An RNA molecular switch: Intrinsic flexibility of 23S rRNA Helices 40 and 68 5'-UAA/5'-GAN internal loops studied by molecular dynamics methods

Kamila Réblová[1,*], **Zora Střelcová**[2], **Petr Kulhánek**[2], **Ivana Beššeová**[1], **David H. Mathews**[3], **Keith Van Nostrand**[3], **Ilyas Yildirim**[4], **Douglas H. Turner**[5], and **Jiří Šponer**[1,*]

[1]Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 61265 Brno, Czech Republic

[2]National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5, 62500 Brno, Czech Republic

[3]Department of Biochemistry & Biophysics, University of Rochester Medical Center, 601 Elmwood Avenue, Box 712, Rochester, New York 14642

[4]Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627

[5]Department of Chemistry, University of Rochester, Rochester, New York 14627-0216

Abstract

Functional RNA molecules such as ribosomal RNAs frequently contain highly conserved internal loops with a 5'-UAA/5'-GAN (UAA/GAN) consensus sequence. The UAA/GAN internal loops adopt distinctive structure inconsistent with secondary structure predictions. The structure has a narrow major groove and forms a *trans* Hoogsteen/Sugar edge (*t*HS) A/G base pair followed by an unpaired stacked adenine, a *trans* Watson-Crick/Hoogsteen (*t*WH) U/A base pair and finally by a bulged nucleotide (N). The structure is further stabilized by a three-adenine stack and base-phosphate interaction. In the ribosome, the UAA/GAN internal loops are involved in extensive tertiary contacts, mainly as donors of A-minor interactions. Further, this sequence can adopt an alternative 2D/3D pattern stabilized by a four-adenine stack involved in a smaller number of tertiary interactions. The solution structure of an isolated UAA/GAA internal loop shows substantially rearranged base pairing with three consecutive non-Watson-Crick base pairs. Its A/U base pair adopts an incomplete *cis* Watson-Crick/Sugar edge (*c*WS) A/U conformation instead of the expected Watson-Crick arrangement. We performed 3.1 µs of explicit solvent molecular dynamics (MD) simulations of the X-ray and NMR UAA/GAN structures, supplemented by MM-PBSA free energy calculations, locally enhanced sampling (LES) runs, targeted MD (TMD) and nudged elastic band (NEB) analysis. We compared parm99 and parmbsc0 force fields and net-neutralizing Na$^+$ vs. excess salt KCl ion environments. Both force fields provide a similar description of the simulated structures, with the parmbsc0 leading to modest narrowing of the major groove. The excess salt simulations also cause a similar effect. While the NMR structure is entirely stable in simulations, the simulated X-ray structure shows considerable widening of the

Corresponding authors: Kamila Réblová, Tel.: +420 541517250, Fax: +420 541212179. kristina@physics.muni.cz and Jiří Šponer, Tel.: +420 541517133, Fax: +420 541212179. sponer@ncbr.chemi.muni.cz.

major groove, loss of base-phosphate interaction and other instabilities. The alternative X-ray geometry even undergoes conformational transition towards the solution 2D structure. Free energy calculations confirm that the X-ray arrangement is less stable than the solution structure. LES, TMD and NEB provide a rather consistent pathway for interconversion between the X-ray and NMR structures. In simulations, the incomplete *c*WS A/U base pair of the NMR structure is water mediated and alternates with the canonical A–U base pair, which is not indicated by the NMR data. Completion of full *c*WS A/U base pair is prevented by the overall internal loop arrangement. In summary, the simulations confirm that the UAA/GAN internal loop is a molecular switch RNA module that adopts its functional geometry upon specific tertiary contexts.

Keywords

Molecular dynamics; RNA; ribosome; force field; internal loop

INTRODUCTION

RNA secondary structures comprise four basic elements such as helices, external loops (hairpin loops), internal loops and junction loops. First crystallographic structures of the ribosome[1−3] determined in this decade have revealed that the internal loops are structured by amazingly variable non-Watson-Crick base pairs and many of them form recurrent structural motifs with distinct shapes.[4] (By internal loops we mean short series of nominally unpaired bases within a longer paired helix, *i.e.*, bases that do not form canonical Watson-Crick base pairs). Thus, the RNA structures can be considered as fascinating combinations of short canonical helices responsible for major part of the thermodynamics stability and various non-canonical (non-Watson-Crick) functional elements with diverse sequences, shapes and flexibilities. Some of the recurrent motifs are autonomous, *i.e.* their structures are within the ribosome independent of context, while others have arrangements affected by surrounding structures, *i.e.* exhibit induced fit binding. In the present work, we investigate one of the most salient recurrent non-autonomous RNA structural motifs that adopt their functional shape only in very specific tertiary contexts. The aim is to complement the existing structural data by analyses utilizing the available computational methods based on classical atomistic explicit solvent simulations and to establish what kind of information can in principle be gathered using modern computations for such RNA structural elements. Typical example of this element occurs in Helix 40 (H40) of the large ribosomal subunit. H40 contains a highly conserved internal loop in all three domains of life with 5'-UAA/5'-GAN (UAA/GAN) consensus sequence (Figure 1A).[5]

This motif is present in seven internal loops of 23S rRNA and in other RNAs such as the RNase P RNAs and group I and II introns with different degrees of conservation.[5] Despite different locations and tertiary interactions, the majority of the UAA/GAN internal loops adopt a distinctive structure with an unpaired stacked adenine, and a bulged nucleotide (N). The three conserved adenines create a characteristic cross-strand AAA stack (Figure 1A).[5] An alternative secondary structure of the loop was seen in the crystal structures of the H68 of *Escherichia coli* (*E.c.*) 23S rRNA[9] (Figure 1B) and the intact RNase P RNA from *Bacillus stearothermophilus*[10] indicating structural plasticity of this motif.

Considering the structural data, spatial arrangement of this loop is likely to be dictated by surrounding ribosomal segments. In the ribosome, the H40 loop forms contacts with the hairpin structure between H39 and H40 by the adenines of the AAA stack via A-minor interactions.[5,11] A-minor interactions represent the most numerous and highly conserved tertiary interactions in large structured RNAs and ribonucleoproteins. The bulged nucleotide is involved in tertiary contacts as well. The bacterial H40 internal loop is additionally a part

of the binding site of the ribosomal protein L20.[12] Recently, the boxed motif in Figure 1A was classified as an UA_handle submotif, which is a highly versatile non-autonomous common RNA building block.[8]

The structure of the UAA/GAA motif flanked by Watson-Crick base pairs was determined in solution by NMR spectroscopy.[13] There are striking differences between the X-ray ribosomal H40 and the solution structure (Figure 1A, 1C). The pairing is restructured so that there are no unpaired or bulged bases. The solution structure is much more consistent with the arrangement expected based on 2D thermodynamics prediction, except that the standard 2D predictions would propose a canonical A–U base pair instead of the observed A/U base pair. The solution structure has a considerably wider (more open) major groove compared to the X-ray H40 UAA/GAA segment, which has a narrow (closed) major groove and wide minor groove (Figure 1). However, this feature may be a trivial consequence of the NMR structure refinement utilizing force field calculations. Thus, the functional structure seen in the crystals differs from the presumably intrinsically preferred arrangement seen in solution and is likely stabilized by tertiary and quaternary interactions.[13]

X-ray and NMR studies of RNAs can be complemented by molecular dynamics (MD), which provides dynamic data and additional insight into structure.[14−30] For instance, MD simulations can characterize isolated RNA building blocks independent of their structural context. The simulations capture intrinsic stochastic fluctuations of geometries and reveal intrinsic elastic properties that are important for function.[31−34] In addition, simulations can disclose whether an RNA building block is entirely internally relaxed or is deformed due to its interactions with the surroundings. In the latter case, simulations can reveal rapid structural relaxation towards low energy conformations.[35,36] These insights are unique and they would be hard to obtain by other methods. Thus, despite being limited by force field approximations and timescale, MD simulations provide valuable data that can complement experiments.[37−42]

In the present study, we run explicit solvent MD simulations on isolated internal loops of H40 starting with solution structure determined by NMR and the structure in ribosomes as determined by X-ray diffraction (Figure 1A, 1C). The study is complemented by MD simulation of the X-ray loop of H68 from the ribosome (Figure 1B). The main aim was to characterize base pairing and local arrangement of the loop on the nanosecond time scale to better understand its structural plasticity. The simulations were supplemented by free energy calculations that extract free energy directly from the trajectories. We also utilized Locally Enhanced Sampling (LES), nudged elastic band (NEB) and targeted MD (TMD) to investigate the pathway for the conformational change between the NMR and X-ray structures. Standard net-neutralizing Na^+ simulations with parm99 AMBER force field[43] were compared with parmbsc0[44] force field and excess salt KCl simulations. This study had two purposes. First, to investigate an important RNA modular block stabilized by tertiary contacts that appears to act as a flexible RNA structural switch. Second, to test capability of explicit solvent simulations and some auxiliary techniques to describe structural dynamics of RNA. The combination of methods provides interesting qualitative insights into the intricate properties of the UAA/GAN RNA internal loop.

Our standard simulations show relaxation of the ribosomal H40 loop. The X-ray conformation opens significantly and adopts an arrangement that resembles the solution structure. However, the X-ray secondary structure of the loop was maintained on the 200–300 ns time scale despite numerous local disturbances evidenced by the simulations. Thus, spontaneous transition of the H40 loop to the solution structure (presumably the global free energy minimum conformation) was not achieved. Disruption of the H40 X-ray secondary structure was detected in the LES simulations, where the NMR-like secondary structure was

sampled, albeit not dominantly. In contrast, almost perfect transition was observed in the standard simulation of the H68 loop, which is probably not as deformed as the ribosomal H40 loop due to reduced number of tertiary contacts. The free energy calculations clearly show that the solution structure of the UAA/GAA internal loop is more stable compared to its ribosomal geometries and reveal free energy change of the conformational transition between the H68 and NMR structures. Finally, LES, NEB and TMD calculations allowed us to propose a plausible mechanism by which the solution conformation could rearrange into "ribosomal" H40 X-ray geometry. Both parmbsc0 and KCl simulations are qualitatively in agreement with standard net-neutralizing Na$^+$ parm99 simulations. The major groove opening is, however, reduced compared to the Na$^+$ parm99 simulations.

MATERIALS AND METHODS

Starting structures

The H40 X-ray structure was taken from the available 50S subunits, *i.e.* from archaeal subunit of *Haloarcula marismortui* (*H.m.*) (pdb code 1S72)[45] and from bacterial subunits of *E.c.* (2AW4),[9] *Deinoccocus radiodurans* (*D.r.*) (1NKW)[2] and *Thermus thermophilus* (*T.t.*) (2J01)[46] (Figure 1A and Figure S1). We have compared conformations of 23S rRNA H40 from the recent structures of Ramakrishnan that include EF-G[47] and EF-Tu[48] to *T.t.* X-ray structure used in the present study. No visible structural differences were detected among the structures. RMSd between X-ray structure used in the present study and the newly released structures were below 1 Å, so they are identical. The H40 NMR structure was taken from pdb 2H49 (we utilized model 1)[13] (Figure 1C). Both X-ray and NMR structures have a UAA/GAA internal loop (only the *H.m.* helix contains UAA/GAG sequence) flanked by various Watson-Crick base pairs (Figure 1). In our study, the Watson-Crick base pairs in the X-ray structures (Figure S2) were mutated to match the NMR structure (Figure 1C), with an additional UAA/GAG->A substitution introduced for the *H.m.* structure. Thus all systems have the same sequence which allowed us to compare free energies in the studied systems (see below). Although all simulated structures have identical sequences, the starting structures still reflect some local structural differences of the X-ray structures. Moreover, we used the NMR nucleotide numbering (1–18) for X-ray H40s throughout the text to unify description of the systems (Figure 1). Original X-ray numbers are indicated in Figure 1 and Figure S2 and Table 1.

All the X-ray H40s have the same secondary structure and almost identical local geometry (Figure 1, Figures S1 and S2).

The X-ray internal loop comprises a sheared A6/G13 pair and rH U4/A14 pair, bulging A15 base and A5 base stacked within the stem without a base pairing partner on the other strand (Figure 1A). A5 forms an intrastrand stack with A6 base and an interstrand "cross"-strand stack with A14 base resulting in the characteristic AAA stack (Figure 1A). These interactions shape the internal loop into a specific arrangement exhibiting a broadened minor groove and narrow major groove (~6–8 Å) stabilized by base-phosphate (BPh) interaction.[49] *E.c.* X-ray structure of H40 exhibits bifurcated binding mode (peculiar alternative of base phosphate interaction type 4BPh) in which N2 and N1 of G13 bind to the same anionic oxygen of the phosphate group.[49] In particular there are G13(N1)-A5(O2P) and G13(N2)-A5(O2P) H-bonds (Figure 2). In contrast, *H.m.*, *D.r.* and *T.t.* X-ray structures exhibit base phosphate interaction type 5BPh including only G13(N1)-A5(O2P) H-bond (Figure 2).[49]

The NMR internal loop also shows the sheared A6/G13 base pair seen in the H40 X-ray structures; however, the loop contains two additional single hydrogen bonded non-canonical base pairs: a sheared A14/A5 and *c*WS A15/U4 base pair (Figure 1C). The latter base pair is classified as a *c*WS[6] although it contains only one direct A(N6)-U(O2) H-bond. The

simulations reveal that the other interaction characteristic for *c*WS A/U base pairs, the A(N1)-U(O2′) interaction, is in fact water-mediated. The major groove of the NMR structure is wide (open) (~17 Å) (Figure 1C) and it does not have any base-phosphate contacts across the groove.[13] The NMR structure reveals another AAA stack (Figure 1C) with a cross-strand A6/A14 and intrastrand A14/15 interactions.

The X-ray UAA/GAA loop of H68 was taken from the *E.c.* 50S (2AW4)[9] (Figure 1B). Base pairs above and below the loop were mutated according to the Watson-Crick base pairs of the NMR structure (Figure 1). The internal loop of H68 consists of unpaired G13 and A6 bases (Figure 1B) where the G13 base is in *syn* conformation. In the ribosomal H40 and in the solution structure the G13 is in *anti* orientation and in both these structures it forms a sheared A6/G13 pair (see Figure 1). Furthermore, the loop of H68 comprises stacked middle adenines A5 and A14 and imperfect *c*WS A15/U4 base pair stabilized by a single H-bond (A(N6)-U(O2)) (Figure 1B). In addition, the A5 and A14 bases form intrastrand stacks with adjacent adenines A6 and A15, respectively, resulting in a four-adenine stack (Figure 1B). The major groove is a little wider (~9–10 Å) compared with the H40 geometry.

Ribosomal contacts of H40

The studied UAA/GAN H40 internal loops are involved in identical A-minor interactions in all four 50S subunits. The three adenines of the cross-strand AAA stack interact with two consecutive highly conserved C=G base pairs of the hairpin between H39 and H40.[5] In particular, the adenine of the sheared A/G pair forms a type I A-minor interaction with one C=G pair while the A5 adenine forms a type II A-minor interaction with the next C=G pair (G=C in *T.t.*). Finally, the adenine of the rH U/A base pair forms a tilted variant of A-minor type I also with the second G=C base pair (Figure 3).[5]

In bacterial ribosomes, the minor groove side of the UAA/GAN motif interacts with the hairpin between H39 and H40 while the major groove side interacts with ribosomal protein L20 (Figure 3). In contrast, in the archaeal *H.m.* 50S the ribosomal protein L20 is substituted by ribosomal protein L30, which interacts with the minor groove side of the UAA/GAN internal loop (Figure 3). Moreover, the *H.m.* UAA/GAN has contact with H25 (Figure 3). Apparently, the UAA/GAN internal loop of H40 exhibits different interactions on its minor groove side in bacteria and archaea.

As noted by Lee et al.,[5] the UAA/GAN motifs in other helices adopting the same 2D/3D arrangement are also involved in extensive molecular contacts resembling those of H40 UAA/GAN. The alternative 2D/3D UAA/GAN conformation (H68 of *E.c.* 23S rRNA and in intact RNase P RNA) contacts just one RNA helix (Figure 3). This is another indication that structural context may alter geometry of the UAA/GAN motif.

Standard MD simulations

Standard explicit solvent simulations were carried out (at 300 K) using the pmemd module of AMBER 9.0[50,51] and force field parm99[43] version of the Cornell et al. force field[52] on a time scale of 200+ ns each. Control simulations of 100 ns were run with parmbsc0.[44] Parmbsc0 is the latest reparametrization of the Cornell et al. force field aimed to stabilize B-DNA simulations by penalizing substates with γ-trans backbone topologies. While parmbsc0 achieves a considerably improved performance at predicting DNA backbone conformations as compared to parm99, no systematic comparative testing was done for RNA with both parm99 and parmbsc0 until now. This work supports the viability of both force fields for RNA simulations. An overview of all simulations is given in Table 1. The total length of performed simulations was 3.1 µs. The simulated molecules were neutralized by standard AMBER Na$^+$ monovalent cations (radius 1.868 Å and well depth 0.00277 kcal/mol),[53]

initially placed using the xleap module of AMBER at the most negative sites around the solute. RNA molecules were solvated by a TIP3P water box to a depth of 10 Å. Net-neutralization corresponds to a cation concentration of ~0.15-0.2 M. Some control simulations were carried out in KCl with ~0.2 M excess salt concentration (Table 1). For this purpose we used either modified parameters for K^+ (radius 1.705 Å and well depth 0.1936829 kcal/mol) and Cl^- (radius 2.513 Å and well depth 0.0355910 kcal/mol), which prevents salt crystallization at low to medium salt concentrations,[54] or Dang's parameters for K^+ (radius 1.87 Å and well depth 0.1 kcal/mol)[55] and Cl^- (radius 2.47 Å and well depth 0.1 kcal/mol)[56] together with SPC/E water box.[57,58] The simulations were carried out using the particle mesh Ewald (PME) technique with a 9 Å nonbonded cutoff and a 2 fs integration time step. Trajectory was saved once a picosecond. Equilibration and production phases were carried using standard protocols.[20] Trajectories were analyzed using the ptraj module of AMBER and structures were visualized using the VMD molecular visualization program, http://www.ks.uiuc.edu/Research/vmd/.[59] The figures were prepared using VMD. The stability of stacking interactions within the AAA stack of H40 was monitored by calculating the van der Waals interaction energies between A5 and A14, A5 and A6, and A5 and G13 bases utilizing the anal module of AMBER.

To enhance sampling of H40, two simulations were carried out at an elevated temperature (400 K) (Table 1). The system was gradually heated from 300 K to 400 K during the first 100 ps using NPT conditions (constant pressure ensemble). The production runs were continued at 400 K using both NPT and NVT (constant volume ensemble). There are no clear guidelines whether NVT or NPT simulations should be preferred, and, in fact, both approaches have drawbacks.[60] Elevated temperature simulations were previously successfully applied to studies of a base substitution in an RNA frameshifting pseudoknot.[61] Additionally, a "no-salt" simulation of H40 in which cations were omitted from the simulation box was carried out (Table 1). The missing counterions were substituted by a net-neutralizing plasma representing a uniform neutralizing charge distribution over the box. This feature is implemented in the AMBER program package for use with the PME method and guarantees the neutrality of the system.[62] The aim of the no-salt simulation was to destabilize the simulated structure.

Locally enhanced sampling (LES) simulations

LES simulations[63–65] were carried out using the sander program of AMBER 9.0[50,51] to enlarge sampling of the internal loop started from the X-ray structures of H40 UAA/GAA internal loops. The addles module of AMBER was used to split the internal loop region into five independent copies, *i.e.* residues 4–6 and 13–15 were copied five times. Force field parameters for the copies were scaled, which lowered the energy barriers on the potential energy surface and increased the flexibility of the given region. Equilibration and production phase were carried out using standard protocols.[20] Heating during the equilibration phase was continued up to 300 K. The LES method along with explicit solvent simulations were successfully utilized in studies of several nucleic acids systems.[20,66–68]

Free energy calculations

The MM-PBSA method (Molecular Mechanics, Poisson-Boltzmann, and Surface Area)[69,70] implemented in AMBER 9.0[50,51] was used for free energy analysis of the explicit solvent MD trajectories. This method is based on a continuum solvent approach that replaces the explicit solvent and utilizes snapshots directly from the simulation. Here it was employed to estimate total free energy of H40 along MD_Ec_99, MD_Hm_99, MD_Dr_99, and MD_Tt_99 simulations and of H68 along MD_H68 simulation (see Table 1). Total free energy of the solution structure along MD_NMR_99 simulation was also obtained. The calculations of MM-PBSA energy included calculations of Molecular Mechanics energy

(EMM) and solvation energy (EPBSA). EMM was calculated by the sander module of AMBER (explicit solvent and ions were not included) with parm99.[43] EPBSA is composed of two types of contributions, electrostatic (EPB) and nonpolar (ESA). The electrostatic part was calculated with a numerical solver with the Poisson-Boltzmann method implemented in the PBSA program.[71] The nonpolar part depends on solvent accessible surface area, which was calculated by the molsurf program implemented in AMBER.[72] Conformational entropy was obtained using the nmode module of AMBER 9.0,[50,51] which performs normal mode analysis[73] to predict the conformational entropy. The program provides a total solute entropic term as a sum of translational, rotational and vibrational entropic contributions. All free energy terms were derived using each consecutive 20th snapshot.

Nudged elastic band (NEB) method

The NEB method was employed to investigate conformational change pathway for the transformation between the NMR and X-ray structures. The original NMR and X-ray structures wre energy minimized using standard methods with AMBER 10.0[74] and parm99.[43] The potential energy for the minimized NMR structure was −4087.4 kcal/mol, and for the X-ray structure −4000.1 kcal/mol. These structures were used as end points in NEB calculations.[75–77] The initial NEB pathway consisted of 16 NMR structures followed by 16 X-ray structures. 21 NEB trajectories were calculated using a simulated annealing protocol and varying the random number seed. The simulated annealing protocol (Table S1)[78] involved quickly heating the system to 1000 K, followed by slow cooling and finally quenched dynamics to remove any remaining kinetic energy from the system.

Targeted MD (TMD)

TMD was used as implemented in AMBER 9.0[50,51] to perform a forced conformational transition between the NMR and X-ray structures. The NMR structure was equilibrated using standard protocol[20] and then it was used as a starting point (reactant structure) and the equilibrated *E.c.* X-ray structure of H40 was used as end point (target structure). The reaction coordinate was defined as RMSD of internal loop (residues: 4–6, 13–15) between the instantaneous reactant structure and the fixed target (product) structure. Since the RMSD is dependent on group of atoms, which are used for the best-fit to the target structure, we run two TMD simulations with different initial setting. In the MD_TMD_1 simulation, modest positional restraints (0.01 kcal/mol.A2) were applied to terminal WC base pairs (residues: 1, 2, 8–11, 17, 18), which were simultaneously used for the best-fit. Simulation was carried out in twenty 1 ns long windows. In each window the molecule was forced to target RMSD, which gradually decreased (in each window about ~0.4 Å increment) (Figure S3). In the MD_TMD_2 simulation, the positional restraints were not used and the best-fit was carried out over the whole structure (*i.e.* over residues 1–18). This simulation included only ten 1 ns windows where target RMSD gradually decreased by ~0.4 Å increments (Figure S3). Both simulations were run in explicit solvent at 300 K using NPT conditions. Control simulations for both MD_TMD_1 and MD_TMD_2 were run with different random number seeds. The force constant was set to 0.1 kcal/mol.A2 in both simulations. Apart from tracking the conformational transition we employed the Weighted Histogram Analysis Method (WHAM)[79] (version 1.0) to estimate the free energy profile of the conversion. Note that targeted MD is substantially affected by the imposed path[80] such that any large-scale conformational changes like those in ref.[22,81] as well as in the present study should always be reviewed carefully and only viewed as crude estimates of the real transitions.

RESULTS

Standard MD simulations

Geometry of the ribosomal H40 relaxes in parm99 MD simulations—Definitely the most striking feature in the 350 ns MD_Ec_99 simulation was considerable opening of the structure due to widening of the major groove (Figure 1A right and Figure 4A).

In the course of the simulation the major groove width was monitored by two inter-phosphate distances (11P-4P and 12P-3P) (Figure 4A). In the time period from 0 to 30 ns, the major groove width oscillated around 9 Å (Figure 4B), then it rapidly increased up to ~16 Å and oscillated around this value until ~100 ns (Figure 4B). In the 100 to 300 ns time period the major groove width fluctuated around 20 Å (Figure 4B). The opening, which was also seen in other simulations of X-ray H40, was coupled with the disruption of the BPh interaction.[49] The two non-canonical base pairs of the internal loop showed instabilities during the simulation. Particularly, opening events of both H-bonds of the sheared A/G pair were detected (Table 2).

Disruption of this pair was seen in the 133–297 ns time period, during which the A6 base flipped out of the helix. Changes were also detected for the rH U/A base pair. An opening event and eventual disruption was seen for the U(N3)-A(N7) H-bond (Table 2). The cross-strand A5A14 stack exhibited fluctuations in the 0–100 ns time period; however, in the rest of the simulation it was essentially stable (Figure 5).

Larger changes were found for the A5A6 intrastrand stack. In the 0–100 ns time period the A5 base alternatively stacked between A6 and G13, and afterwards established stable stacking with G13 (Figure 5). The A15 bulge fluctuated outside the helix over the whole simulation. Its insertion into the stem was obstructed by A14, which was involved in cross-strand stacking. The MD_Ec_99 simulation was extended up to 350 ns. However, the canonical segment (residues 7–12) including the sheared A6/G13 base pair was disrupted at ~300 ns (Figure S4). Thus the simulated structure was ultimately lost and the 300 to 350 ns time period was not considered in our analyses.

The MD_Hm_99 and MD_Dr_99 simulations showed picture similar to the MD_Ec_99 simulation, *i.e.* marked opening of the major groove (Figure 4B), fluctuation of the A15 base outside the helix, instability of the two non-Watson-Crick base pairs (Table 2) and changes in base stacking (Figure 5). Hence full description of these simulations is given in Supporting Information. In the MD_Tt_99 simulation the major groove did not convert into a permanently open conformation in contrast to the previous simulations. The major groove width oscillated and it was stabilized by temporarily formed X-ray G13(N1)-A5(O2P) H-bond (Figure 2). Apart from this contact additional H-bond formed between G13(N2) and A5(O2P), which however oscillated in larger range compared to the G13(N1)-A5(O2P). This binding mode in which N2 and N1 of G bind to the same anionic oxygen of the phosphate group represents base phosphate interaction type 4BPh, which can be seen in the *E.c.* X-ray structure of H40 (Figure 2). In the course of the MD_Tt_99 simulation, the original 5BPh interaction thus alternated with the 4BPh interaction. Both the intrastrand and cross-strand stack fluctuated but replacement of the A5A6 stack by the A5G13 stack was not seen. Similarly to the previous simulations, instabilities of the two non-canonical base pairs were detected (Table 2). Details of this simulation are also indicated in Supporting Information.

MD simulations of the ribosomal H40 with mutated residues—The A5A14 cross-strand stack is likely key for maintaining the stability of the X-ray loop. Thus we carried out the MD_A5U simulation with an A5U mutation aimed at destabilizing this stack (Table 1).

The simulation showed expulsion of the uracil out of the stem and subsequent stacking of the sheared pair and the rH pair. This caused shortening of the helix by one base pair level and formation of a more compact and presumably more stable structure. This substitution apparently destabilizes the functional structure of the UAA/GAN internal loop. Two additional simulations, including the MD_A14U simulation with A14U substitution and the MD_A14G_U4C simulation with A14G and U4C substitutions (Table 1), did not show disturbance of the cross-strand stacking, although instabilities of substituted base pairs were detected (data not shown). In the later simulation we initially attempted to introduce a canonical G14C4 base pair but this base pairing was not stable in this context.

Elevated temperature and no-salt simulations of H40—Both simulations performed at 400 K led to disruption of the whole structure within the first 10 ns (data not shown). During the first 5 ns of the no-salt simulation the major groove rapidly widened to 18 Å and changes occurred in the internal loop. In particular, the sheared A/G pair and the rH pairs disrupted after 120 ns and 95 ns, respectively; however, the A5A14 cross-strand stack remained stable indicating this stack has considerable stability.

Geometry of the solution structure is stable in MD simulations—The MD_NMR_99 simulation of the NMR solution structure is significantly more stable than that of the X-ray structure (see low RMSD value in Table 1). The major groove remained as wide as in the experimental structure and the inter-phosphate distances oscillated around the starting values (Figure 4C). The sheared A/G and A/A pairs were stable. It must be admitted that the experimental structure was refined with the same force field, albeit using 500 K in vacuo annealing instead of using explicit solvent at 300 K.

The peculiar *c*WS A/U base pair that is incomplete in the NMR structure exhibits interesting behavior. Its presence conflicts with secondary structure predictions that would place a Watson-Crick base pair there. In addition, the Watson-Crick A–U base pair would be achievable from the incomplete *c*WS starting geometry by a modest rearrangement. This base pair may be essential to understanding the internal loop. The simulation reveals that the incomplete *c*WS base pair is in fact water-bridged, as its sugar-base A15(N1)-U4(O2´) interaction is mediated by water, a substate not apparent from the classification by Leontis and Westhof.[6] The bridging water molecules do not show anomalously long residency times[82] and exchange typically on the time scale of hundreds of picoseconds. Further development of the trajectory revealed two alternative geometries. In the 0–47 ns and 130–160 ns time periods, it was seen in the starting geometry, but it assumed standard Watson-Crick (*c*WW) conformation in the rest of the simulation. It never sampled geometry with a fully completed *c*WS base pair with direct A15(N1)-U4(O2´). Therefore, the simulation appears to be consistent with the unusual experimental structure, albeit perhaps subtly biased towards canonical pairing. Note that very small bias of the force field (in terms of free energy) would be sufficient to change the balance between these two substates if they are close in energy. Thus, the simulation behavior does not indicate any large imbalance of the force field.

In the MD_NMR_restr simulation, we imposed a direct A15(N1)-U4(O2´) H-bond in the *c*WS A/U base pair via a restraint. At 10 ns, the restraint was released after which the H-bond changed immediately to the water mediated bond. This is another indication of the overall (qualitative) correctness of the force field, which clearly eliminates the fully paired *c*WS structure, in line with the experiments. This reflects very good balance of the AMBER force field for stacking interactions and non-canonical RNA base pairing.[83,84] At 17 ns the base pair adopted standard *c*WW geometry, which was stable until the end of the simulation.

MD simulations of H40 with parmbsc0 are similar to the simulations with parm99—The 100 ns control simulations of the ribosomal H40 run with parmbsc0 force field[44] (Table 1) showed similar albeit reduced widening of the groove as the corresponding parm99 simulations (cf. Figure 4B). However, the A5A6 and A5A14 stacks were not disrupted in three out of four simulations (cf. Figure 4B top). Full details are given in Supporting Information. The control simulation of the solution structure with parmbsc0 (MD_NMR_bsc0, see Table 1) provided an identical picture as the MD_NMR_99 simulation (Supporting Information and Figure 4C).

MD simulations of H40 in excess of KCl—In the simulations carried out with excess of KCl (MD_Ec_K1-3, Table 1), instabilities in base pairing were mostly seen for the U(N3)-A(N7) H-bond of the rH pair in agreement with the parm99 Na$^+$ simulations (Table 2). However, the two stacks (A5A6 and A5A14) were stable similarly to the parmbsc0 force field simulations.

In the MD_Ec_K1 simulation, (KCl, Dang's parameters and parm99, see Table 1), the opening of the major groove was reduced by ~4 Å compared to the parm99 Na$^+$ simulation (Figure S5). In the MD_K2 simulation (KCl, Dang's parameters, parmbsc0) and in the MD_K3 simulation (KCl, Joung's parameters and parm99), the widening of the major groove coupled with disruption of the BPh G13(N1)-A5(O2P) contact was only seen during the first 18 and 30 ns, respectively, (Figure S5). After that we observed narrowing of the major groove and restoration of the X-ray BPh H-bond, in a form of the bifurcated G13(N1, N2)-A5(O2P) 4BPh interaction. This H-bond was seen in the MD_Tt_99 simulation (see above). In addition, in the MD_Ec_K3 simulation, the structure expelled the unpaired A5 base from the stem at 23 ns, which was accompanied by subsequent stacking of the sheared pair and the rH pair. A similar event has been detected in the net-neutralizing Na$^+$ simulation with A5U mutation (see above).

The MD_NMR_K simulation provided a picture close to identical to the MD_NMR_99 simulation. Similarly to the MD_NMR_bsc0, subtle compaction of the major groove by ~1 Å compared to the parm99 result was detected (Figure S5). The sheared A/G and A/A base pairs were stable. The A/U base pair showed the starting geometry till 70 ns while after 70 ns it converted to the canonical *c*WW conformation.

MD simulation of the H68 X-ray structure converts to the solution structure—In the simulation of the ribosomal H68 (MD_H68), the widening of the major groove was also detected (Figure 1B right and Figure 6A) similarly to the H40 simulations (see above and Figure 4).

The inter-phosphate distances quickly increased up to 20 Å and then fluctuated around this value (Figure 6B). At the beginning of the simulation, single H-bond (A(N6)-G(N7)) formed between A6 and G13 bases and it was stable till the end of the simulation. This pairing does not correspond to any established base pair family.[6] Around 30 ns the middle stacking adenines A14 and A5 formed a sheared pair and the A15 and U4 bases formed a *c*WS pair with one direct bond (A(N6)-U(O2)) and one water-mediated bond (A(N1)-U(O2´)) (Figure 6C). Both these pairs occur in the solution structure (Figure 1C). The sheared pair was stable by the end of the simulation, while the *c*WS pair alternated with the *c*WW geometry, similarly to the simulations of the solution structure (see above). The *c*WS geometry was seen in the time periods of 30–45 and 52–57 ns, and the *c*WW geometry in the time periods of 45–52 and 57–100 ns. Importantly, the final transformed H68 geometry is very close to the solution structure, with RMSd of only 1.6 Å (Figure S6). The solution structure A6/G13 base pair was, however, not formed. This is due to the fact that G13 in the X-ray structure is in unusual *syn* conformation. The simulation was not long enough to flip the G13 to the *anti*

conformation, which would lead to entire agreement with the NMR structure. In fact, it cannot be ruled out that the initial G13 *syn* conformation is an experimental refinement error. In the RNase P RNA X-ray structure, the equivalent guanine is indeed in *anti* orientation. We have attempted three 150 ns simulations (two with parm99 and one with parmbsc0, data not shown) where the G13 was initially flipped to *anti*. These simulations, however, did not reveal larger transitions towards the solution structure. It may reflect both stabilizing effect of the initial G13 *anti* conformation as well as sampling limitations. Despite this, we still consider the above-analyzed MD_H68 simulation as a solid piece of evidence of the tendency of the X-ray structure to convert spontaneously to the solution structure, which is also supported by the free energy computations (see below).

LES simulations of the ribosomal H40

The aim of the LES simulations was to achieve transition from the X-ray H40 structure to the NMR structure, which did not occur in multiple standard simulations. The solution structure internal loop arrangement was sampled only in two LES simulations, which are described below. Other LES simulations are presented in Supporting Information.

In the first 12 ns of the LES_Ec simulation (Table 1) the original X-ray base pairing of the internal loop was disrupted and the major groove width increased to ~20 Å. In the 12–40 ns time period, the "multiplied" bases of the internal loop (nucleotides involved in LES) adopted various rapidly changing arrangements and did not form stable base pairs. At 41 ns, G13 and A6, A14 and A5, and A15 and U4 became coplanar (Figure 7), which markedly resembled the solution structure (Figure 1C).

However, the LES bases failed to establish stable pairs. This "NMR-like" arrangement was maintained in the rest of the simulation except for several short disruptions. We started standard MD simulation from this "NMR-like" geometry (see Table 1). After 40 ns, single H-bonds formed between G13 and A6, and between A14 and A5, however, the A15 and U4 bases were expelled from the stem. After 70 ns the internal loop was disrupted resulting in disturbing of the whole structure (data not shown).

During the first 5 ns of the LES_Dr simulation (Table 1) internal loop base pairs were disrupted and the width of the major groove increased to ~20 Å, similar to the LES_Ec simulation. At 7 ns, the bulging A15 base flipped into the stem and during the 10–12 ns time period the LES bases formed an arrangement where G13 and A6, A14 and A5, and A15 and U4 were coplanar, like in the LES_Ec simulation (Figure 7). In the 12–40 ns time period the LES bases sampled various unstable arrangements (data not shown). Thus, in summary, LES may show some signs of the transition but no complete transition was achieved.

Free energy calculations

Figure 8 summarizes the MM-PBSA free energy calculations for the parm99 simulations of the UAA/GAN internal loop X-ray and NMR structures.

For H40 simulations, the initial rapid expansion of the major groove (Figure 4) is accompanied by free energy drop by about 3–6 kcal/mol (Figure 8). Total free energy time course of the H68 loop revealed two marked decreases of free energy. The first one (by ~10 kcal/mol) can be seen after the first 5 ns and it corresponds to the rapid expansion of the major groove (Figure 6B), similarly to the H40 total free energy time courses. The second one (by ~12 kcal/mol) can be seen around 30 ns and it corresponds to the transition of the 2D structure (*i.e.* formation of the sheared A/A and *c*WS A/U base pairs, see above).

Comparing the averaged total free energies for the 1–100 ns time periods, the NMR structure is predicted to be more stable than the X-ray H40 structure by about 17 kcal/mol

(*E.c.* H40 simulation), 19 kcal/mol (*H.m.* H40), 11 kcal/mol (*D.r.* H40) and 18 kcal/mol (*T.t.* H40). The solute entropic term favors the X-ray structure by ~4–5 kcal/mol which is consistent with the expectation that the X-ray structure is intrinsically less rigid in isolation. The H68 after the transition (since ~30 ns) is on average by ~5 kcal/mol more stable than the NMR structure. With exclusion of the entropic term the averaged free energies of final H68 and NMR structures would be identical. In summary, the free energy computations give a clear hint that the NMR structure of the UAA/GAN internal loop is indeed intrinsically more stable than the X-ray H40 and H68 structures, albeit the energy difference is probably over-estimated as usual with this kind of highly approximate free energy calculations. Note that the free energy computations should in no case be taken quantitatively, despite abundant such attempts in contemporary literature.

We have further tested an alternative potential of mean force method of free energy computations which was recently used for 16S ribosomal decoding bases 1492 and 1493.[85] We could not use it for the transition between X-ray and solution structures, as we did not see a full transition. However, we used the method to investigate the free energy basin around the H40 X-ray structure while comparing sampling with the parmbsc0 and parm99 force fields. The results are in full detail in Supporting Information.

NEB and TMD reveal a possible pathway for the transformation between the NMR and X-ray H40 structures

The NEB calculations provide a 32 image pathway where the end points are fixed conformations, with the first image being the energy-minimized NMR structure (Figure 9A) and the final structure being the X-ray structure.

Observations of all pathways reveal they pass through similar intermediates. Potential energy profiles for the pathways also are similar (Figure S7). The pathways involve a particular order of structural events. First, A14 breaks its pairing with A5 and moves away from being stacked with A6 and A15 (Figure 9B). A15 slides rapidly out of the helix to become the bulged A15 observed in the X-ray structure. A14 moves to pair with U4 as soon as A15 is out of the way and stacks with G16 (Figure 9C). A5 is left unpaired, and loses its stacking interaction with U4 to end up hovering over the pairing region of the U4-A14 pair. A5 also shifts further from G13 to become stacked with A6 (Figure 9D) and forms a hydrogen bond with the backbone of the opposite strand as it is no longer base paired.

The potential energy profile is a plot of the potential energy for each of the 32 images along the pathway (Figure S7). The potential energy difference between the product and reactant structure, 87 kcal/mol, is large compared to the above-noted free energy differences. Note that the NEB potential energy does not include the entropic effect of conformational freedom, and it would require a sampling method such as umbrella sampling to relate the NEB potential energy to free energy.

The NEB calculations were run from the NMR structure, which turned out to have the lowest potential energy, to the high energy X-ray structure. All pathways start with a slight increase in potential energy to about −4060 kcal/mol and remain there until the last few images, where there is a sudden increase in energy to the X-ray structure. There is variation in the slight peaks and valleys throughout the region where the potential energy is about −4060 kcal/mol, but no clear or consistent transition states or intermediates are observed for the 21 NEB pathways. In summary, the limited variation between NEB trials suggests that the conformational change occurs using a single predominant pathway (Figure 9). There is some minor energetic variation as the balance of molecular forces is slightly different for the different pathways as indicated by the occasional and inconsistent energy minima and maxima in the potential energy profile.

The conformational transitions between NMR and X-ray structures obtained from the TMD simulations are basically identical to the NEB data (Figure 9). In the 20 ns long MD_TMD_1 simulation, the conversion started at 11 ns (Figure S3) with disruption of the sheared A14/A5 pair, similar to the NEB result, followed by disruption of the cWS A15/U4 pair. Then, the A15 base bulged out of the helix and the A14 base moved by one base in the strand and created a pair with U4. The A5 base remained unpaired and formed a stack first with G13 and then it moved and stacked with A6. In the MD_TMD_2 simulation, the transition started directly with disruption of the A15/U4 base pair. Otherwise the transition was identical to the MD_TMD_1 simulation. Additional control simulations revealed identical pictures as the MD_TMD_1 and MD_TMD_2 simulations (data not shown). The energy profiles extracted from the MD_TMD_1 and MD_TMD_2 simulations indicate that the X-ray geometry of the H40 UAA/GAN internal loop has about 30 kcal/mol higher free energy than the NMR structure (Figure S3). However, the extracted energies must be taken with care because step changes in the RMSD profiles (mainly in the MD_TMD_2 simulation) can be seen (Figure S3). This indicates insufficient overlap between windows in these regions, which may bias the potential.[86] The trend in the free energy is, however, entirely consistent with the MM-PBSA and NEB data.

DISCUSSION AND CONCLUSIONS

We employed MD methods to investigate the highly conserved UAA/GAN internal loop of 23S rRNA H40 (Figure 1)[5] which also occurs in six other 23S rRNA helices and in other RNAs.[5] It consists of rH U4/A14 and sheared A6/G13 base pairs interconnected via an unpaired A5 base by two stacks, the A5A14 cross-strand stack and the A5A6 intrastrand stack, as well as a bulged N15 base (Figure 1). The UAA/GAN internal loop in the ribosomal structure has a narrow major groove and wide minor groove (Figure 1). This functional conformation is involved in tertiary contacts with surrounding ribosomal elements that drastically rearrange the base pairing and stacking of the loop compared to its solution structure.[13] These contacts include involvement of conserved adenines in A-minor interactions (Figure 3).

The solution structure contains three non-canonical base pairs (the A6/G13 sheared pair, the A14/A5 sheared pair, and the cWS A15/U4 pair) with no unpaired base. The cWS A15/U4 base pair observed in the NMR structure is surprising because secondary structure predictions posit a canonical A–U base pair at this position. In addition, the cWS A15/U4 pair is incompletely paired.

We studied the H40 UAA/GAN internal loop taken from available X-ray bacterial and archaeal 50S subunits along with the NMR solution structure[13] (Table 1 and Figure 1, Figures S1 and S2). Furthermore, we investigated the less frequent UAA/GAN geometry from 23S rRNA H68 which adopts yet another (alternative) 2D/3D arrangement (Figure 1). Therefore, the UAA/GAN internal loop is an RNA molecular switch that has functional geometry in folded RNAs that differs from its optimal geometry in isolation.

H40 UAA/GAN basic simulations

Unrestrained explicit solvent MD simulations revealed relaxation of the X-ray H40 UAA/GAN internal loop on a scale of tens of nanoseconds. In particular, considerable expansion of the major groove width from the original value of 6–8 Å up to 16–22 Å was detected, coupled with disruption of the X-ray base-phosphate interaction across the major groove (Figure 2 and Figure 4). Further, widening of the major groove was accompanied by replacement of the X-ray A5A6 intrastrand stack by a new A5G13 stack (Figure 5). The newly formed stack was probably more compatible with the wide major groove than the original X-ray one. The relaxed X-ray geometry (Figure 4A) partially resembles the solution

structure with its wide major groove with a width of ~17 Å (Figure 1 and Figure 4). However, the open conformation of the solution structure may be a result of the NMR refinement procedure, which utilizes NMR-restrained molecular dynamics and energy minimization[13] because the NMR experiment does not provide precise structural information about the sugar phosphate backbone. Further, all standard simulations of the X-ray H40 UAA/GAN exhibited opening events and fluctuations of base pairs in the internal loop (Table 2). Irreversible disruptions of H-bonds were also detected in these pairs (Table 2). The increased dynamics of the sheared A/G pairs can be because of the large number of potential hydrogen bonds which cannot all be made simultaneously, which was proposed in an NMR study of GNRA hairpin loops.[87] Despite such dynamics and changes, the overall X-ray H40 secondary structure of the internal loop was maintained in the simulations, *i.e.* the base pairing does not spontaneously rearrange towards the solution structure. Occasionally, in some long simulations, the structure was ultimately disrupted.

The cross-strand A5A14 stack shows only modest fluctuations in our simulations in comparison with the intrastrand A5A6 stack and may be one of the key stabilizing elements of the X-ray secondary structure. It has been suggested[13] that the cross-strand stack allows base pairing of A14 and U4 and additionally compensates for H-bonds lost between A15 and U4 and between A14 and A5. We attempted to disrupt the cross-strand stack in the simulations with several mutations (A5U, A14U, and A14G together with U4C, Table 1). The first substitution led to expulsion of the U5 from the stack and subsequent substantial rearrangements, hinting at the key role of A5 not only for the tertiary interactions but also for the stability of the functional X-ray structure of the UAA/GAN loop. The other substitutions had an inconclusive impact on the simulations. Likewise, simulations at elevated temperature and assuming no-salt condition did not provide any insights into the properties of the UAA/GAN internal loop.

H68 simulations

In the standard simulation of the ribosomal H68 UAA/GAA loop, we observed a large spontaneous transition clearly towards the solution structure (Figure S6), except that the A6 and G13 bases did not form any classified base pair, which probably relates with initial *syn* orientation of the G13 base. The simulation was not able to overcome this initial *syn* orientation.

Solution structure

Simulations of the solution UAA/GAN loop structure were stable (Table 1). The wide major groove remained unchanged and the base pairs of the loop were stable except for fluctuations of the A15/U4 base pair. In particular, the experimental *c*WS A15/U4 base pair is stabilized by only one direct H-bond, despite that two direct bonds are assumed by standard classification.[6] The simulations reveal that there is an additional stabilizing interaction in this base pair, namely a sugar-base water-bridge. In the simulations, this partially paired base pair alternates with the canonical (*c*WW) geometry expected from thermodynamic considerations,[88,89] but not indicated by the NMR experiment.[13] On the other hand, a fully paired *c*WS A15/U4 base pair never formed in the simulations and was immediately disrupted even when initially imposed by restraints, entirely in agreement with NMR. Thus, simulations of the NMR structure of the UAA/GAN internal loop indicate a satisfactory performance of the simulation force field, albeit the balance of the simulation might be subtly shifted towards formation of the canonical A15-U4 base pair. We suggest that the solution structure of the UAA/GAN internal loop represents an interesting test molecule for verification of simulation methods and force fields.

Structural plasticity of the UAA/GAA internal loop

To obtain additional insights we applied a range of auxiliary methods that can enhance the capabilities of standard simulations. Note, however, that all these methods necessarily introduce additional approximations and are thus inherently less reliable than standard simulations. The LES technique was applied to enhance sampling of bases in the X-ray H40 loop. All the LES simulations revealed widening of the major groove and disruption of the internal loop. The internal loop adopted various arrangements where bases involved in the LES region mutually stacked or formed temporary contacts. However, LES was not robust enough to converge into a stable prevalent conformation. Occasionally, the bases formed a secondary structure arrangement similar to the solution conformation (Figure 1 and Figure 7), although no stable base pairing was established.

Total free energies were extracted from the standard simulations utilizing the MM-PBSA method. Free energy time courses along the UAA/GAN H40 X-ray loop trajectories showed an ~3–6 kcal/mol free energy improvement during the significant 10–15 Å increases in the major groove width (see Figure 4 and Figure 8). Furthermore, MM-PBSA data predicted the free energy of the X-ray H40 UAA/GAN internal loop structure to be ~10–20 kcal/mol less favorable compared to the NMR structure. A previous experimental study[13] predicted the internal loop of the NMR structure to be favorable by ~5 kcal/mol when compared to the X-ray H40 functional structure. This estimate was based on an experimental measurement of free energy of the internal loop in the solution structure and a prediction of free energy of the X-ray ribosomal internal loop utilizing a nearest neighbor model.[89,90] Thus, the free energy calculations identify the correct trend but do not reach quantitative accuracy.

The conformational transition between NMR and H40 X-ray structures was investigated by the NEB and TMD methods. Results are mutually consistent. The conversion could start with breaking the sheared A14/A5 pair. This is in accord with the NMR study,[13] which suggested structural dynamics of the A5 base and proposed that the dynamics may provide a pathway for conformational conversion. The transition continues with disruption of the *c*WS A15/U4 pair and bulging out of A15 base, and eventually formation of rH U5/A14 pair, the A5A6 stack and the cross-strand A5A14 stack (Figure 9). Calculations thus predict a likely mechanism for rearrangement of the solution conformation into the functional "ribosomal" X-ray geometry. In the ribosome, the conversion could be induced by an adjacent rRNA (hairpin structure between H39 and H40) and ribosomal protein L20, which binds *E.c.* 23S rRNA at an early stage of ribosomal assembly.[91] The presence of two single H-bond pairs, sheared A/A and *c*WS A/U pair, in the internal loop of the solution structure suggests that the loop structure may be internally weak and easily disrupted by external forces. The weakness of the pairing in the solution structure, which is the global minimum of the UAA/GAN internal loop, may be one of the important pre-requisites for its smooth transition to the functionally important substate. This may contribute to determination of the consensus sequence of the UAA/GAN internal loop whose sequence signature is otherwise primarily determined by the X-ray architecture and the tertiary interactions it is involved in.

Force field choice

We see a modest difference between the parm99 and parmbc0 force fields, but this does not affect any key conclusions of the paper. Overall, the control MD simulations carried out with parmbsc0 force field[44] provided a similar picture as the simulations run with parm99[43] (400 ns of comparable trajectories for the H40 initial structure). Nevertheless, in simulations of the ribosomal H40, the major groove width was reduced by 2–4 Å when using parmbsc0 compared to parm99 simulations (Figure 4). In the simulation of the solution structure, the major groove width was reduced only by 1 Å. There are two competing α/γ backbone substates, canonical geometry and t/t conformation (two established A-RNA families 20 and

24).[92] During 100 ns portion of the *E.c.* parm99 simulation we detected 17 % population of nucleotides in the α/γ t/t conformation while the t/t flips are reversible. This is comparable with 10–15 % population reported in the MD study of 16S rRNA H44 and canonical A-RNA. [36] In contrast, the γ-trans states are fully suppressed using the parmbsc0 force field. The suppression of the γ-trans substates allows the major groove narrowing. As the α/γ t/t substate occurs occasionally also in experimental structures and is compatible with overall A-RNA topology, both force fields have satisfactory performance, despite the above-noted difference. The actual propensity of A-RNA to populate the α/γ t/t substates is likely in between the parm99 and parmbsc0 propensities while the overall difference between the two force fields is for the present RNA system small.

In the excess salt KCl simulations, the major groove width was also reduced compared to the net-neutralizing Na^+ simulations (by ~2–4 Å for the H40 structure). In addition, in some KCl simulations the major groove after the initial widening returned back to closed X-ray conformation stabilized by the BPh interaction. We suggest that the results are explained by better screening of phosphates with higher ionic strength which allows their closer approach across the groove. Thus the stability of the functional H40 conformation may be affected by ionic conditions or other interactions reducing the inter-phosphate repulsion.

Which of the force field options is better? The answer is not unambiguous. The simulations clearly show a tendency of the major groove to widen for the UAA/GAA internal loop when it is taken out of its ribosomal context. We have earlier reported widening of the major groove width for the 5S rRNA Loop E in parm99 Na^+ simulations. Loop E is an internal loop with seven consecutive non-canonical RNA base pairs. Shorter, ~10 ns, trajectories gave increases in the P-P distances for the 5S rRNA Loop E system by only a few Angstroms compared to the X-ray structure.[82] In contrast to H40, however, loop E is in its global minimum in its X-ray structure. All these results may give an impression that simulations tend to overshoot the RNA major groove width. This effect is larger with parm99 than with parmbsc0 and can be reduced by using excess salt condition instead of net-neutralization. Our very recent reference simulations on canonical A-RNA,[93] however, show that the picture is more complex. These reference simulations also usually show tendency for widening of the major groove in simulations compared to the X-ray structures, which is larger with the parm99 than with the parmbsc0 force field. These simulations, however, also show that the A-RNA major groove width and its relaxation depend on the base sequence and are different for different X-ray structures. Therefore, there is no unambiguous experimental target value of the major groove width, as the experimental values depend on the sequence and crystallization conditions. In some cases parm99 with net-neutralization can remain closest to the experimental structures. Therefore, the right interpretation is that the RNA major groove width is sensitive to the sequence, environment and molecular interactions and the simulations reflect this groove width plasticity. We suggest that all force field options utilized in the present study are justified for RNA simulations. Nevertheless, RNA simulations can perhaps indirectly profit from the parmbsc0 force field choice, since its tendency to keep the major groove more closed may reduce likelihood of irreversible structural disruptions during some large temporary major groove width fluctuation events. On the other hand, the parmbsc0 force field appears to somewhat rigidify the simulated structures compared to parm99, as evidenced by less frequent changes of the adenine stacks (see above). We cannot tell, however, whether this behavior is improvement or not compared with the parm99 force field.

The H40 UAA/GAA simulations further show some local instabilities and some long simulations result in entire loss of the X-ray structure topology for the UAA/GAA internal loop without approaching closer to the solution structure. Development of the simulations on a much longer time scale is thus uncertain. The perturbation may reflect either the genuine

internal instability of the functional (ribosomal) structure of UAA/GAA when considered in isolation, or it may be accumulation of force field imbalances in our long simulations. Most likely, both factors contribute.

Concluding remarks

The simulations provide atomistic characterization of the structural dynamics of UAA/GAA internal loop in three distinct experimental topologies. The simulations appear to be consistent with experimental data and give new insights. Our study is probably the first simulation study of a recurrent RNA non-Watson-Crick element that is not autonomous, *i.e.*, it folds only in specific contexts. The H40 simulations do not spontaneously transform to the solution (ground state) structure and such transition is probably beyond the limits of contemporary computational chemistry. However, almost complete transformation was seen for the alternative H68 X-ray structure. Methods like TMD of NEB can achieve transformation between the H40 and solution UAA/GAA topologies, however, they impose artificially selected path and require a priori knowledge of both starting and final structures. Free energy computations can provide some very crude estimates of the free energy trends but are probably far from reaching even qualitative accuracy.

The results suggest that the H40 and H68 internal loops are under stress due to tertiary and quaternary interactions, and that H68 can relax to its conformation in isolation much faster than H40 if the interactions with its surroundings are relieved or altered. Thus the MD results suggest that the different structures induced by tertiary and quaternary interactions may also have implications for temporal control of events. Both the MD and NMR results indicate there is no significant population of higher free energy structures in isolated RNA. This suggests that approach of other parts of the ribosomal RNA or of protein induce a conformational change rather than trapping a minor species. This type of conformational switch may be important for assembly and/or movement in molecular machines such as the ribosome. We demonstrate that despite the above-explained limitations, modern MD-based computations can complement experimental techniques and provide insights into the role of molecular interactions in shaping RNA building blocks.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science 2000;289:905–920. [PubMed: 10937989]

2. Harms JM, Schluenzen F, Zarivach R, Bashan A, Gat S, Agmon I, Bartels H, Franceschi F, Yonath A. High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. Cell 2001;107:679–688. [PubMed: 11733066]

3. Wimberly BT, Brodersen DE, Clemons WMJ, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V. Structure of the 30S ribosomal subunit. Nature 2000;407:327–339. [PubMed: 11014182]

4. Leontis NB, Westhof E. Analysis of RNA motifs. Curr. Opin. Struct. Biol 2003;13:300–308. [PubMed: 12831880]

5. Lee JC, Gutell RR, Russell R. The UAA/GAN internal loop motif: A new RNA structural element that forms a cross-strand AAA stack and long-range tertiary interactions. J. Mol. Biol 2006;360:978–988. [PubMed: 16828489]

6. Leontis NB, Stombaugh J, Westhof E. The non-Watson-Crick base pairs and their associated isostericity matrices. Nucleic Acids Res 2002;30:3497–3531. [PubMed: 12177293]

7. Sarver M, Zirbel CL, Stombaugh J, Mokdad A, Leontis NB. FR3D: finding local and composite recurrent structural motifs in RNA 3D structures. J. Math. Biol 2008;56:215–252. [PubMed: 17694311]

8. Jaeger L, Verzemnieks EJ, Geary C. The UA_handle: a versatile submotif in stable RNA architectures. Nucleic Acids Res 2009;37:215–230. [PubMed: 19036788]

9. Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, Cate JH. Structures of the bacterial ribosome at 3.5 Å resolution. Science 2005;310:827–834. [PubMed: 16272117]

10. Kazantsev AV, Krivenko AA, Harrington DJ, Holbrook SR, Adams PD, Pace NR. Crystal structure of a bacterial ribonuclease P RNA. Proc. Natl. Acad. Sci. U. S. A 2005;102:13392–13397. [PubMed: 16157868]

11. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. Proc. Natl. Acad. Sci. U. S. A 2001;98:4899–4903. [PubMed: 11296253]

12. Guillier M, Allemand F, Raibaud S, Dardel F, Springer M, Chiaruttini C. Translational feedback regulation of the gene for L35 in Escherichia coli requires binding of ribosomal protein L20 to two sites in its leader mRNA: A possible case of ribosomal RNA-messenger RNA molecular mimicry. RNA 2002;8:878–889. [PubMed: 12166643]

13. Shankar N, Kennedy SD, Chen G, Krugh TR, Turner DH. The NMR structure of an internal loop from 23S ribosomal RNA differs from its structure in crystals of 50S ribosomal subunits. Biochemistry 2006;45:11776–11789. [PubMed: 17002278]

14. Auffinger P, Bielecki L, Westhof E. Symmetric K^+ and Mg^{2+} ion-binding sites in the 5S rRNA loop E inferred from molecular dynamics simulations. J. Mol. Biol 2004;335:555–571. [PubMed: 14672663]

15. Auffinger P, Hashem Y. Nucleic acid solvation: from outside to insight. Curr. Opin. Struct. Biol 2007;17:325–333. [PubMed: 17574833]

16. Cojocaru V, Klement R, Jovin TM. 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. [PubMed: 15956103]

17. Deng NJ, Cieplak P. Molecular dynamics and free energy study of the conformational equilibria in the UUUU RNA hairpin. J. Chem. Theory Comput 2007;3:1435–1450.

18. Krasovska MV, Sefcikova J, Spackova N, Sponer J, Walter NG. Structural dynamics of precursor and product of the RNA enzyme from the hepatitis delta virus as revealed by molecular dynamics simulations. J. Mol. Biol 2005;351:731–748. [PubMed: 16045932]

19. Li W, Sengupta J, Rath BK, Frank J. Functional conformations of the L11-ribosomal RNA complex revealed by correlative analysis of cryo-EM and molecular dynamics simulations. RNA 2006;12:1240–1253. [PubMed: 16682558]

20. Reblova K, Fadrna E, Sarzynska J, Kulinski T, Kulhanek P, Ennifar E, Koca J, Sponer J. Conformations of flanking bases in HIV-1 RNA DIS kissing complexes studied by molecular dynamics. Biophys. J 2007;93:3932–3949. [PubMed: 17704156]

21. Romanowska J, Setny P, Trylska J. Molecular dynamics study of the ribosomal A-site. J. Phys. Chem. B 2008;112:15227–15243. [PubMed: 18973356]

22. Sanbonmatsu KY, Joseph S, Tung CS. Simulating movement of tRNA into the ribosome during decoding. Proc. Natl. Acad. Sci. U. S. A 2005;102:15854–15859. [PubMed: 16249344]

23. Vaiana AC, Westhof E, Auffinger P. A molecular dynamics simulation study of an aminoglycoside/A-site RNA complex: conformational and hydration patterns. Biochimie 2006;88:1061–1073. [PubMed: 16824662]

24. Kopitz H, Zivkovic A, Engels JW, Gohlke H. Determinants of the unexpected stability of RNA fluorobenzene self pairs. Chembiochem 2008;9:2619–2622. [PubMed: 18823057]

25. Mazier S, Genest D. Insight into the intrinsic flexibility of the SL1 stem-loop from genomic RNA of HIV-1 as probed by molecular dynamics simulation. Biopolymers 2008;89:187–196. [PubMed: 18008323]

26. Mokdad A, Krasovska MV, Sponer J, Leontis NB. Structural and evolutionary classification of G/ U wobble basepairs in the ribosome. Nucleic Acids Res 2006;34:1326–1341. [PubMed: 16522645]

27. Rhodes MM, Reblova K, Sponer J, Walter NG. Trapped water molecules are essential to structural dynamics and function of a ribozyme. Proc. Natl. Acad. Sci. U. S. A 2006;103:13380–13385. [PubMed: 16938834]

28. Schneider C, Brandl M, Suhnel J. Molecular dynamics simulation reveals conformational switching of water-mediated uracil-cytosine base-pairs in an RNA duplex. J. Mol. Biol 2001;305:659–667. [PubMed: 11162082]

29. Villa A, Wöhnert J, Stock G. Molecular dynamics simulation study of the binding of purine bases to the aptamer domain of the guanine sensing riboswitch. Nucleic Acids Res 2009;37:4774–4786. [PubMed: 19515936]

30. Zhuang Z, Jaeger L, Shea JE. Probing the structural hierarchy and energy landscape of an RNA T-loop hairpin. Nucleic Acids Res 2007;35:6995–7002. [PubMed: 17940098]

31. Razga F, Koca J, Sponer J, Leontis NB. 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. [PubMed: 15722438]

32. Spackova N, Sponer J. Molecular dynamics simulations of sarcin-ricin rRNA motif. Nucleic Acids Res 2006;34:697–708. [PubMed: 16456030]

33. Reblova K, Razga F, Li W, Gao H, Frank J, Šponer J. Dynamics of the Base of Ribosomal A-site Finger Revealed by Molecular Dynamics Simulations and Cryo-EM. Nucleic Acids Res. 2009 doi: 10.1093/nar/gkp1057.

34. 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. [PubMed: 19804732]

35. 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. [PubMed: 17553840]

36. Reblova K, Lankas F, Razga F, Krasovska MV, Koca J, Sponer J. Structure, dynamics, and elasticity of free 16S rRNA helix 44 studied by molecular dynamics simulations. Biopolymers 2006;82:504–520. [PubMed: 16538608]

37. Cheatham TE III, Young MA. Molecular dynamics simulation of nucleic acids: successes, limitation, and promise. Biopolymers 2001;56:232–256. [PubMed: 11754338]

38. McDowell SE, Spackova N, Sponer J, Walter NG. Molecular dynamics simulations of RNA: An in silico single molecule approach. Biopolymers 2007;85:169–184. [PubMed: 17080418]

39. Orozco M, Perez A, Noy A, Luque FJ. Theoretical methods for the simulation of nucleic acids. Chem. Soc. Rev 2003;32:350–364. [PubMed: 14671790]

40. Sponer, J.; Lankas, F. Challenges and Advances in Computational Chemistry and Physics: Computational Studies of RNA and DNA. Dordrecht, Netherlands: Springer; 2006.

41. MacKerell AD, Nilsson L. Molecular dynamics simulations of nucleic acid-protein complexes. Curr. Opin. Struct. Biol 2008;18:194–199. [PubMed: 18281210]

42. Orozco M, Noy A, Perez A. Recent advances in the study of nucleic acid flexibility by molecular dynamics. Curr. Opin. Struct. Biol 2008;18:185–193. [PubMed: 18304803]

43. Wang J, Cieplak P, Kollman PA. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? J. Comput. Chem 2000;21:1049–1074.

44. Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE III, Laughton CA, Orozco M. Refinement of the amber force field for nucleic acids. Improving the description of α/γ conformers. Biophys. J 2007;92:3817–3829. [PubMed: 17351000]

45. Klein DJ, Moore PB, Steitz TA. The Roles of Ribosomal Proteins in the Structure, Assembly and Evolution of the Large Ribosomal Subunit. J. Mol. Biol 2004;340:141–177. [PubMed: 15184028]

46. Selmer M, Dunham CM, Murphy FV, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. Structure of the 70S ribosome complexed with mRNA and tRNA. Science 2006;313:1935–1942. [PubMed: 16959973]

47. Gao YG, Selmer M, Dunham CM, Weixlbaumer A, Kelley AC, Ramakrishnan V. The Structure of the Ribosome with Elongation Factor G Trapped in the Posttranslocational State. Science 2009;326:694–699. [PubMed: 19833919]

48. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy FV, Weir JR, Ramakrishnan V. The Crystal Structure of the Ribosome Bound to EF-Tu and Aminoacyl-tRNA. Science 2009;326:688–694. [PubMed: 19833920]

49. Zirbel CL, Sponer JE, Sponer J, Stombaugh J, Leontis NB. Classification and energetics of the base-phosphate interactions in RNA. Nucleic Acids Res 2009;37:4898–4918. [PubMed: 19528080]

50. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE III, DeBolt S. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecule. Comput. Phys. Commun 1995;91:1–41.

51. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Merz, KM.; Pearlman, DA.; Crowley, M.; Walker, RC.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, KF.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, DH.; Schafmeister, C.; Ross, WS.; Kollman, PA. AMBER 9. San Francisco: University of California; 2006.

52. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. J. Am. Chem. Soc 1995;117:5179–5197.

53. Aqvist J. Ion water interaction potentials derived from free-energy perturbation simulations. J. Phys. Chem 1990;94:8021–8024.

54. Joung IS, Cheatham TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. J. Phys. Chem. B 2008;112:9020–9041. [PubMed: 18593145]

55. Dang LX, Kollman PA. Free energy of association of the K+:18-crown-6 complex in water: a new molecular dynamics study. J. Phys. Chem 1995;99:55–58.

56. Smith DE, Dang LX. Computer simulations of NaCl association in polarizable water. J. Chem. Phys 1994;100:3757–3766.

57. Auffinger P, Cheatham TE, Vaiana AC. Spontaneous formation of KCl aggregates in biomolecular simulations: A force field issue? J. Chem. Theory Comput 2007;3:1851–1859.

58. Berendsen HJC, Grigera JR, Straatsma TP. The missing term in effective pair potentials. J. Phys. Chem 1987;91:6269–6271.

59. Humphrey W, Dalke A, Schulten K. VMD - Visual Molecular Dynamics. J. Mol. Graph. Model 1996;14:33–38.

60. Zhou R, Berne BJ, Germain R. The free energy landscape for beta hairpin folding in explicit water. Proc. Natl. Acad. Sci. USA 2001;18:14931–14936. [PubMed: 11752441]

61. Csaszar K, Spackova N, Stefl R, Sponer J, Leontis NB. Molecular dynamics of the frame-shifting pseudoknot from beet western yellows virus: The role of non- Watson-Crick base-pairing, ordered hydration, cation binding and base mutations on stability and unfolding. J. Mol. Biol 2001;313:1073–1091. [PubMed: 11700064]

62. Darden T, Pearlman D, Pedersen L. Ionic charging free energies: spherical versus periodic boundary condition. J. Phys. Chem 1998;109:10921–10935.

63. Elber R, Karplus M. Enhanced sampling in molecular dynamics: use of the time dependent hartree approximation for a simulation of carbon monoxide diffusion through myoglobin. J. Am. Chem. Soc 1990;112:9161–9175.

64. Roitberg A, Elber R. Modeling side chains in peptide and proteins: application of the locally enhanced sampling and the simulated annealing methods to find minimum energy conformations. J. Chem. Phys 1991;95:9277–9287.

65. Simmerling C, Elber R. Hydrophobic "collapse" in a cyclic hexapeptide: Computer simulations of CHDLFC and CAAAAC in water. J. Am. Chem. Soc 1994;116:2534–2547.

66. Fadrna E, Spackova N, Stefl R, Koca J, Cheatham TE, Sponer J. Molecular dynamics simulations of guanine quadruplex loops: Advances and force field limitations. Biophys. J 2004;87:227–242. [PubMed: 15240460]

67. Simmerling C, Miller JL, Kollman PA. Combined locally enhanced sampling and Particle Mesh Ewald as a strategy to locate the experimental structure of a nonhelical nucleic acid. J. Am. Chem. Soc 1998;120:7149–7155.

68. Fadrna E, Spackova N, Sarzynska J, Koca J, Orozco M, Cheatham TE III, Kulinski T, Sponer J. Single stranded loops of quadruplex DNA as key benchmark for testing nucleic acids force fields. J. Chem. Theor. Comput 2009;5:2514–2530.

69. Kollman PA, Massova I, Reyes C, Kuhn B, Huo SH, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. Accounts Chem. Res 2000;33:889–897.

70. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices. J. Am. Chem. Soc 1998;120:9401–9409.

71. Luo R, David L, Gilson MK. Accelerated Poisson-Boltzmann calculations for static and dynamic systems. J. Comput. Chem 2002;23:1244–1253. [PubMed: 12210150]

72. Sitkoff D, Sharp KA, Honig B. Accurate calculation of hydration free energies using macroscopic solvent models. J. Phys. Chem 1994;98:1978–1988.

73. Brooks B, Karplus M. Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsin inhibitor. Proc. Natl. Acad. Sci. U.S.A 1983;80:6571–6575. [PubMed: 6579545]

74. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Crowley, M.; Ross, WS.; Zhang, W.; Merz, KM.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, KF.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, SR.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, DH.; Seetin, MG.; Sagui, C.; Babin, V.; Kollman, PA. AMBER 10. San Francisco: University of California; 2008.

75. Alfonso DR, Jordan KD. A flexible nudged elastic band program for optimization of minimum energy pathways using ab initio electronic structure methods. J. Comput. Chem 2003;24:990–996. [PubMed: 12720320]

76. Jonsson, H.; Mills, G.; Jacobsen, KW. Nudged elastic band method for finding minimum energy paths of transitions. In: Berne, BJ.; Ciccoti, G.; Coker, DF., editors. Classical and Quantum Dynamics in Condensed Phase Simulations. Singapore: World Scientific; 1998. p. 384-404.

77. Henkelman G, Jonsson H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. J. Chem. Phys 2000;113:9978–9985.

78. Mathews DH, Case DA. Nudged elastic band calculation of minimal energy paths for the conformational change of a GG non-canonical pair. J. Mol. Biol 2006;357:1683–1693. [PubMed: 16487974]

79. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. J. Comput. Chem 1992;13:1011–1021.

80. Noy A, Pérez A, Laughton CA, Orozco M. Theoretical study of large conformational transitions in DNA: the B<->A conformational change in water and ethanol/water. Nucleic Acids Res 2007;35:3330–3338. [PubMed: 17459891]

81. Aci S, Mazier S, Genest D. Conformational pathway for the kissing complex - Extended dimer transition of the SL1 stem-loop from genomic HIV-1 RNA as monitored by targeted molecular dynamics techniques. J. Mol. Biol 2005;351:520–530. [PubMed: 16023135]

82. Reblova K, Spackova N, Stefl R, Csaszar K, Koca J, Leontis NB, Sponer J. Non-Watson-Crick basepairing and hydration in RNA motifs: Molecular dynamics of 5S rRNA loop E. Biophys. J 2003;84:3564–3582. [PubMed: 12770867]

83. Hobza P, Kabelac M, Sponer J, Mejzlik P, Vondrasek J. Performance of empirical potentials (AMBER, CFF95, CVFF, CHARMM, OPLS, POLTEV), semiempirical quantum chemical methods (AM1, MNDO/M, PM3), and ab initio Hartree-Fock method for interaction of DNA bases: Comparison with nonempirical beyond Hartree-Fock results. J. Comput. Chem 1997;18:1136–1150.

84. Sponer JE, Spackova N, Kulhanek P, Leszczynski J, Sponer J. Non-Watson-Crick base pairing in RNA. Quantum chemical analysis of the cis Watson-Crick/sugar edge base pair family. J. Phys. Chem. A 2005;109:2292–2301. [PubMed: 16838999]

85. Vaiana AC, Sanbonmatsu KY. Stochastic Gating and Drug–Ribosome Interactions. J. Mol. Biol 2009;386:648–661. [PubMed: 19146858]

86. Wood RH, Muhlbauer WCF, Thompson PT. Systematic errors in free energy perturbation calculations due to a finite sample of configuration space: sample size hysteresis. J. Phys. Chem 1991;95:6670–6675.

87. Jucker FM, Heus HA, Yip PF, Moors EHM, Pardi A. A network of heterogeneous hydrogen bonds in GNRA tetraloops. J. Mol. Biol 1996;264:968–980. [PubMed: 9000624]

88. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. Proc. Natl. Acad. Sci.U.S.A 2004;101:7287–7292. [PubMed: 15123812]

89. Mathews DH, Sabina J, Zuker M, Turner DH. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol 1999;288:911–940. [PubMed: 10329189]

90. Xia T, SantaLucia J Jr, Burkard ME, Kierzek R, Schroeder SJ, Jiao X, Cox C, Turner DH. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry 1998;37:14719–14735. [PubMed: 9778347]

91. Nowotny V, Nierhaus KH. Protein L20 from the large subunit of Escherichia coli ribosomes is an assembly protein. J. Mol. Biol 1980;137:391–397. [PubMed: 7021848]

92. Schneider B, Moravek Z, Berman HM. RNA conformational classes. Nucleic Acids Res 2004;32:1666–1677. [PubMed: 15016910]

93. Besseova I, Otyeopka M, Reblova K, Sponer J. Dependence of A-RNA simulations on the choice of the force field and salt strength. Phys. Chem. Chem. Phys 2009;11:10701–10711. [PubMed: 20145814]
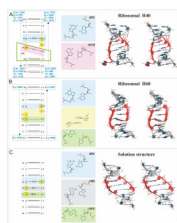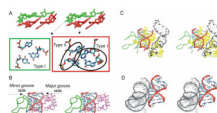
Figure 1.

2D structures (Left) and 3D stereo views (Right) of studied segments including non-Watson-Cricks base pairs in the internal UAA/GAA loop (Middle). A. Left: 2D X-ray structure of H40 with unified sequence flanking the internal loop (see Materials and Methods). Middle: sheared A/G and rH U/A base pairs. Right: Stereo view of *E.c.* H40 X-ray structure. B. Left: 2D X-ray structure of *E.c.* 23S rRNA H68 exhibiting an alternative conformation of UAA/GAA motif with unified canonical flanking sequence. Black dashed line indicates single H-bond. Middle: unpaired G and A bases, stacking middle adenines and single bonded A/U base pair. Right: Stereo view of this structure. C. Left: 2D NMR structure. Middle: sheared A/G, sheared A/A and incomplete *c*WS A/U base pairs. Right: Stereo view of the NMR structure. In all Figures, bases of the UAA/GAA internal loop are in red, 3D structures are colored accordingly, hydrogens are not shown in the X-ray structures, bases in yellow boxes in the 2D structures are involved in stacking and the marks between the bases indicate base paring family according to the Leontis & Westhof classification (*t*HS = *trans* Hoogsteen/Sugar edge A/G or A/A, known also as "sheared" base pairs, *t*WH = *trans* Watson-Crick/Hoogsteen U/A, known also as reverse Hoogsteen (rH) base pair, and *c*WS = *cis* Watson-Crick/Sugar edge A/U).[6,7] X-ray nucleotide numbers are in blue, NMR numbers are in black. The green rectangular trapezium for H40 structure marks bases forming the UA_handle submotif.[8]
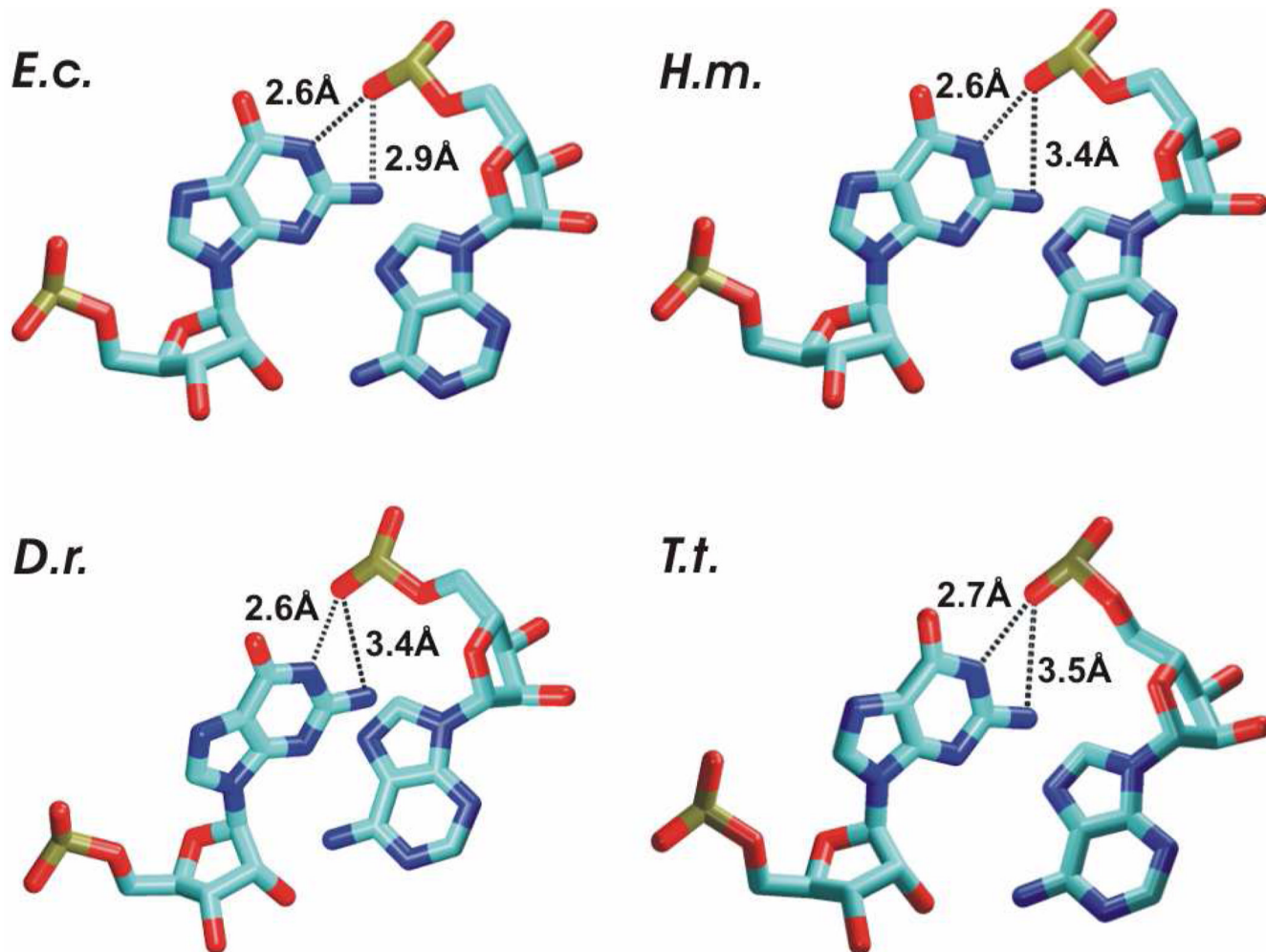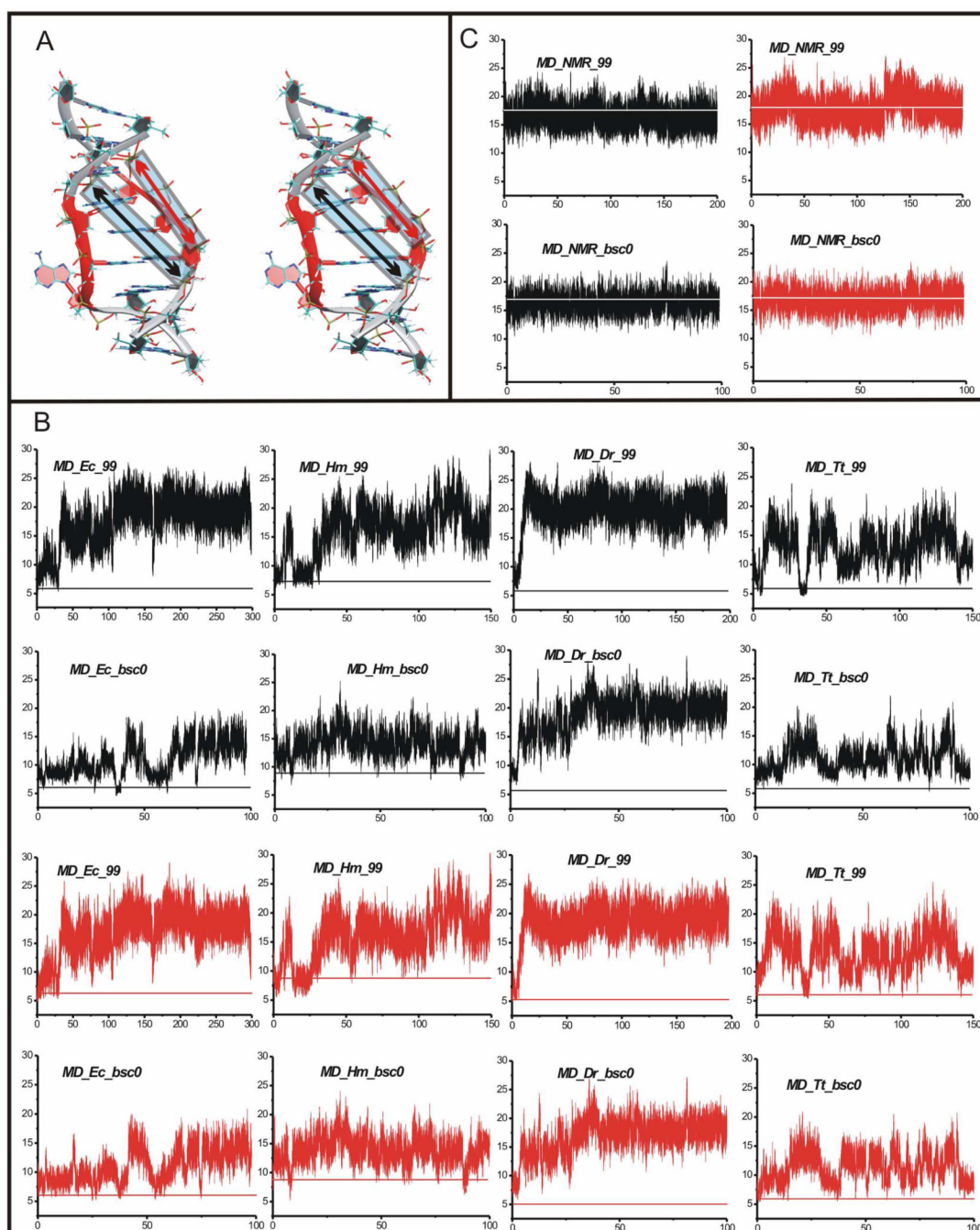
Figure 2.
Base phosphate interactions observed in the ribosomal X-ray structures of H40. *E.c.*
structure exhibits bifurcated binding mode (base phosphate interaction type 4BPh) in which
N2 and N1 of G13 bind to the same anionic oxygen of the phosphate group A5(O2P). *H.m.*,
D.r. and *T.t.* structures exhibit only G13(N1)-A5(O2P) H-bond, which represents base
phosphate interaction type 5BPh. The differences might reflect limits of the resolution of the
experimental structures.

Figure 3.
A. Stereo view of three adenines of the UAA/GAA motif from *E.c.* H40 forming an AAA stack which interacts with two C≡G base pairs from the hairpin between H39 and H40 via A-minor interactions. C≡G pair and the corresponding interacting adenine(s) are highlighted with the same color. Details of these interactions are visualized below the stereo view in corresponding green and red boxes, including a description of the A-minor interaction type. B. Stereo view of bacterial *E.c.* H40 interacting with hairpin between H39 (in green) and H40 and ribosomal protein L20 (in magenta). C. Stereo view of archaeal *H.m.* H40 interacting with the hairpin between H39 (in green) and H40, ribosomal protein L30 (in yellow) and H25 (in grey). D. Stereo view of bacterial *E.c.* H68 (exhibiting the alternative conformation of UAA/GAN internal loop) interacting with H75 (in grey).

Figure 4.
A. Stereo view of the averaged 55–57 ns MD structure of the simulated X-ray H40 UAA/GAA internal loop from the MD_Ec_99 simulation. The structure exhibits wider (more open) major groove compared to the original geometry (see Figure 1A). Monitored inter-phosphate distances across the major groove are indicated by black (12P-3P) and red arrows (11P-4P) in blue transparent boxes. B. Time courses of two inter-phosphate distances (12P-3P in black and 11P-4P in red) in standard MD simulations of X-ray H40 run with the parm99 force field and in control simulations run with the parmbsc0 force field, all with net-neutralizing Na$^+$ (Table 1). The x-axis stands for time (in nanoseconds) while y-axis stands for inter-phosphate distance (in Å). Horizontal lines show experimental distances. C. Time

course of two inter-phosphate distances (12P-3P in black and 11P-4P in red) in standard MD simulation of NMR structure run with the parm99 force field and in control simulation run with the parmbsc0 force field.
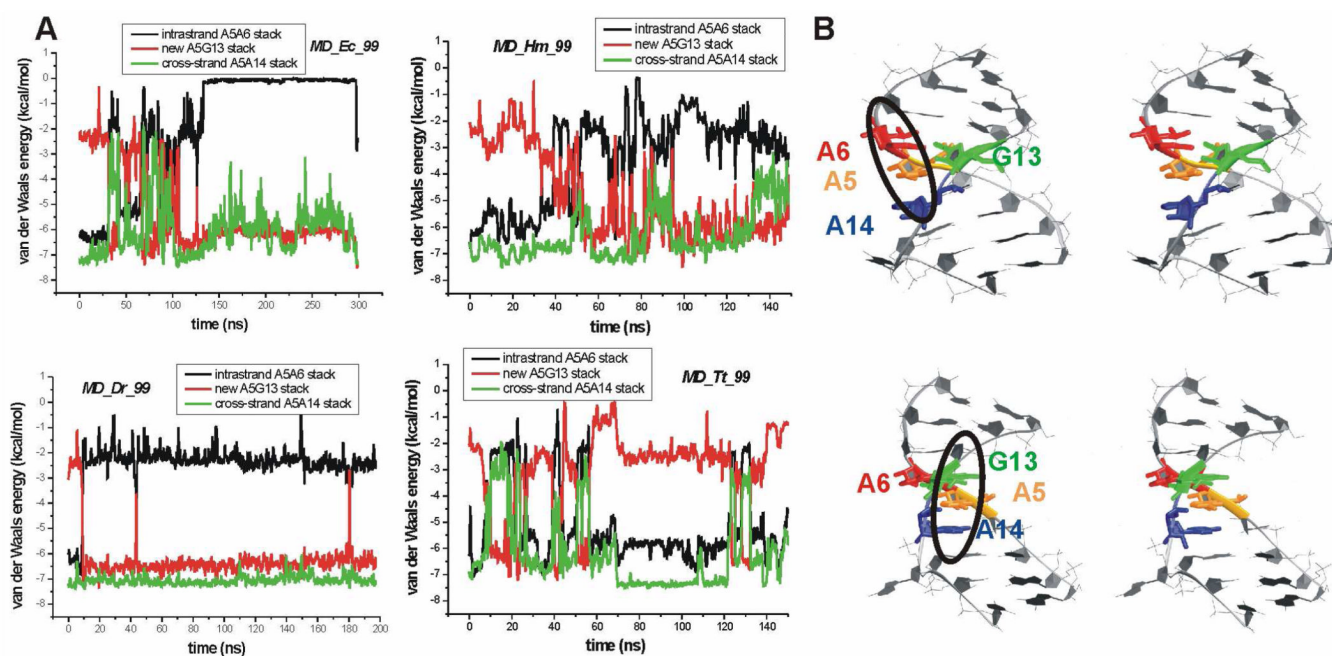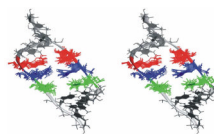
Figure 5.
A. Time courses of the van der Waals interaction energy calculated between bases forming A5A6, A5A14 and A5G13 stacks in the standard simulations of X-ray H40 run with parm99 force field. B. Stereo views of X-ray H40 UAA/GAA internal loop with colored bases forming stacks. Top – the original stacking pattern highlighted by the black oval, *i.e.* A5 forms an intrastrand stack with A6, and simultaneously cross-strand stack with A14. Bottom – the stacking pattern formed in the course of the simulations (highlighted by the black oval) where A5 stacks with G13 and also with A14.
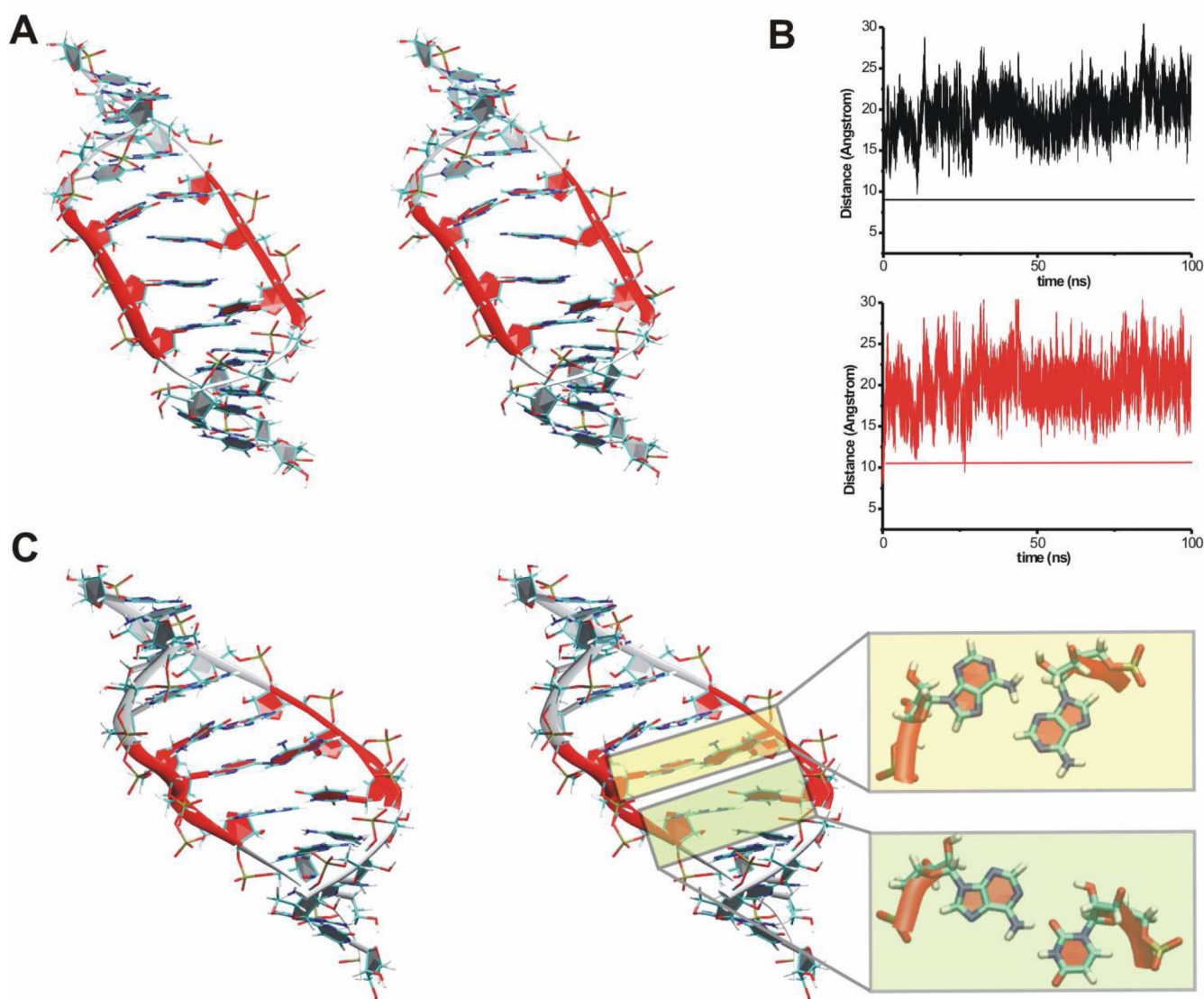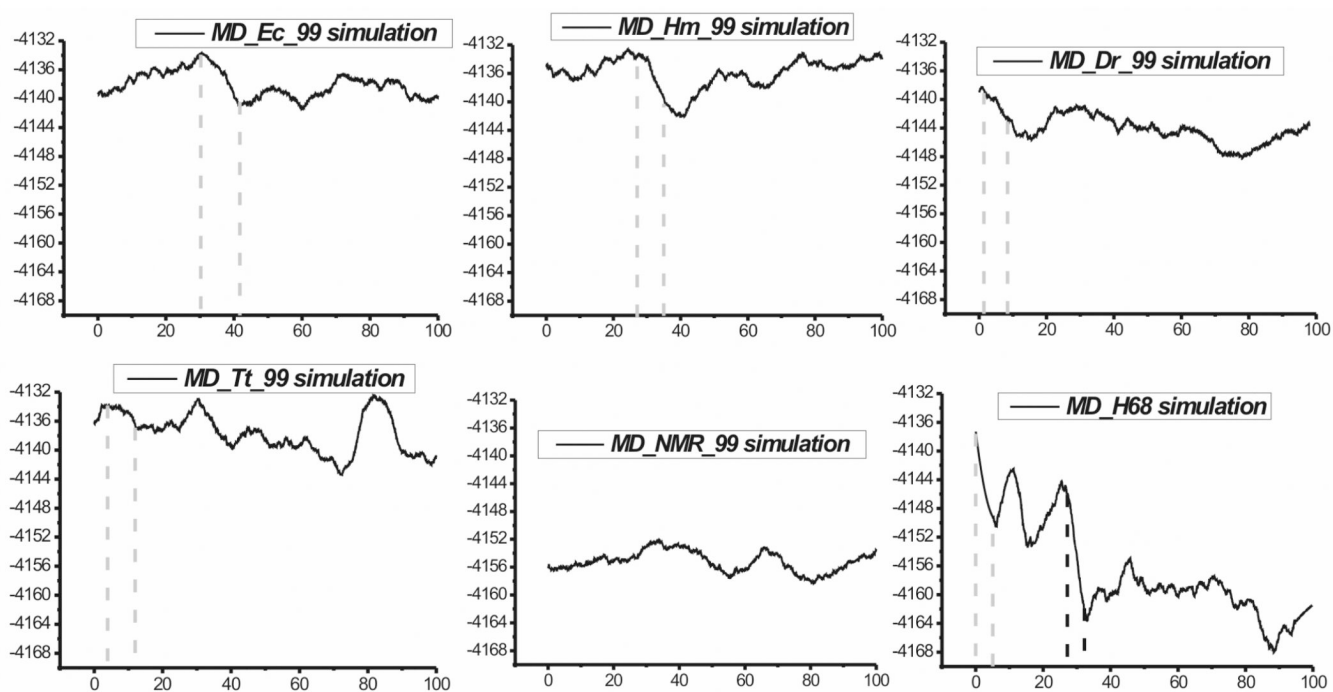
Figure 6.
A. Stereo view of the snapshot structure of the H68 UAA/GAA internal loop from the MD_68 simulation at 5 ns. The structure exhibits a wider (more open) major groove compared to the original geometry (see Figure 1B) B. Time courses of two inter-phosphate distances (12P-3P in black and 11P-4P in red) along the MD_68 simulation. C. Stereo view of the snapshot structure of the H68 UAA/GAA internal loop from the MD_68 simulation at 40 ns with formed sheared A14/A5 and *c*WS A15/U4 pairs, which are highlighted in the color transparent boxes. This structure closely resembles the NMR structure.
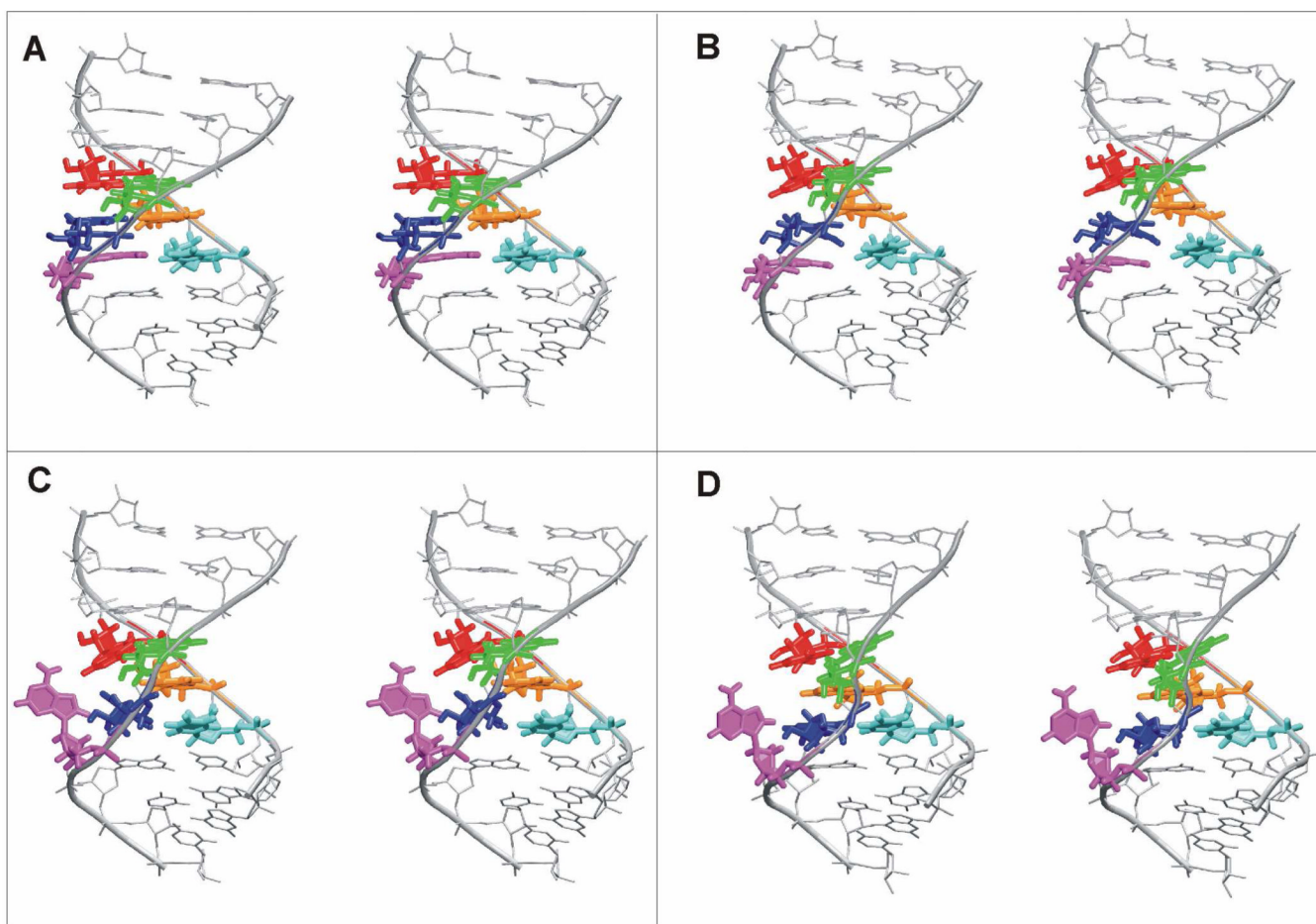
Figure 7.
3D stereo view of a snapshot of the UAA/GAA structure from the LES_Ec simulation with multiple copies of nucleotides in the internal loop (LES region). The structure resembles the solution structure (Figure 1C), *i.e.* it has wide major groove and coplanar A6 and G13 (red), A5 and A14 (blue) and U4 and A15 (green) bases.

Figure 8.
Total free energy time courses in standard net-neutralizing Na$^+$ simulations with parm99 force field. The x-axes stand for time (in nanoseconds) while y-axes stand for total free energy (in kcal/mol). The grey vertical dashed lines mark the time period when initial opening of the major groove was observed. In the MD_Tt_99 simulation the major groove oscillated back and forth with the grey lines indicating the first opening. In the time course of the MD_H68 simulation the vertical black dashed lines indicate the time period when the sheared A/A and cWS A/U pairs formed.

Figure 9.
3D stereo view of structures showing the conformational transition between NMR and X-ray structures predicted by NEB calculations. G13 is highlighted in red, A6 in green, A14 in blue, A5 in yellow, A15 in magenta and U4 in cyan. A) Starting NMR structure. B) Intermediate structure where the A14/A5 pair breaks and moves away from being stacked with A6 and A15. C) Intermediate structure where A15 slides out of the helix to become bulged out. D) Final structure where A14 pairs with U4 while A5 is unpaired and stacks with A6. This structure corresponds to the arrangement of the H40 X-ray structure.

Table 1

Survey of performed simulations. The sequence of the simulated molecules was unified to match the sequence used in the NMR study (see Materials and Methods and Supporting Information).

Organism	Simulated segment (original experimental numbering)	Simulation name	Resolution (Å) and pdb code	Length of simulation (ns)	RMSD$^+$ (Å)	Force field
Standard MD simulations with net-neutralizing Na$^+$ atmosphere						
E.c.	996–1004, 1151–1159	MD_Ec_99	3.5, 2AW4	350*	4.4±1.2	Parm99
E.c.	996–1004, 1151–1159	MD_Ec_bsc0	3.5, 2AW4	100	2.3±0.4	Bsc0
H.m.	1093–1101,1255–1263	MD_Hm_99	2.4,1S72	250**	3.5±1.0	Parm99
H.m.	1093–1101,1255–1263	MD_Hm_bsc0	2.4, 1S72	100	2.7±0.5	Bsc0
D.r.	1007–1015, 1162–1170	MD_Dr_99	3.1, 1NKW	200	4.8±0.7	Parm99
D.r.	1007–1015, 1162–1170	MD_Dr_bsc0	3.1, 1NKW	100	4.3±0.8	Bsc0
T.t.	996–1004, 1151–1159	MD_Tt_99	2.8, 2J01	200**	2.4±0.7	Parm99
T.t.	996–1004, 1151–1159	MD_Tt_bsc0	2.8, 2J01	100	2.3±0.5	Bsc0
E.c.	996–1004, 1151–1159	MD_A5U%	3.5, 2AW4	100	2.4±0.5	Parm99
E.c.	996–1004, 1151–1159	MD_A14U&	3.5, 2AW4	100	2.2±0.3	Parm99
H.m.	1093–1101,1255–1263	MD_A14G_U4C@	2.4, 1S72	50	1.9±0.4	Parm99
E.c.	996–1004, 1151–1159	MD_nosalt#	3.5, 2AW4	150	4.3±1.0	Parm99
E.c.	996–1004, 1151–1159	MD_400K$	3.5, 2AW4	20	6.7±3.2	Parm99
E.c.	996–1004, 1151–1159	MD_400K$$	3.5, 2AW4	20	5.1±2.7	Parm99
E.c.	996–1004, 1151–1159	MD_LES_Ec~	3.5, 2AW4	80	4.2±1.7	Parm99

Organism	Simulated segment (original experimental numbering)	Simulation name	Resolution (Å) and pdb code	Length of simulation (ns)	RMSD+ (Å)	Force field
N/A	1–18	MD_NMR_99	N/A, 2H49	200	1.6±0.3	Parm99
N/A	1–18	MD_NMR_bsc0	N/A, 2H49	100	1.2±0.2	Bsc0
N/A	1–18	MD_NMR_rest $r^{\#\#}$	N/A, 2H49	200	1.7±0.3	Parm99
E.c.	1885–1893, 1849–1857	MD_H68	3.5, 2AW4	100	3.9±0.5	Parm99
Standard MD simulations in excess of KCl						
E.c.	996–1004, 1151–1159	MD_Ec_K1 ^{x}	3.5, 2AW4	100	2.1±0.5	Parm99
E.c.	996–1004, 1151–1159	MD_Ec_K2 ^{x}	3.5, 2AW4	100	2.0±0.3	Bsc0
E.c.	996–1004, 1151–1159	MD_Ec_K3 ^{x,x}	3.5, 2AW4	100	2.9±0.4	Parm99
N/A	1–18	MD_NMR_K ^{x}	N/A, 2H49	100	1.5±0.3	Parm99
LES simulations						
E.c.	996–1004, 1151–1159	LES_Ec	3.5, 2AW4	60	6.7±2.2	Parm99
D.r.	1007–1015, 1162–1170	LES_Dr	3.1, 1NKW	40	5.9±1.3	Parm99
T.t.	996–1004, 1151–1159	LES_Tt	2.8, 2I01	40	6.4±1.3	Parm99
H.m.	1093–1101,1255–1263	LES_Hm	2.4, 1S72	40	6.4±2.0	Parm99

+ RMSD values are calculated along the trajectory for the individual snapshots with respect to the starting structure.

* Due to disruption of the structure we considered only 0–300 ns trajectory portion in the analyses.

** Due to disruption of the structure we considered only 0–150 ns in the analyses.

% Simulation run with A5U mutation.

& Simulation run with A14U mutation.

@Simulation run with A14G and U4C mutations.

#Simulation run under no-salt condition.

$Simulation run at 400K (NVT).

$$simulation run at 400K (NPT).

~Standard MD simulation that started from the NMR-like conformation observed in the LES_Ec simulation.

##Simulation run with restraint, which enforced direct A(N1)–U(O2´) H-bond of the cWS A/U base pair (instead of the water-mediated one) for 10 ns.

xSimulation run with Dang's parameters for K^+ and Cl^- (see Materials and Methods).

xxSimulation run with Joung and Cheatham's parameters for K^+ and Cl^- (see Materials and Methods).

Table 2

Base pairing changes detected in the standard simulations performed with the X-ray H40 UAA/GAA structure.

Simulation name	sheared A/G pair		reverse Hoogsteen U/A pair	
	A(N7)-G(N2) H-bond	A(N6)-G(N3) H-bond	U(O2)-A(N6) H-bond	U(N3)-A(N7) H-bond
MD_Ec_99	75–78 ns [oe]	33–40 ns [oe]	stable	89.2–90.1 ns [oe]
	133–297 ns [d]	65–78 ns [oe]		at 101 [d]
		133–297 ns [d]		
MD_Ec_bsc0	[f]	Stable	stable	18.5–23.2 ns [oe]
				41.3–46.8 ns [oe]
				at 65 ns [d]
MD_Hm_99	at 150 ns [d]	39–45 ns [oe]	stable	56–81 ns [oe]
		45–97 [f]		
		97–109 [oe]		
		150 ns [d]		
MD_Hm_bsc0	1–5 ns [oe]	[f]	stable	5–15 ns ns [oe]
	90–94 ns [oe]			26–37 ns [oe]
				75–96 ns [oe]
MD_Dr_99	stable	[f]	stable	at 7 ns [d]
MD_Dr_bsc0	stable	[f]	stable	At 30 ns [d]
MD_Tt_99	63–65 ns [oe]	63–65 ns [oe]	stable	70–122 ns [oe]
MD_Tt_bsc0	58–69 ns [oe]	58–69 ns [oe]	stable	13–17 ns ns [oe]
				24–32 ns [oe]
				60–78 ns [oe]
				83–96 ns [oe]
MD_Ec_K1	[f]	stable	77–80 ns [oe]	4–22 ns [oe]
				77–80 ns [oe]
				96–98 ns [oe]
MD_Ec_K2	stable	[f]	stable	11–28 ns [oe]
MD_Ec_K3	stable	stable	stable	8–9 ns [oe]
				16–20 ns [oe]

"oe", "d" and "f" marks stand for temporary opening, disruption (till the end of the simulation) and considerable fluctuations, respectively.