyses of resulting trajectories were performed using the ptraj module of AMBER, and the VMD program[70] was used for visualization. We used the S2S program[71] to generate maps of secondary structure.

The kink-turn interstem angle was measured as described earlier.[72] We defined the angle of Kt-77 as the angle between residue groups [2124, 2125, 2173, 2174]; [2126, 2162]; and [2127, 2128, 2160, 2161]. The angle of Kt-78 was defined as the angle between residue groups [2134, 2135, 2156, 2157]; [2130, 2133, 2158]; and [2128, 2129, 2159, 2160].

A bioinformatics study of L1 stalk sequences (especially the tetraloops) was done using Ribostral.[73] We used aligned sequences of the large ribosomal subunit 23S (28S) rRNA from the CRW Web site.[74]

RESULTS

Bioinformatics Analysis of the L1 Stalk Tetraloop. *Bacterial L1 Stalk Tetraloop Satisfies the GNNA and UNNG Patterns.* There is great variability in the sequence of the L1 stalk tetraloop in all phylogenetic families. Among bacteria, GCCA (50%) and GCUA (35%) tetraloops are favored, both deviating from the GNRA consensus[26] which requires the third base to be purine to form the $G_{L1}(O2')/R_{L3}(N7)$ interaction. The "UNCG-like" tetraloops (as in *T.t.*) are rather rare (7% of bacteria) and include in decreasing frequency UUUG, UGUG, UCUG, UCCG, and UGCG. Over 75% of these tetraloops do not follow the UNCG consensus which requires the third base to be a cytosine to form the $C_{L3}(N4)/N_{L2}(pro\text{-}R_p)$ interaction. Thus, the L1 stalk tetraloops in bacteria can be classified as either "GNNA" or "UNNG". We suggest that the UNNG tetraloops are preferred by thermophilic organisms.

The alignments indicate a large preference of pyrimidines (cytosine more than uracil) as a second and third nucleotide of the tetraloop. Preference for cytosine in the third position would be understandable for UNNG tetraloops to form the $C_{L3}(N4)/N_{L2}(pro\text{-}R_p)$ interaction. However, in the context of GNNA tetraloops, this particular preference makes the tetraloop inconsistent with the widespread GNRA class. Preference for cytosine in the second position is most likely caused by the tertiary C2145(N4)/C2111(pro-R_p) 6BPh interaction which is present in the L1 stalk and seems to be largely conserved in bacteria. We also observed near total tertiary interaction coupling (over 99%) between nucleotides 2111 and 2147 (fourth nucleotide from the tetraloop), where there is always either U2111/A2147 or C2111/G2147 combination. This observation does not discriminate which of the L1 stalk experimental structures is correct (2WRJ − G2147 *anti* or 3U4M − G2147 *syn*). Both *c*WW C/G → *c*WW U/A and *c*WH C/G → *c*WH U/A substitutions are isosteric, although the CG and UA covariation is perhaps more consistent with the WH pattern. The WW pattern would also include AU and GC combinations, although these could be eliminated by the internal tetraloop requirements.

L1 Tetraloop in Archaea and Eukaryotes. In archaea, the UNNG tetraloop has ∼55% occurrence with the UCCG variant being the most common. This likely reflects the fact that many archaea are extremophilic organisms and that the UNNG tetraloop provides superior thermodynamic stability over GNNA as witnessed in thermophilic bacteria preferring the UNNG. The other observations are the same as those for the bacterial sequences.

The situation in eukaryotes is different. Namely, the H78 helix is often missing from the structure (placing the tetraloop

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

Table 1. List of Simulations of the L1 Stalk Carried out with the bsc0χ_{OL3} Force Field

simulation name	simulated structure	ions[k]	water model	trajectory length (ns)
	Bacterial L1 Stalk (pdb: 2WRJ)			
RNA_Na	L1 RNA[a]	81 Na^+	TIP3P	100
RNA_KCl	L1 RNA[a]	136 K^+, 55 Cl^-	SPC/E	1000
PR_KCl	protein	76 K^+, 84 Cl^-	SPC/E	200
L1_KCl	L1 RNA[a] + protein	136 K^+, 63 Cl^-	SPC/E	180
L1m_KCl	L1 RNA[a] + protein[i]	136 K^+, 63 Cl^-	SPC/E	100
L1tRc_KCl	L1 RNA[b] + protein + tRNA[c]	193 K^+, 44 Cl^-	SPC/E	200
KT77_Na	Kt-77[e]	23 Na^+	TIP3P	100
KT77wl_Na	Kt-77[d]	26 Na^+	TIP3P	100
KT77_KCl	Kt-77[e]	35 K^+, 12 Cl^-	SPC/E	100
KT77wl_KCl	Kt-77[d]	38 K^+, 12 Cl^-	SPC/E	1000
KT78_Na	Kt-78[f]	21 Na^+	TIP3P	400
KT78_Mg	Kt-78[f]	10 Mg^{2+}, 1 Na^+	TIP3P	1000
KT78_KCl	Kt-78[f]	31 K^+, 12 Cl^-	SPC/E	1000
UNCG_KCl	UCCG tetraloop[h]	20 K^+, 7 Cl^-	SPC/E	1093
LoopA_KCl	loop 2111/2119[g]	17 K^+, 7 Cl^-	SPC/E	100
GCCA_anti_isol	GCCA tetraloop[h,j]	20 K^+, 7 Cl^-	SPC/E	300
GCCA_anti	L1 RNA[a,j]	136 K^+, 55 Cl^-	SPC/E	100
	Bacterial L1 Stalk (pdb: 3U4M)			
3U4M_RNA	L1 RNA[a]	135 K^+, 57 Cl^-	SPC/E	730
3U4M_complete	L1 RNA[a] + protein	5 Mg^{2+}, 124 K^+, 62 Cl^-	SPC/E	200
3U4M_UNCG	UCCG tetraloop[h]	20 K^+, 7 Cl^-	SPC/E	300
GCCA_syn_isol	GCCA tetraloop[h,j]	20 K^+, 7 Cl^-	SPC/E	300
GCCA_syn	L1 RNA[a,j]	135 K^+, 57 Cl^-	SPC/E	100
	L1 Stalk from Nikulin et al. (pdb: 1MZP)			
euk_KMgCl	L1 RNA	3 Mg^{2+}, 81 K^+, 32 Cl^-	SPC/E	1000
euk_KCl	L1 RNA	81 K^+, 28 Cl^-	SPC/E	100
euk_KCl_2	L1 RNA	81 K^+, 28 Cl^-	SPC/E	300

[a]Nucleotides 2104−2185 (2WRJ) and 2105−2184 (3U4M). [b]Nucleotides 2093−2196. [c]Nucleotides 2−28 and 44−72 of chain W. [d]Nucleotides 2121−2129 and 2159−2177, two additional GC pairs were modeled into the structure beyond nts 2129 and 2159 to allow stable simulations. [e]The long bulge of Kt-77 was reduced to a three-nucleotide bulge (using model building) to obtain a structure more resembling a consensual kink-turn, leaving nucleotides 2121−2129, 2159−2163, and 2172−2177. Two additional GC pairs were modeled into the structure beyond nts 2129 and 2159 to allow stable simulations. [f]Nucleotides 2128−2138 and 2153-2160. [g]Nucleotides 2111−2119. [h]Nucleotides 2139−2152. [i]Methionine was added by modeling (see Methods). [j]The UCCG tetraloop (nts 2144−2147) was replaced with GCCA. Base of nucleotide C2111 was replaced with uracil. [k]The K^+ concentration in KCl excess salt simulations was 0.30−0.35 M.

right after the Kt-77 canonical stem). Additionally, some eukaryotes (for example, *H. sapiens*) seem to contain a different structure in place of the H78 and the tetraloop could not be identified by the alignments. Nevertheless, for the organisms where the tetraloop could be found based on alignment with the bacterial sequence, there is a preference for tetraloops belonging to the classical GNRA or YNMG families. No distinct preference for pyrimidines in the second or third nucleotide was observed. There is also no clear correlation between nucleotides 2111 and 2147, contrasting the bacterial rules.

Thus, bacteria and archaea have a strong evolutionary preference for the conservation of contacts between the tetraloop and L2111/2119 while sacrificing the tetraloop intrinsic thermodynamic stability caused by deviations from the UNCG and GNRA tetraloop families. Since there is most likely no contact between L2111/2119 and the tetraloop in eukaryotes, the intrinsically stable tetraloops from UNCG and GNRA families are preferred.

Complete bioinformatics data are in the Supporting Information, Tables S2 and S3.

Bioinformatics Analysis of the L1 Stalk Internal Loops. Phylogenetic study of the Kt-77 bulge and L2111/2119 regions was done in the past.[18,43,75] Using structural data, we add several points to these studies. Namely, (i) all nucleotides participating in Kt-77 bulge/L2111/2119 interactions are highly conserved (over 95%) in all domains of life, and if there is a substitution, it is isosteric or near-isosteric; (ii) both Kt-77 and Kt-78 essential interactions are conserved in all domains (except for Kt-78 in eukaryotes); and (iii) H78 is the least conserved element in the L1 stalk, randomly using canonical base pairs and GU wobbles in bacteria and archaea and being absent in eukaryotes. Note, however, that while the actual sequence of helix H78 is not conserved, its length (in terms of number of nucleotide base pairs) is highly conserved, conserving the positioning of the attached elements and the flexibility.[76]

Conservation of pairing between L2111/2119 and the Kt-77 bulge is necessary because these segments form the complex platform that serves as a binding site for the tRNA. Both Kt-77 and Kt-78 are also conserved; however, while Kt-77 is evidently recognized by L1 protein with high fidelity, we think the main function of Kt-78 is to change the direction of helix H78 in order to bring the capping tetraloop into the vicinity of L2111/2119. Similarly, the low conservation of the H78 sequence (but high conservation of its length) indicates that it is not specifically recognized by any element and its primary function is to bring the tetraloop into the right place. This can be achieved with various A-RNA sequences. Recently, it was also suggested that the minor groove of H78 might be interacting with the C-domain of protein L9 during the proteosynthesis;[14] however, as we note, the helix H78 sequence is poorly conserved.

Molecular Simulation Analysis of the L1 Stalk Tetraloop. *UCCG Tetraloop with G anti.* In simulation of an isolated UCCG tetraloop having G in *anti*, we see quick loss of the UNCG tetraloop essential interactions. Nucleotide G2147 has the tendency to bulge out of the tetraloop and then fluctuates between *anti* and *syn* conformations. It is possible that longer simulations of the isolated structure would lead to reintegration of G into the structure and formation of the native tetraloop with the G2147 base in *syn*, but it did not occur in our 1.1 μs simulation. Our results are consistent with recent extensive REMD simulation of folding of the UUCG tetraloop.[77] In simulations considering the tetraloop in a broader context (contextual simulations), the structure maintains essential interactions of the tetraloop, with the exception of simulation L1_KCl (see Table 1 for the list of simulations) where the tetraloop disintegrates and simulation RNA_Na where the $C_{L3}(N4)/C_{L2}(pro\text{-}R_p)$ hydrogen bond is lost.

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

UCCG Tetraloop with G syn. All essential interactions of the UCCG tetraloop were maintained for the entire simulation of the isolated structure. One additional $C_{L2}(O2')/G_{L4}(N7)$ interaction which is also occasionally mentioned as the UNCG signature interaction was sometimes temporarily formed in the isolated structure simulation. The tetraloop was completely stable in contextual simulations.

Interactions between Kt-77 Bulge, L2111/2119, and UCCG Tetraloop—the 2WRJ Structure. In our simulations, interactions between loops are mostly very well maintained. However, we universally observe loss of the 6BPh C2145/C2111 interaction. This interaction either changes to a C2145(N4)/C2111(O2′) hydrogen bond or disappears completely. This either indicates a problem in the force field description or reflects a lower resolution of the starting structure. However, the observed behavior can also be related to the *c*WW C2111/G2147 base pairing imposed by the *anti* conformation of G_{L4} G2147 (see above). This canonical base pair is quite perturbed in the X-ray structure, having a C1′/C1′ distance of 8.5 Å. In simulations, this *c*WW base pair stabilizes and the C1′/C1′ distance increases to the expected 10.6 Å value. It simultaneously "nudges" the tetraloop backbone slightly further away from the L2111/2119 to create space for the *c*WW base pair and most likely disrupts the 6BPh C2145/C2111 interaction.

Interactions between Kt-77 Bulge, L2111/2119, and UCCG Tetraloop—the 3U4M Structure. We did not observe loss of any interactions between the loops in simulations of the 3U4M structure. Even the 6BPh C2145/C2111 interaction was fully maintained in all simulations. This indicates that the 6BPh interaction is only compatible with G2147 of the UCCG tetraloop in standard *syn* conformation. The 6BPh C2145/C2111 interaction is further bolstered by sugar–base C2111(O2′)/C2145(O2) interaction and by stacking of both bases (Figure 8). There is a delicate balance of these three interactions, and any significant change in the vicinity (like *syn* → *anti* base swap of the G2147) leads to structural perturbations seen in simulations. Since the G2147 base is in *syn* conformation, the C2111/G2147 base pair adopts a *c*WH arrangement with N3-protonated cytosine. This base pair is buckled in the crystal structure as well as in the simulations. The buckling reduces the C1′/C1′ distance in the base pair which may be important for the overall balance between the 6BPh C2145/C2111 interaction and the C2111/G2147 base pair.

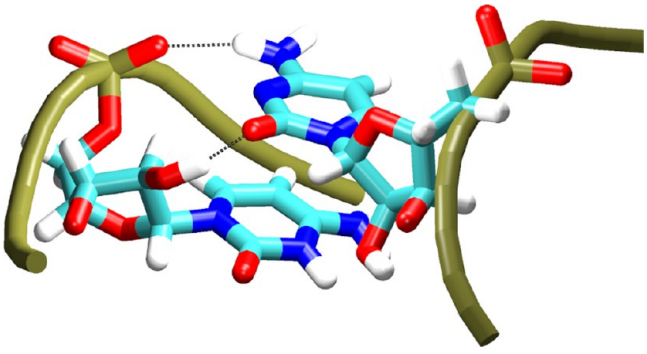


Figure 8. The 6BPh C2145/C2111 interaction. C2111(O2′)/C2145(O2) and C2145(N4)/C2111(*pro*-Rp) interactions are indicated. The RNA backbone is brown.

Is G2147 syn or anti? Even after all investigations, we are unable to make a final conclusion about the *syn* and *anti* orientation of the G2147 in the UCCG tetraloop, though we tend to suggest that the G2147 should rather be *syn*. *Syn–anti* mis-refinements are common in X-ray structures of ribosome.[78] The *syn* variant is supported by the higher resolution structure; it is compatible with the native UNCG tetraloop structure, and the simulations with *anti* G2147 lead to loss of important nearby BPh interaction. On the other hand, both G2147 conformations are compatible with the C2111/G2147 tertiary interaction, structurally as well as based on sequence analysis, albeit the covariation pattern would more straightforwardly support G2147 to be *syn* (see the Bioinformatics part). Further, the tertiary base pair interaction with *anti* G2147 is deformed in the X-ray structure compared to ideal *c*WW base pairing. On the other side, *syn* G2147 requires protonation of C2111 and also the visual inspection of the electron density of the 2WRJ structure is quite consistent with the suggested *anti* conformation. Further, the UNCG tetraloop is evolutionarily dispensable (see above). Thus, we hypothesize that both states are viable, may coexist, and may be pH dependent, due to the need to protonate C2111 in the *syn* variant. The pH difference between the crystal structures (6.5 for 2WRJ, and 5.6 for 3U4M) supports this view.

Apart from the differences related to the *anti*/*syn* conformation of G2147 in the 2WRJ and 3U4M structures, respectively, the systems behave in a similar fashion. We always observe formation of an additional *c*HW G2116/G2165 base pair between the Kt-77 bulge and L2111/2119. This pair is located at the base of the structure formed by the Kt-77 bulge and L2111/2119 (Figure 6 and Table 2). It is formed from the otherwise unpaired nucleotides of 2116 and 2165 and can be considered the zeroth base pair layer of the structure. It sometimes breaks down but usually reforms very quickly again.

Table 2. Layers of Interactions between Kt-77 Bulge and L2111/2119

level	nucleotides	interactions
zeroth[a]	G2116 G2165	*c*HW G2116/G2165
first	A2117 G2166 U2172	*c*WW A2117/U2172 *t*WH G2166/U2172
second	G2115 U2167 A2171	*t*WS A2171/G2115 *t*WH U2167/A2171
third	A2114 A2119 G2168	cSS U2167/A2114 *t*WS A2119/G2168 *t*WH A2114/A2119
fourth	U2113 A2170	*t*HH U2113/A2170
fifth	G2112 A2169	*t*HH G2112/A2169

[a]Usually formed in simulations but not present in the X-ray structure.

We also observed an interaction between the Kt-77 bulge and L2111/2119 that is maintained in simulations only in the presence of a bound monovalent cation and during short intervals of ion exchange can become perturbed. This ion site is near atoms U2113(O4) and A2170(N1). An ion located at this site stabilizes the *t*HH U2113/A2170 base pair. Note that this base pair is slightly perturbed from the typical *t*HH base pair family in all three experimental structures, which probably explains the necessity for additional stabilization by a cation (Figure S1 and Table S4, Supporting Information). The Nikulin et al. structure shows a magnesium ion in this position,

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX–XXX

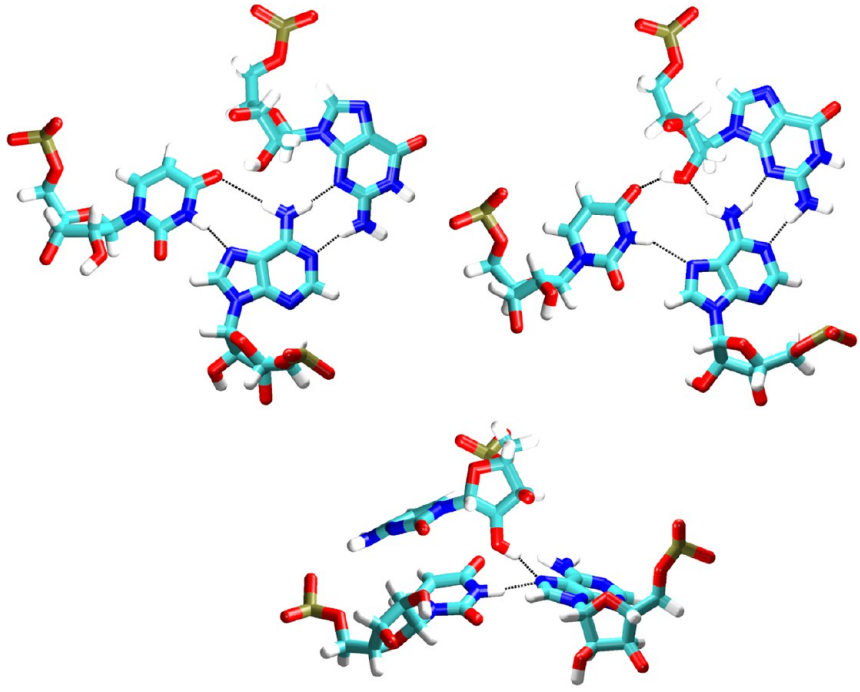


Figure 9. Perturbations of the U2111/A2147 base pair observed in the GCCA(*syn*) simulations. (top) *c*WH U2111/A2147 and *t*WS A2147/G2144 base pairs; the A2147(N6)/U2111(O4) interaction can be either direct (left) or mediated (right) by the G2144(O2′) hydroxyl group; (bottom) bifurcated U2111(N3)/A2147(N7) and C2145(O2′)/A2147(N7) interactions that are deforming the *c*WH U2111/A2147 base pair.

while the structures of the *T.t.* ribosomal L1 stalk do not contain it, which may be due to disorder and resolution limit. In our simulations, monovalent Na^+ or K^+ ions seem to be able to substitute magnesium at this site. It confirms earlier results for kissing-loop complexes, loop E and HDV ribozyme that RNA simulations can provide rather good assessment of binding of monovalent ions. Monovalent ions in simulations can readily occupy Mg^{2+} binding sites seen in the experiments.[79−83]

GCCA Simulations. Since our bioinformatics study revealed that the most common tetraloop in the L1 stalk is GCCA, we transformed the UCCG tetraloop into the GCCA one and carried out a set of simulations. We tested the A2147 base being in either *anti* of *syn* conformation (modeling based on the 2WRJ and 3U4M structures, respectively). Concomitantly, the C2111 base was transformed into uracil, to follow the 2111/2147 covariation pattern (see above). We did simulations of the isolated GCCA tetraloop to see its intrinsic stability and contextual simulations of the whole RNA structure to see differences in its interaction with the rest of the structure. Since the GCCA tetraloop was to our knowledge not previously described in the literature, we also discuss its putative signature interactions.

In the system with A2147 in *syn*, there are *t*WS A_{L4}/G_{L1}, $C_{L3}(N4)/C_{L2}(pro\text{-}R_p)$ 6BPh, $G_{L1}(O2')/C_{L2}(O5')$, and $C_{L2}(O2')/A_{L4}(N7)$ hydrogen bonds. The *t*WS pair is equivalent to the *t*WS G/U pair which is a signature interaction for the UNCG tetraloops. Also, the C_{L3}/C_{L2} 6BPh interaction is found in the UNCG tetraloops, albeit it is somewhat fluctuating here. The $G_{L1}(O2')/C_{L2}(O5')$ interaction is formed due to lack of O6 atom on the A2147 base (normally, in the UNCG tetraloop, the U_{L1} hydroxyl group would form a signature interaction with the O6 atom of G_{L4}). $C_{L2}(O2')/A_{L4}(N7)$ is another interaction that is sometimes mentioned as a signature interaction of the UNCG tetraloops. Curiously, while we see this interaction in simulations of the GCCA(*syn*) tetraloop, it was not found in the L1 stalk UCCG tetraloop structure or simulations of larger systems with UCCG. It is a minor substate in the simulation of the isolated UCCG tetraloop. In conclusion, we predict three UNCG-like interactions in the GCCA(*syn*) tetraloop that are stable in the simulations. This indicates that the GCCA tetraloop is capable of adopting a topology similar to the UNCG tetraloop.

In the contextual simulations of the entire L1 stalk RNA, the *c*WH U2111/A2147 base pair is mostly stable, although not as much as the *c*WH C2111/G2147 base pair in the UCCG system. There are frequent insertions of nearby functional groups between the two bases, distorting the pair in various ways (Figure 9). Namely, there are bifurcated U2111(N3)/A2147(N7) and C2145(O2′)/A2147(N7) interactions. The latter interaction was permanent in the simulation of isolated structure and was competing with the U2111 base for the contact with A2147. It was effectively increasing the buckle of the U2111/A2147 base pair. Additionally, the hydroxyl group of G2144 was repeatedly being inserted between A2147(N6) and U2111(O4) atoms, causing the A2147(N6)/U2111(O4) interaction to be either direct or mediated by the G2144(O2′) group (Figure 9).

The GCCA system with A2147 *anti* shows similar instabilities as the UCCG(*anti*) tetraloop. The *t*HS A_{L4}/G_{L1} base pair is noticeably less stable than the *t*WS A_{L4}/G_{L1} pair in the GCCA(*syn*) systems. Namely, the A2147 has larger fluctuations, perturbing the base pair, especially the $A_{L4}(N6)/G_{L1}(N3)$ hydrogen bond, which is often lost due to base pair opening. Other interactions are $C_{L3}(N4)/C_{L2}(pro\text{-}R_p)$ 6BPh and $G_{L1}(O2')/C_{L2}(O5')$ hydrogen bond. There is no interaction between the hydroxyl group of C_{L2} and the base of A_{L4} (as was the $C_{L2}(O2')/A_{L4}(N7)$ in GCCA(*syn*)), which might explain the increased fluctuations of the A_{L4}. Additionally, there is no contact between the N2 group of G_{L1} and the phosphate group downstream of the A_{L4} nucleotide which is usually observed in

dx.doi.org/10.1021/jp401482m | *J. Phys. Chem. B* XXXX, XXX, XXX−XXX

traditional GNRA tetraloops, further contributing to the increased A_{L4} fluctuations.

In contextual simulations, the interactions between the tetraloop and L2111/2119 were lost early in the simulation but re-established after 57 ns. However, similarly to the simulations with the UCCG(*anti*) tetraloop, the 6BPh C2145/C2111 interaction was permanently lost and replaced by the C2145(N4)/C2111(O2′) interaction. The *c*WW U2111/A2147 base pair remained stable after 57 ns. There were no bifurcations or insertions of other groups between the base pair similar to those that were seen in GCCA(*syn*) simulations.

In both GCCA(*syn*) and GCCA(*anti*) systems, the *c*HS U2118/G2148 interaction, i.e., the contact between the L2111/2119 and the tetraloop enclosing CG pair, was fully maintained once stable pairing between U2111 and A2147 was established.

In conclusion, simulations with the modeled GCCA tetraloop paint a similar picture as the simulations with UCCG. More stable trajectories are achieved with the fourth nucleotide base being in *syn* conformation. This conformation provides stability to the tetraloop both in isolated simulations (where it allows additional stabilizing hydrogen bond interaction to form) and in contextual simulations (where it is better maintaining interactions between L2111/2119 and the tetraloop). This makes the GCCA tetraloop more similar to the YNMG than to the GNRA family. When still considering the *anti* possibility, the simulations suggest the GCCA(*anti*) would be more stable than UCCG(*anti*).

Simulations of the Internal Loops. *L2111/2119.* Simulation of the isolated L2111/2119 loop leads to immediate loss of the internal base pairs (Figure 3). No new stable base pairing is established, and the structure forms a base stack of the neighboring unpaired bases. Thus, the L2111/2119 is structured only as part of a larger context (Figure 1) when it is sandwiched between the Kt-77 bulge and UCCG tetraloop in the bacterial ribosome (Figure 5). The loop is entirely stable in all other simulations (full RNA, RNA + protein, and RNA + protein + tRNA).

Kt-77 Bulge. There are several simulations of kink-turns in the literature[21−23,84−86] but none of them with a bulge similar to Kt-77. The preceding simulations revealed hinge-like (elbow-like) fluctuations of the kink-turns. In simulations of isolated Kt-77, the bulge somewhat attenuates the hinge-like fluctuations, especially large-scale openings as the long bulge represents a structural hindrance for such movements. The bulge changes conformation due to loss of its interactions with L2111/2119 (Figure 5). Further, the bulge comes into contact with the Kt-77 stems during the elbow-like motions (Figure 4). The changes in the structure of the bulge in simulations of isolated Kt-77 do not seem to have a significant effect on the structure or dynamics of the main part of the kink-turn. The mean simulation interstem angle (see Methods) is close to the experiment (104−105° in all three X-ray structures) in all simulations (Supporting Information, Table S5). The Supporting Information describes simulations of system where we replaced the long bulge by the common three-nucleotide bulge.

In contextual simulations (simulations including also some other elements), the bulge is held in place by L2111/2119 and its structure is fully maintained. All data support the view that the long bulge is an additional independent RNA element inserted into the kink-turn and extruding from it to establish specific tertiary interactions. In contrast to the kink-turn, the bulge does not fold autonomously and is adapted to a specific environment in the L1 stalk region.

Kt-77 Essential and A-Minor Interactions. The kink-turn A-minor I interaction and essential interaction were mostly well maintained. In simulation RNA_Na, the A-minor I interaction falls apart at 3.5 ns, but the structure is restored at 69 ns and then stabilized. The essential interaction of the kink-turn is permanently disrupted immediately after the start in simulation L1tRc_KCl. Both problems appeared at the simulation start, and could be due to the lower resolution of the 2WRJ structure. There were no such problems with the 3U4M structure simulations.

The kink-turn essential interaction G2162(O2′)/A2126(N1) is stable in other simulations of Kt-77. There is no tendency for it to be disrupted or to become water-mediated. It contrasts simulation data for other kink-turns as local instability of the essential interaction of kink-turn is common (see the Discussion below).[72]

Another observation specific to the Kt-77 simulations is fluctuation of the G2127(O2′)/A2173(N1) interaction of A-minor which is periodically falling apart and reforming. Population of the disrupted form is ∼25%. It does not affect other parts of the system. Simulations of other kink-turns show instead fluctuations of O2′/O2′ (C2161(O2′)/A2173(O2′) in Kt-77) interaction of the A-minor between direct and water-mediated state. This contributes to kink-turn hinge-like dynamics.[21−23] We also observe such dynamics in our Kt-77 simulations; however, for the most part, the fluctuations mainly involve the G2127(O2′)/A2173(N1) interaction. The inserted Kt-77 bulge subtly modifies fluctuations of the longer strand of the kink-turn, which includes the O2′/O2′ interaction. The hinge-like properties of Kt-77 are more visible for the shorter strand and the O2′/N1 interaction. In summary, the long bulge seems to modestly modulate the fluctuations of the Kt-77 compared to other kink-turns simulated before.[21−23,50,72,84−87]

Kt-78 Simulations. The simulation behavior of Kt-78 is similar as described in the literature for other kink-turns[21−23] and is summarized in the Supporting Information. In some simulations of Kt-78, we observed formation of a hitherto unknown alternative Kt-78 structure (Figure 10) further discussed in the Supporting Information.

Force Field Performance. *Local Instability of the Essential Interaction in Kt-78 Is Not Corrected by the χ_{OL3} Modification.* A common feature of all simulations (including simulations in larger contexts) of Kt-78 is local rearrangement of the U2130(O2′)/A2158(N1) essential interaction. It tends to be disrupted at the very beginning, and then, it is replaced by A2158(N6)/U2130(O2′) interaction (Supporting Information, Figure S2). There is usually a period of fluctuations (tens of ns) before the shift becomes permanent. Such behavior has been reported in earlier MD studies of kink-turns;[50] however, the rearranged geometry is not known from experimental structures. Although it is just a local perturbation of the structure which does not affect its overall arrangement (so the kink-turn simulations can be still considered as valid), we consider it as a sign of force field imperfection. It reflects a struggle between the interactions and backbone conformation in the kink region. Although it does not seem to have an immediate impact on kink-turn stability, it could potentially lead to structure disruption or major rearrangement of the structure in longer simulations. As mentioned above, this structural change is absent in Kt-77 simulations, where the essential interaction is perfectly stable. It is likely that the long bulge of Kt-77 allows a more subtle adaptation of the bulge. It may reduce the force field imbalance between the overall

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

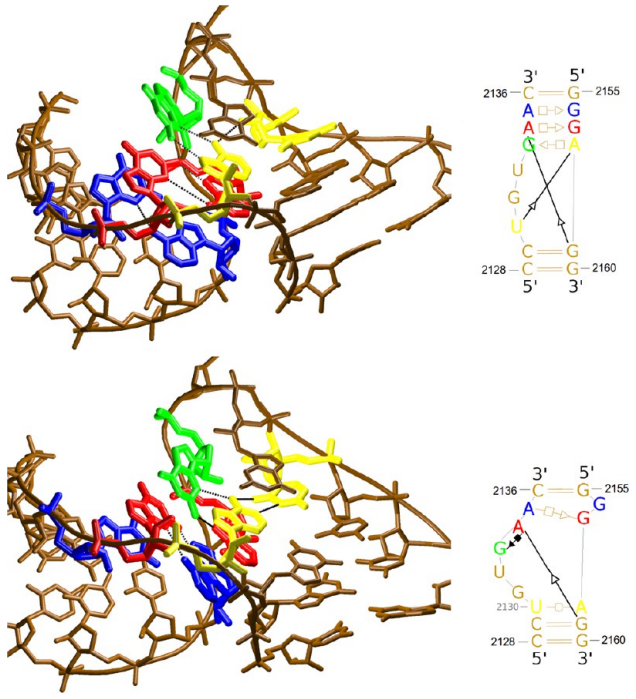


Figure 10. Alternative structure of Kt-78 seen in some simulations. The third *t*HS A/G pair is blue, the second *t*HS A/G pair is red, the kink-turn essential interaction is yellow, and guanine from the first *t*HS A/G pair is green. Black dotted lines indicate pairing and hydrogen bonds.

kink-turn topology and the essential interaction which is responsible for the loss of the essential interaction in kink-turn simulations. An important result of our study is that the problem is not eliminated with the bsc0χ_{OL3} force field which has been for the first time used for kink-turns here. We tried to identify an eventual force field term which could be responsible for the rearrangement, but no conclusive suggestions could be made.

Older Force Field Version Is Inferior to bsc0χ_{OL3}. Although the bsc0χ_{OL3} force field refinement does not improve the simulation behavior for the kink-turn essential interaction, improvement is seen for some other structural features of the L1 stalk RNA, as summarized in the Supporting Information. For more systematic analysis of RNA force field performance, see refs 55, 87, and 88.

■ ADDITIONAL RESULTS

Dynamics of the L1 Protein and the Mechanism of tRNA Release. The secondary structure of L1 protein was largely stable in simulations, and its dynamics were expressed mainly by its interdomain movements.[49] In simulations without tRNA, we observed significant movement of protein domains in relation to each other. This movement occurred through opening of the domains and also through rotation of the domains in relation to each other. It has been suggested that similar movements of the protein are required for the release of bound tRNA.[49] In comparison with the experimental structure, we observed both closing and opening of the domains. However, we observed no opening or rotation of L1 protein domains in simulations with tRNA. If this opening is indeed associated with tRNA release, such a process would be far beyond our simulation time scales. We also observed interactions between domain II of the protein and Kt-77 bulge. These interactions are lost when the L1 protein opens. This also leads to increased instability in bases G2112 and A2169 (bases with direct contact to tRNA). Therefore, opening of L1 protein which in itself destabilizes interaction between L1 protein and tRNA could also destabilize interaction between L1 rRNA and tRNA. There might be a system of positive feedback loops related to tRNA release where destabilization of one of the contacts between tRNA and L1 stalk leads subsequently to or is coupled with weakening of the remaining interactions (Figure 11). It could similarly work in the opposite direction for the tRNA binding.

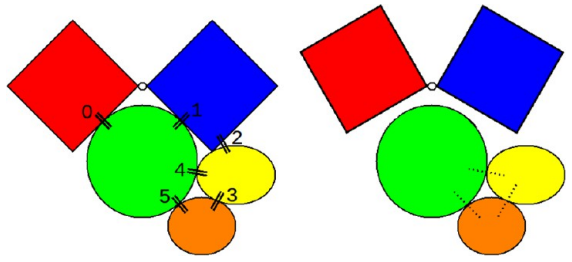


Figure 11. (left) Interactions between L1 protein domain I (red) and II (blue), tRNA (green), Kt-77 bulge (yellow), and L2111/2119 (orange). (right) Opening of the L1 protein domains disrupts the protein/tRNA (0, 1) and protein/Kt-77 bulge interactions (2). The loss of the latter interaction (2) also leads to increased fluctuations of Kt-77 bulge/L2111/2119 *t*HH G2112/A2169 base pair (3) which also interacts with tRNA (4, 5) and could ultimately weaken this contact between tRNA and rRNA. Stable contacts are indicated by double lines. Weakened contacts are indicated by dotted lines.

Simulations with tRNA Fragment. The tRNA fragment (see Methods) was stable in simulations. The interactions between tRNA and L1 stalk were largely unchanged from the crystal structure state. Our simulations are too short to address processes associated with tRNA release. Nevertheless, we made one observation regarding the *c*WW U2109/U2180 base pair in helix 76. This base pair is strongly evolutionarily conserved in bacteria and archaea (but not in eukaryotes), suggesting it may be recognized by some ligands. In simulations, it was generally less stable than the rest of the base pairs in the helices. In L1tRc_bsc0_KCl simulation, we observed its reversible loss associated with formation of an additional interaction between the L1 stalk and tRNA. Uracil 2180 bulged out of the helix to form either a base pair or stacking interaction with U18 from the tRNA D-loop (Figure 12). We also observed instability of the *c*WW U2109/U2180 pair in the newer experimental structure of *T.t.* ribosome,[89] albeit without forming an interaction with tRNA. This suggests intrinsically reduced stability of this base pair. However, since we observed its interaction with tRNA in only one simulation, we refrain from making any specific suggestions regarding its role.

Mg^{2+} Ions. Some of our simulations included Mg^{2+} cations. Their behavior is described in the Supporting Information.

Simulations Based on the Nikulin et al. Structure. *Basic Behavior.* The main feature of simulations of the Nikulin et al. 2003 structure missing the H78 was lower stability (as judged by structural fluctuations seen in the simulations) of L2111/2119 compared to simulations of the bacterial structures. Although its internal interactions did not disintegrate permanently, there were larger fluctuations in comparison with the bacterial structures. These fluctuations also permeated into the Kt-77 bulge where the base pairs became partially

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

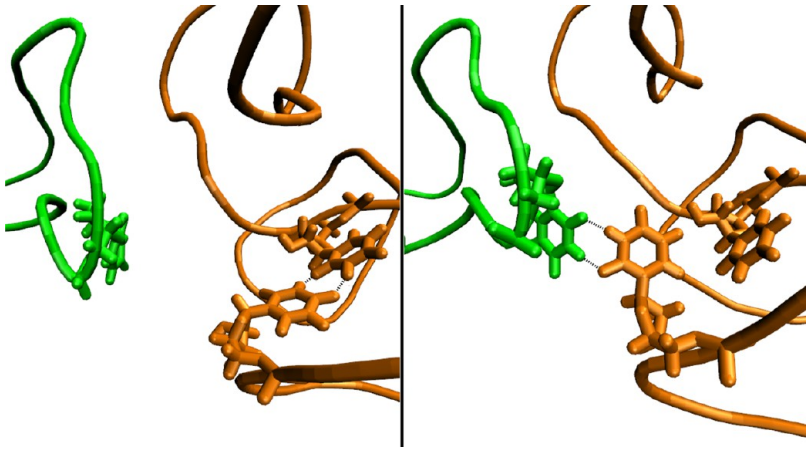


Figure 12. Disintegration of the rRNA (orange) *c*WW U2109/U2180 base pair and subsequent interaction of U2180 with U18 from tRNA (green). Hydrogen bond interactions are indicated by black lines.

perturbed. There was particular increase in fluctuations in nucleotides C7 and U14 (which are located at the base of L2111/2119). Note that, in full bacterial structure, these nucleotides are interacting with the tetraloop extensively, but, in the Nikulin et al. structure, they are unpaired, being only engaged in mutual stacking interaction. In simulations, the base stack itself was mostly maintained, but both bases fluctuated. There were also large fluctuations of the sugar−phosphate backbone around the C7 nucleotide, which in full bacterial structure normally interacts with the tetraloop. In other words, the particular shape of the backbone around C7 nucleotides was not maintained.

Additional increase in fluctuations was found near the canonical stem of Kt-77, which continues as H78 in bacterial structure but is abruptly stubbed with a GCAA tetraloop in Nikulin et al. structure. While internally the tetraloop was fairly stable in classical GNRA topology in all simulations, globally there was increase in fluctuations in comparison to the equivalent region in the full bacterial structure. These fluctuations were smaller than those for L2111/2119, since most of their impact on the structure is probably adsorbed by the Kt-77 hinge-like movements and flexibility.

Inclusion of Mg^{2+} Deteriorates the Simulation. In one simulation, we included three crystallographic Mg^{2+} ions. They have an incomplete coordinate shell in the crystal structure, and the second Mg^{2+} has initially only one inner-shell contact with the G21(O6) atom. One Mg^{2+} forms an inner shell contact with the G16(O1P) atom early in the simulation and does not interact with RNA in any other way. The second Mg^{2+} forms two orthogonal inner shell contacts with G21(O6) and U43(O1P) atoms. In comparison with non-Mg^{2+} simulations, we observed reduced fluctuations of G21 base (the non-canonical stem of Kt-77) and its different position in relation to the U43 phosphate group. In absence of Mg^{2+}, it interacts with the U43 phosphate group. The third Mg^{2+} is experimentally located in the vicinity of the U9/A41 base pair; however, in our simulation, it moves away from the pair and forms an inner shell contact with the A15(O1P) atom. It then disrupts the base stack formed by nucleotides C7 and U14. The C7 nucleotide flips completely over, and its base comes to interact with the *t*HH G8/A40 base pair (the tRNA interface). This leads to complete loss of the *t*HH G8/A40 and *t*HH U9/A41 base pairs, and perturbation of the remaining interactions between the L2111/2119 and the Kt-77 bulge. In conclusion, the first Mg^{2+} seems relatively unimportant, the second might be important for Kt-77 structure, and the third introduces (in our opinion) profound errors in the simulation. Thus, inclusion of Mg^{2+} into RNA simulations, due to both force field and sampling limitations, is quite a controversial issue.[90,91] The structural changes in the vicinity of the third Mg^{2+} are almost certainly caused by a generally unsatisfying MM description of divalent cations and may also indicate inaccuracy in its initial placement. Nevertheless, the simulation behavior also suggests that there is more structural flexibility in the vicinity of nucleotides C7 and U14 in the absence of H78, compared with full bacterial structure.

Possible Role of H78. One of the evolutionary reasons for formation of helix 78 (and subsequent placing of the tetraloop next to L2111/2119) could be to provide base pairing for bases C7 and U14 with stabilization of the backbone topology of L2111/2119. Note that, in full bacterial structure, there are no unpaired bases in the L2111/2119. However, without helix H78, there would be two unpaired bases in what is most likely the critical part of the L1 stalk (the tRNA binding interface). Instability in that region could affect function of the L1 stalk and ribosome. An obvious hurdle for this suggestion is the fact that helix H78 is naturally missing in most eukaryotes. Since there are insufficient structural data for the eukaryotic L1 structure, it is difficult to speculate on the nature of its interactions. Perhaps, the *c*WW G2452/A2494 base pair (experimental numbering)[15] which lies next to the unpaired nucleotides of L2111/2119 might play a role in compensating instability arising from the absence of the tetraloop. Note that this is only a putative base pair arrangement, since the base pairs are distorted in the X-ray structure;[15] cf. Figure 2. While only canonical base pairs are found in this position in bacteria and archaea, the non-canonical G/A base pair is highly conserved in this position in eukaryotes.

Apart from the above-mentioned differences, the Nikulin et al. structure behaves in simulations similarly to the bacterial structures.

DISCUSSION

We have carried out series of bioinformatics and MD simulation investigations to better understand the structural dynamics and conservation patterns of the L1 stalk RNA from the large ribosomal subunit. A difficulty in computational studies of dynamical parts of the ribosome, such as the L1 stalk,

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

is the limited number of experimental structures with atomistic resolution. We base our study on all the available atomic-resolution X-ray structures of the L1 stalk.[11,13,14] We concentrate our efforts on non-canonical elements of the L1 stalk RNA (Figure 1), which includes two kink-turns (Figure 4), a tetraloop, and an internal loop (Figure 3). Over 11.5 μs of MD simulations are carried out for a number of systems, ranging from simulations of the individual RNA building blocks up to contextual simulations of the whole L1 stalk RNA and its complexes with L1 protein and relevant tRNA fragment.

Phylogenetic analysis indicates high evolutionary conservation of L1 stalk non-canonical elements. The exception is kink-turn Kt-78 which is deleted along with the entire helix H78 in eukaryotes and the tetraloop which displays highly variable sequences.

Our analysis shows that the tetraloop nucleotide sequences for bacteria and archaea follow the GNNA and UNNG sequence patterns, with the GCCA tetraloop having ~50% occurrence in bacteria. This sequence pattern differs from the most common GNRA and YNMG tetraloop families, which means that L1 tetraloops necessarily lack some signature interactions of the dominant RNA tetraloop families (Supporting Information, Tables S2 and S3). The L1 tetraloop sequence patterns are likely enforced by a strongly conserved tertiary 2111/2147 base pair between the tetraloop and the L2111/2119 loop, which allows one to sacrifice the otherwise strict GNRA and UNCG sequence requirements (Figure 5). This tertiary interaction is missing in eukaryotes, and concomitantly, the tetraloop sequences in eukaryotes follow the standard GNRA or YNMG patterns.

We explore a major difference between the *T.t.* X-ray L1 stalk structures, where the G_{L4} base of the UCCG tetraloop is *anti* in one structure[13] and *syn* in the other.[14] It cannot be ruled out that the *anti* variant is a mis-refinement. However, straightforward rejection of the *anti* conformation is not possible due to the G_{L4} (i.e., G2147) tertiary base pair with L2111/2119 which can be formed by both G_{L4} conformations, albeit using different base pairing family (Figure 5). On the basis of the overall simulation data, we nevertheless suggest that G_{L4} should be rather *syn*. However, coexistence of both *syn* and *anti* G_{L4} substates cannot be ruled out and may be pH dependent, since the intrinsically more stable *syn* variant of the tetraloop requires protonated C2111 to form the tertiary interaction.

Since most bacteria have a GCCA tetraloop instead of UCCG, we also executed several simulations with a GCCA tetraloop (a previously undescribed type of tetraloop) modeled into the L1 stalk structure. More stable trajectories were evidently achieved with A_{L4} in *syn*. We suggest that the GCCA tetraloop is capable of structurally adapting to the YNMG (UNCG) tetraloop family type of fold.

Kink-turn Kt-77 is unusual because of the 8 nt insertion into its bulge (Figure 4). Kt-77 is very stable in simulations, while its dynamics is slightly modulated compared to common kink-turns.[50] The inserted bulge reduces large scale openings of the structure and modifies subtly the direction of the hinge-like properties and A-minor dynamics of the kink-turn. The bulge internal structure (except for the nucleotides that correspond to common kink-turn consensus) is lost in simulations of isolated Kt-77 but well maintained once the surrounding RNA elements are included. All data support the view that the long bulge is an additional independent RNA element inserted into the kink-turn and extruding from it to establish specific tertiary interactions.

The L2111/2119 element (Figure 3) loses its native structure in isolation, while it is well maintained in simulations where its context is included.

Simulations of near-consensual kink-turn Kt-78 (Figure 4) resemble simulations of other kink-turns reported in the literature.[21−23,50,72,84−87] Overall, the simulation behavior is quite satisfactory, but there is local instability of the key kink-turn sugar-edge essential interaction between the bulge and the non-canonical stem (Supporting Information, Figure S2). We suggest this is a currently unresolved force field issue. The essential interaction is rearranged after a few ns with every tested force field, including the latest bsc0χ_{OL3} RNA force field. The rearranged geometry has never been seen in experimental structures. Interestingly, we did not observe this problem in Kt-77 simulations, where the kink-turn essential interaction was perfectly stable. This suggests that the inserted nucleotides in the bulge of Kt-77 likely alleviate some of the internal struggles between various signature structural features of the kink-turn, making the force field description more robust.

Another common feature of Kt-78 simulations was random transformation of A-minor interaction from 0 into I in several simulations. In later parts (after hundreds of ns) of some simulations, we additionally observed the formation of an alternative kink-turn structure or substate with a rearranged non-canonical stem. We suggest this could be a beginning of the kink-turn unfolding path (Figure 10 and the Supporting Information).

We suggest that the helix H78 evolutionary role in bacteria and archaea might be to bring the tetraloop into the vicinity of L2111/2119 to provide additional stabilization for otherwise unpaired nucleotides. This is evidenced by low conservation of the H78 sequence, its highly conserved length, and the lower stability of L2111/2119 in simulations where H78 is missing. This suggestion is not applicable to eukaryotes where the H78 is naturally missing.

We suggest a possible mechanism for tRNA release, where disruption of one of the contacts between tRNA and the L1 stalk is translated through the L1 stalk internal interactions into disruption of the remaining interactions (Figure 11). This could accelerate both tRNA binding and release events by ensuring that all interactions between the L1 stalk and tRNA are stabilized or disrupted in a coordinated manner.

Our study brings also some methodological insights. Overall, the simulation behavior is very satisfactory. We confirm improved performance of the bsc0χ_{OL3} force field over the earlier force field version for several structural aspects seen in the L1 stalk RNA. The local problems with essential interaction in the kink-turns (see above), however, persist. We also illustrate that interpretation of Mg^{2+} behavior in molecular dynamics simulations is difficult (Supporting Information). While they mostly did not affect the structure in any way (at least on the simulation time scale), in some cases, they introduced perturbation into the simulated structures. Monovalent ions in simulations can often occupy Mg^{2+} sites seen in X-ray structures (Supporting Information, Figure S1 and Table S4), as demonstrated for other RNA systems in the past.[79−83]

CONCLUSIONS

Our study analyzes unusual GNNA and UNNG conservation patterns for the L1 stalk tetraloop in bacteria and archaea. We suggest they are enforced by highly conserved tertiary contacts. There is a difference between available structures of the UCCG L1 stalk tetraloop, where the fourth base of the tetraloop is

dx.doi.org/10.1021/jp401482m | J. Phys. Chem. B XXXX, XXX, XXX−XXX

either in *anti*[13] or *syn*[14] conformation. We suggest this difference may be a mis-refinement in the case of the *anti* conformation. However, we also provide evidence that the *syn* and *anti* states can be coexisting in a pH-dependent manner. We used molecular modeling and simulations to explore the structure of the hitherto unobserved GCCA tetraloop, which is dominant in bacteria based on sequence analysis. We suggest it is structurally similar to the YNMG (UNCG) tetraloop family.

Simulations of Kt-77 show that its behavior is only slightly modified by its unusually long bulge. Kt-78 is a near-consensual kink-turn with standard behavior in simulations. We confirm the instability of the Kt-78 essential interaction, which was identified in previous kink-turn computational studies.[50,72] We conclude that it is a force field inaccuracy persistent even with the latest bsc0χ_{OL3} RNA force field. Nevertheless, the overall performance of the bsc0χ_{OL3} force field in our simulations (more than 11 μs in total) is quite satisfactory.

On the basis of our simulations, we suggest a possible evolutionary role of helix H78, which is the least conserved part of the L1 stalk rRNA. The mechanism of tRNA binding and release from the L1 stalk is briefly discussed.

The behavior of Mg^{2+} ions in molecular dynamics simulations and errors that the divalent ions can introduce are briefly commented on.

■ ASSOCIATED CONTENT

Supporting Information

Description of kink-turn and tetraloop structure; simulations of Kt-77 with bulge reduction; description of Kt-78 simulations and alternative structure; summary and discussion about simulations done with other force fields; behavior of Mg^{2+} in simulations; tables and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sponer@ncbr.muni.cz. Phoone: +420 541 517 133. Fax: +420 541 212 179.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been supported by the Grant Agency of the Czech Republic (Grant No. P305/12/G034). Institutional support has been provided by project "CEITEC - Central European Institute of Technology" (CZ.1.05/1.1.00/02.0068) from European Regional Development Fund.

■ REFERENCES

(1) Blanchard, S. C.; Kim, H. D.; Gonzalez, R. L.; Puglisi, J. D.; Chu, S. tRNA Dynamics on the Ribosome during Translation. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12893−12898.

(2) Fei, J.; Kosuri, P.; MacDougall, D. D.; Gonzalez, R. L. Coupling of Ribosomal L1 Stalk and tRNA Dynamics during Translation Elongation. *Mol. Cell* **2008**, *30*, 348−359.

(3) Aitken, C. E.; Puglisi, J. D. Following the Intersubunit Conformation of the Ribosome during Translation in Real Time. *Nat. Struct. Mol. Biol.* **2010**, *17*, 793−U735.

(4) Cornish, P. V.; Ermolenko, D. N.; Staple, D. W.; Hoang, L.; Hickerson, R. P.; Noller, H. F.; Ha, T. Following Movement of the L1 Stalk between Three Functional States in Single Ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 2571−2576.

(5) Munro, J. B.; Wasserman, M. R.; Altman, R. B.; Wang, L. Y.; Blanchard, S. C. Correlated Conformational Events in EF-G and the Ribosome Regulate Translocation. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1470−1477.

(6) Fei, J. Y.; Richard, A. C.; Bronson, J. E.; Gonzalez, R. L. Transfer RNA-Mediated Regulation of Ribosome Dynamics during Protein Synthesis. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1043−1051.

(7) Tama, F.; Valle, M.; Frank, J.; Brooks, C. L. Dynamic Reorganization of the Functionally Active Ribosome Explored by Normal Mode Analysis and Cryo-Electron Microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9319−9323.

(8) Munro, J. B.; Altman, R. B.; Tung, C. S.; Cate, J. H. D.; Sanbonmatsu, K. Y.; Blanchard, S. C. Spontaneous Formation of the Unlocked State of the Ribosome is a Multistep Process. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 709−714.

(9) Chen, C. L.; Stevens, B.; Kaur, J.; Cabral, D.; Liu, H. Q.; Wang, Y. H.; Zhang, H. B.; Rosenblum, G.; Smilansky, Z.; Goldman, Y. E.; et al. Single-Molecule Fluorescence Measurements of Ribosomal Translocation Dynamics. *Mol. Cell* **2011**, *42*, 367−377.

(10) Korostelev, A.; Ermolenko, D. N.; Noller, H. F. Structural Dynamics of the Ribosome. *Curr. Opin. Chem. Biol.* **2008**, *12*, 674−683.

(11) Nikulin, A.; Eliseikina, I.; Tishchenko, S.; Nevskaya, N.; Davydova, N.; Platonova, O.; Piendl, W.; Selmer, M.; Liljas, A.; Drygin, D.; et al. Structure of the L1 Protuberance in the Ribosome. *Nat. Struct. Biol.* **2003**, *10*, 104−108.

(12) Drygin, D.; Zimmermann, R. A. Magnesium Ions Mediate Contacts between Phosphoryl Oxygens at Positions 2122 and 2176 of the 23S rRNA and Ribosomal Protein L1. *RNA* **2000**, *6*, 1714−1726.

(13) Gao, Y. G.; Selmer, M.; Dunham, C. M.; Weixlbaumer, A.; Kelley, A. C.; Ramakrishnan, V. The Structure of the Ribosome with Elongation Factor G Trapped in the Posttranslocational State. *Science* **2009**, *326*, 694−699.

(14) Tishchenko, S.; Gabdulkhakov, A.; Nevskaya, N.; Sarskikh, A.; Kostareva, O.; Nikonova, E.; Sycheva, A.; Moshkovskii, S.; Garber, M.; Nikonov, S. High-Resolution Crystal Structure of the Isolated Ribosomal L1 Stalk. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 1051−1057.

(15) Ben-Shem, A.; Jenner, L.; Yusupova, G.; Yusupov, M. Crystal Structure of the Eukaryotic Ribosome. *Science* **2010**, *330*, 1203−1209.

(16) Ben-Shem, A.; de Loubresse, N. G.; Melnikov, S.; Jenner, L.; Yusupova, G.; Yusupov, M. The Structure of the Eukaryotic Ribosome at 3.0 Angstrom Resolution. *Science* **2011**, *334*, 1524−1529.

(17) Klein, D. J.; Schmeing, T. M.; Moore, P. B.; Steitz, T. A. The Kink-Turn: A New RNA Secondary Structure Motif. *EMBO J.* **2001**, *20*, 4214−4221.

(18) Gourse, R. L.; Thurlow, D. L.; Gerbi, S. A.; Zimmermann, R. A. Specific Binding of a Prokaryotic Ribosomal-Protein to a Eukaryotic Ribosomal-RNA - Implications for Evolution and Auto-Regulation. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2722−2726.

(19) Lescoute, A.; Leontis, N. B.; Massire, C.; Westhof, E. Recurrent Structural RNA Motifs, Isostericity Matrices and Sequence Alignments. *Nucleic Acids Res.* **2005**, *33*, 2395−2409.

(20) Reblova, K.; Razga, F.; Li, W.; Gao, H. X.; Frank, J.; Sponer, J. Dynamics of the Base of Ribosomal A-Site Finger Revealed by Molecular Dynamics Simulations and Cryo-EM. *Nucleic Acids Res.* **2010**, *38*, 1325−1340.

(21) Razga, F.; Koca, J.; Sponer, J.; Leontis, N. B. Hinge-Like Motions in RNA Kink-Turns: The Role of the Second A-Minor Motif and Nominally Unpaired Bases. *Biophys. J.* **2005**, *88*, 3466−3485.

(22) Razga, F.; Spackova, N.; Reblova, K.; Koca, J.; Leontis, N. B.; Sponer, J. Ribosomal RNA Kink-Turn Motif - A Flexible Molecular Hinge. *J. Biomol. Struct. Dyn.* **2004**, *22*, 183−193.

(23) Razga, F.; Zacharias, M.; Reblova, K.; Koca, J.; Sponer, J. RNA Kink-Turns as Molecular Elbows: Hydration, Cation Binding, and Large-Scale Dynamics. *Structure* **2006**, *14*, 825−835.

(24) Tuerk, C.; Gauss, P.; Thermes, C.; Groebe, D. R.; Gayle, M.; Guild, N.; Stormo, G.; Daubentoncarafa, Y.; Uhlenbeck, O. C.; Tinoco, I.; et al. CUUCGG Hairpins - Extraordinarily Stable RNA Secondary Structures Associated with Various Biochemical Processes. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1364−1368.

(25) Uhlenbeck, O. C. Nucleic-Acid Structure - Tetraloops and RNA Folding. *Nature* **1990**, *346*, 613−614.
(26) Woese, C. R.; Winker, S.; Gutell, R. R. Architecture of Ribosomal-RNA - Constraints on the Sequence of Tetra-Loops. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8467−8471.
(27) Bevilacqua, P. C.; Blose, J. M. Structures, Kinetics, Thermodynamics, and Biological Functions of RNA Hairpins. *Annu. Rev. Phys. Chem.* **2008**, *59*, 79−103.
(28) Schroeder, K. T.; McPhee, S. A.; Ouellet, J.; Lilley, D. M. J. A Structural Database for K-Turn Motifs in RNA. *RNA* **2010**, *16*, 1463−1468.
(29) Ennifar, E.; Nikulin, A.; Tishchenko, S.; Serganov, A.; Nevskaya, N.; Garber, M.; Ehresmann, B.; Ehresmann, C.; Nikonov, S.; Dumas, P. The Crystal Structure of UUCG Tetraloop. *J. Mol. Biol.* **2000**, *304*, 35−42.
(30) Tishchenko, S.; Nikulin, A.; Fomenkova, N.; Nevskaya, N.; Nikonov, O.; Dumas, P.; Moine, H.; Ehresmann, B.; Ehresmann, C.; Piendl, W.; et al. Detailed Analysis of RNA-Protein Interactions within the Ribosomal Protein S8-rRNA Complex from the Archaeon Methanococcus Jannaschii. *J. Mol. Biol.* **2001**, *311*, 311−324.
(31) Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. Functional Insights from the Structure of the 30S Ribosomal Subunit and its Interactions with Antibiotics. *Nature* **2000**, *407*, 340−348.
(32) Nozinovic, S.; Furtig, B.; Jonker, H. R. A.; Richter, C.; Schwalbe, H. High-Resolution NMR Structure of an RNA Model System: The 14-mer cUUCGg Tetraloop Hairpin RNA. *Nucleic Acids Res.* **2010**, *38*, 683−694.
(33) Williams, D. J.; Hall, K. B. Experimental and Theoretical Studies of the Effects of Deoxyribose Substitutions on the Stability of the UUCG Tetraloop. *J. Mol. Biol.* **2000**, *297*, 251−265.
(34) Allain, F. H. T.; Varani, G. Structure of the P1 Helix from Group-I Self-Splicing Introns. *J. Mol. Biol.* **1995**, *250*, 333−353.
(35) Paladino, A.; Zangi, R. Ribose 2′-Hydroxyl Groups Stabilize RNA Hairpin Structures Containing GCUAA Pentaloop. *J. Chem. Theory Comput.* **2013**, *9*, 1214−1221.
(36) Brosius, J.; Dull, T. J.; Noller, H. F. Complete Nucleotide-Sequence of a 23S Ribosomal-RNA Gene from Escherichia-Coli. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 201−204.
(37) Leontis, N. B.; Stombaugh, J.; Westhof, E. The Non-Watson-Crick Base Pairs and Their Associated Isostericity Matrices. *Nucleic Acids Res.* **2002**, *30*, 3497−3531.
(38) Zirbel, C. L.; Sponer, J. E.; Sponer, J.; Stombaugh, J.; Leontis, N. B. Classification and Energetics of the Base-Phosphate Interactions in RNA. *Nucleic Acids Res.* **2009**, *37*, 4898−4918.
(39) Nissen, P.; Ippolito, J. A.; Ban, N.; Moore, P. B.; Steitz, T. A. RNA Tertiary Interactions in the Large Ribosomal Subunit: The A-Minor Motif. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4899−4903.
(40) Doherty, E. A.; Batey, R. T.; Masquida, B.; Doudna, J. A. A Universal Mode of Helix Packing in RNA. *Nat. Struct. Biol.* **2001**, *8*, 339−343.
(41) Szewczak, L. B. W.; Gabrielsen, J. S.; DeGregorio, S. J.; Strobel, S. A.; Steitz, J. A. Molecular Basis for RNA Kink-Turn Recognition by the h15.5K Small RNP Protein. *RNA* **2005**, *11*, 1407−1419.
(42) Liu, J.; Lilley, D. M. J. The Role of Specific 2′-Hydroxyl Groups in the Stabilization of the Folded Conformation of Kink-Turn RNA. *RNA* **2007**, *13*, 200−210.
(43) Hanner, M.; Mayer, C.; Kohrer, C.; Golderer, G.; Grobner, P.; Piendl, W. Autogenous Translational Regulation of the Ribosomal MVAL1 Operon in the Archaebacterium Methanococcus-Vannielii. *J. Bacteriol.* **1994**, *176*, 409−418.
(44) Said, B.; Cole, J. R.; Nomura, M. Mutational Analysis of the L1 Binding-Site of 23S Ribosomal-RNA in Escherichia-Coli. *Nucleic Acids Res.* **1988**, *16*, 10529−10545.
(45) Yates, J. L.; Arfsten, A. E.; Nomura, M. Invitro Expression of Escherichia-Coli Ribosomal-Protein Genes - Autogenous Inhibition of Translation. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 1837−1841.
(46) Dean, D.; Nomura, M. Feedback-Regulation of Ribosomal-Protein Gene-Expression in Escherichia-Coli. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 3590−3594.
(47) Baughman, G.; Nomura, M. Translational Regulation of the L11 Ribosomal-Protein Operon of Escherichia-Coli - Analysis of the Messenger-RNA Target Site Using Oligonucleotide-Directed Mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 5389−5393.
(48) Nevskaya, N.; Tishchenko, S.; Gabdoulkhakov, A.; Nikonova, E.; Nikonov, O.; Nikulin, A.; Platonova, O.; Garber, M.; Nikonov, S.; Piendl, W. Ribosomal Protein L1 Recognizes the Same Specific Structural Motif in Its Target Sites on the Autoregulatory mRNA and 23S rRNA. *Nucleic Acids Res.* **2005**, *33*, 478−485.
(49) Trabuco, L. G.; Schreiner, E.; Eargle, J.; Cornish, P.; Ha, T.; Luthey-Schulten, Z.; Schulten, K. The Role of L1 Stalk-tRNA Interaction in the Ribosome Elongation Cycle. *J. Mol. Biol.* **2010**, *402*, 741−760.
(50) Reblova, K.; Sponer, J. E.; Spackova, N.; Besseova, I.; Sponer, J. A-Minor Tertiary Interactions in RNA Kink-Turns. Molecular Dynamics and Quantum Chemical Analysis. *J. Phys. Chem. B* **2011**, *115*, 13897−13910.
(51) Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C. SCOP - A Structural Classification of Proteins Database for the Investigation of Sequences and Structures. *J. Mol. Biol.* **1995**, *247*, 536−540.
(52) Frishman, D.; Argos, P. Knowledge-Based Protein Secondary Structure Assignment. *Proteins: Struct., Funct., Bioinf.* **1995**, *23*, 566−579.
(53) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* **1996**, *100*, 19824−19839.
(54) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. A Modified Version of the Cornell et al. Force Field with Improved Sugar Pucker Phases and Helical Repeat. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.
(55) Banas, P.; Hollas, D.; Zgarbova, M.; Jurecka, P.; Orozco, M.; Cheatham, T. E.; Sponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.* **2010**, *6*, 3836−3849.
(56) Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E.; Jurecka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886−2902.
(57) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinenement of the AMBER Force Field for Nucleic Acids: Improving the Description of Alpha/Gamma Conformers. *Biophys. J.* **2007**, *92*, 3817−3829.
(58) 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.; Kollman, P. A. A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(59) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(60) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(61) Jorgensen, W. L. Quantum and Statistical Mechanical Studies of Liquids 0.10. Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers - Application to Liquid Water. *J. Am. Chem. Soc.* **1981**, *103*, 335−340.
(62) Aqvist, J. Ion Water Interaction Potentials Derived from Free-Energy Perturbation Simulations. *J. Phys. Chem.* **1990**, *94*, 8021−8024.
(63) Dang, L. X.; Kollman, P. A. Free-Energy of the Association of the K+18-CROWN-6 Complex in Water - A New Molecular-Dynamics Study. *J. Phys. Chem.* **1995**, *99*, 55−58.

(64) Smith, D. E.; Dang, L. X. Computer-Simulations of NaCl Association in Polarizable Water. *J. Chem. Phys.* **1994**, *100*, 3757−3766.
(65) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; et al. *AMBER 10*; University of California: San Francisco, CA, 2008.
(66) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - An N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(67) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(68) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(69) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(70) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(71) Jossinet, F.; Westhof, E. Sequence to Structure (S2S): Display, Manipulate and Interconnect RNA Data from Sequence to Structure. *Bioinformatics* **2005**, *21*, 3320−3321.
(72) Spackova, N.; Reblova, K.; Sponer, J. Structural Dynamics of the Box C/D RNA Kink-Turn and Its Complex with Proteins: The Role of the A-Minor 0 Interaction, Long-Residency Water Bridges, and Structural Ion-Binding Sites Revealed by Molecular Simulations. *J. Phys. Chem. B* **2010**, *114*, 10581−10593.
(73) Mokdad, A.; Leontis, N. B. Ribostral: An RNA 3D Alignment Analyzer and Viewer Based on Basepair Isostericities. *Bioinformatics* **2006**, *22*, 2168−2170.
(74) Cannone, J. J.; Subramanian, S.; Schnare, M. N.; Collett, J. R.; D'Souza, L. M.; Du, Y. S.; Feng, B.; Lin, N.; Madabusi, L. V.; Muller, K. M.; et al. The Comparative RNA Web (CRW) Site: An Online Database of Comparative Sequence and Structure Information for Ribosomal, Intron, and Other RNAs. *BMC Bioinf.* **2002**, *3*, 1−31.
(75) Thomas, M. S.; Nomura, M. Translational Regulation of the L11 Ribosomal-Protein Operon of Escherichia-Coli - Mutations that Define the Target Site for Repression by L1. *Nucleic Acids Res.* **1987**, *15*, 3085−3096.
(76) Razga, F.; Koca, J.; Mokdad, A.; Sponer, J. Elastic Properties of Ribosomal RNA Building Blocks: Molecular Dynamics of the GTPase-Associated Center rRNA. *Nucleic Acids Res.* **2007**, *35*, 4007−4017.
(77) Kührová, P.; Banáš, P.; Best, R. B.; Šponer, J.; Otyepka, M. Computer Folding of RNA Tetraloops? Are We There Yet? *J. Chem. Theory Comput.* **2013**, *9*, 2115−2125.
(78) Stombaugh, J.; Zirbel, C. L.; Westhof, E.; Leontis, N. B. Frequency and Isostericity of RNA Base Pairs. *Nucleic Acids Res.* **2009**, *37*, 2294−2312.
(79) Krasovska, M. V.; Sefcikova, J.; Reblova, K.; Schneider, B.; Walter, N. G.; Sponer, J. Cations and Hydration in Catalytic RNA: Molecular Dynamics of the Hepatitis Delta Virus Ribozyme. *Biophys. J.* **2006**, *91*, 626−638.
(80) Reblova, K.; Spackova, N.; Stefl, R.; Csaszar, K.; Koca, J.; Leontis, N. B.; Sponer, J. Non-Watson-Crick Basepairing and Hydration in RNA Motifs: Molecular Dynamics of 5S rRNA Loop E. *Biophys. J.* **2003**, *84*, 3564−3582.
(81) Reblova, K.; Spackova, N.; Sponer, J. E.; Koca, J.; Sponer, J. Molecular Dynamics Simulations of RNA Kissing-Loop Motifs Reveal Structural Dynamics and Formation of Cation-Binding Pockets. *Nucleic Acids Res.* **2003**, *31*, 6942−6952.
(82) Singh, A.; Sethaphong, L.; Yingling, Y. G. Interactions of Cations with RNA Loop-Loop Complexes. *Biophys. J.* **2011**, *101*, 727−735.
(83) Auffinger, P.; Bielecki, L.; Westhof, E. Symmetric K+ and Mg2+ Ion-Binding Sites in the 5 S rRNA Loop E Inferred from Molecular Dynamics Simulations. *J. Mol. Biol.* **2004**, *335*, 555−571.
(84) Cojocaru, V.; Klement, R.; Jovin, T. M. Loss of G-A Base Pairs is Insufficient for Achieving a Large Opening of U4 snRNA K-Turn Motif. *Nucleic Acids Res.* **2005**, *33*, 3435−3446.
(85) Cojocaru, V.; Nottrott, S.; Klement, R.; Jovin, T. M. The snRNP 15.5K Protein Folds its Cognate K-turn RNA: A Combined Theoretical and Biochemical Study. *RNA* **2005**, *11*, 197−209.
(86) Curuksu, J.; Sponer, J.; Zacharias, M. Elbow Flexibility of the Kt38 RNA Kink-Turn Motif Investigated by Free-Energy Molecular Dynamics Simulations. *Biophys. J.* **2009**, *97*, 2004−2013.
(87) Sklenovsky, P.; Florova, P.; Banas, P.; Reblova, K.; Lankas, F.; Otyepka, M.; Sponer, J. Understanding RNA Flexibility Using Explicit Solvent Simulations: The Ribosomal and Group I Intron Reverse Kink-Turn Motifs. *J. Chem. Theory Comput.* **2011**, *7*, 2963−2980.
(88) Banas, P.; Sklenovsky, P.; Wedekind, J. E.; Sponer, J.; Otyepka, M. Molecular Mechanism of preQ(1) Riboswitch Action: A Molecular Dynamics Study. *J. Phys. Chem. B* **2012**, *116*, 12721−12734.
(89) Voorhees, R. M.; Schmeing, T. M.; Kelley, A. C.; Ramakrishnan, V. The Mechanism for Activation of GTP Hydrolysis on the Ribosome. *Science* **2010**, *330*, 835−838.
(90) Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. Molecular Dynamics and Quantum Mechanics of RNA: Conformational and Chemical Change We Can Believe In. *Acc. Chem. Res.* **2010**, *43*, 40−47.
(91) Sponer, J.; Cang, X. H.; Cheatham, T. E. Molecular Dynamics Simulations of G-DNA and Perspectives on the Simulation of Nucleic Acid Structures. *Methods* **2012**, *57*, 25−39.