BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Structure and Conformational Dynamics of the Domain 5 RNA Hairpin of a Bacterial Group II Intron Revealed by Solution Nuclear Magnetic Resonance and Molecular Dynamics Simulations

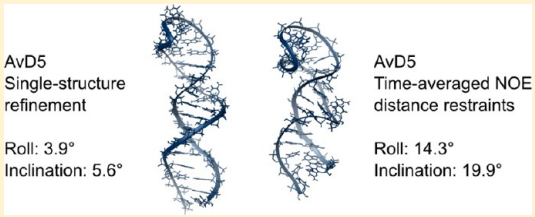
Maria Pechlaner,[†] Roland K. O. Sigel,[†] Wilfred F. van Gunsteren,[‡] and Jožica Dolenc*[,‡]

[†]Institute of Inorganic Chemistry, University of Zurich, CH-8057 Zurich, Switzerland

[‡]Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, CH-8093 Zurich, Switzerland

Supporting Information

ABSTRACT: Nuclear magnetic resonance (NMR) nuclear Overhauser enhancement (NOE) data obtained for a 35-nucleotide RNA segment of a bacterial group II intron indicate a helical hairpin structure in which three parts, a terminal pentaloop, a bulge, and a G-A mismatch, display no Watson−Crick base pairing. The 668 NOE upper distance bounds for atom pairs are insufficient to uniquely determine the conformation of these segments. Therefore, molecular dynamics simulations including time-averaged distance restraints have been used to obtain a conformational ensemble compatible with the observed NMR data. The ensemble shows alternating hydrogen bonding patterns for the mentioned segments. In particular, in the pentaloop and in the bulge, the hydrogen bonding networks correspond to distinct conformational clusters that could not be captured by using conventional single-structure refinement techniques. This implies that, to obtain a realistic picture of the conformational ensemble of such flexible biomolecules, it is necessary to properly account for the conformational variability in the structure refinement of RNA fragments.



Noncoding RNAs can fold into complex tertiary structures that allow them to perform highly specific catalytic and regulatory tasks in gene expression, translation, and splicing and make them molecules of interest in medicine, biotechnology, and nanotechnology.[1−4] The overall fold is based on a network of tertiary contacts mediated largely by the bulges and loops that connect or cap the regular base-paired helices.[5] Often those unpaired or noncanonically paired segments are rather flexible and can adapt or even significantly rearrange upon binding to an interaction partner.[6−9] A static crystal structure of these parts of RNA might therefore not give a complete picture of its structural variability, and neither might a set of structures derived from NMR data if averaging effects are not considered in the structure refinement procedure.

Group II introns are large self-splicing ribozymes[10] composed of six domains, the largest of which is domain 1 (D1), which folds first and serves as a scaffold for the other domains. In particular, it provides the docking site for domain 5 (D5), a small but exceptionally conserved RNA hairpin. Interestingly, D1 and D5 alone have been shown to be sufficient to perform the first step of splicing even when they are not covalently linked to each other (i.e., in *trans*).[11]

Domain 5 is characterized by two helical regions that are connected by an asymmetric internal loop (we call it the "bulge") and capped by a terminal loop (Figure 1A).[12,13] Three nucleotides at the 5′ side of the lower helix are known as the "catalytic triad" because this motif (in particular, the invariant central guanine) has been shown to play an important role in catalysis *in vivo*.[14−16] While the bulge is also essential for intron activity,[17−19] the terminal loop, on the other hand, has a merely structural role, forming a tertiary contact to D1.[20] The similarities of this small D5 domain to spliceosomal snRNA U6[21,22] have evoked strong interest in its structure and function. Two highly related D5 structures derived from solution NMR data of *Saccharomyces cerevisiae* group II intron ai5γ (ScD5) and *Pylaiella littoralis* group II intron LSU/2 (PlD5) are available[12,13] in addition to the crystal structure of a D5D6 construct from ai5γ[23] and crystal structures of a group IIC intron from *Oceanobacillus iheyensis*.[24−28] Studies of D5 in solution indicate pronounced flexibility of the catalytic triad, of the loop, and in particular of the bulge region.[12,13,29] Differences observed between the bulge regions in the solution and in the crystal structures hint at an intrinsic flexibility that might be an important requirement to allow for a structural transition upon D5 docking. While in the solution structures the bulge adenine A24 is stacked in and the guanine G26 is turned outward, in all crystal structures A24 is turned outward, G26 is engaged in a GU wobble pair, and the helical parts below and above the bulge stack onto each other (Figure S1 of the Supporting Information).

We have been interested in D5 from *Azotobacter vinelandii* intron 5 (AvD5) for the evaluation of the structural impact of two significant alterations compared to the previously studied D5: AvD5 has a GA mismatch at the center of the catalytic triad (GU in all other structures studied), and an unusual UAGUU

Received: June 18, 2013
Revised: August 27, 2013
Published: September 3, 2013

   dx.doi.org/10.1021/bi400784r | *Biochemistry* 2013, 52, 7099−7113

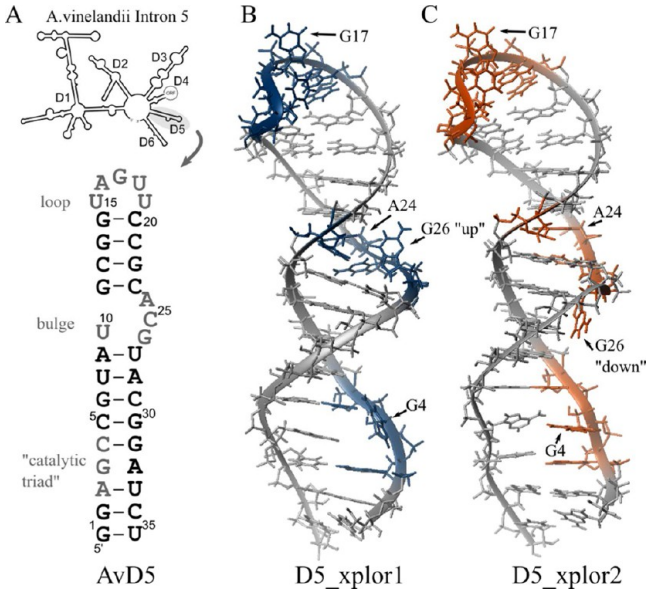


Figure 1. (A) Domain architecture of *A. vinelandii* intron 5 (top). Sequence and secondary structure of AvD5 (bottom). The loop, bulge, and "catalytic triad", which hosts the GA mismatch, are colored gray. (B) Starting structures used in the restrained MD simulations. D5_xplor1 and D5_xplor2 are the two-lowest energy structures of a 200-member set of structures calculated using Xplor-NIH.[47,48] The crucial residues of the loop, bulge, and catalytic triad (G17, G26, and G4) are marked.

pentaloop replaces the canonical GAAA tetraloop (Figure 1). The bulge region is identical to that in ScD5 and *O. iheyensis* D5 and almost identical to that in PlD5 (Figure S2 of the Supporting Information).

Both ScD5 and PlD5 solution structures were obtained using standard simulated annealing and refinement protocols in CNS[30] using a rather simple force field and no electrostatics. This approach is still widely used in the determination of structures of biological macromolecules based on NMR data.[31−34] Usually, calculations are performed *in vacuo*, i.e., in the absence of solvent and counterions. Only lately has the use of force fields that include charges and account for the solvent environment in the final refinement stage become more common.[35−37] The difference that including electrostatics and an explicit solvent environment can make in NMR structure refinement of RNA hairpins is demonstrated by the recently re-refined structures of ScD5 and PlD5 that indicate that the two structures are much more alike than reported previously.[38]

NMR spectroscopy records ensemble-averaged signals,[39−44] which is however seldom taken into account in the conventional derivation of structures, despite its great importance in flexible systems such as nucleic acids.[45,46] Usually, instantaneous distance restraints are employed in NMR structure refinement, penalizing each deviation from an experimentally determined distance range. Time-averaged restraints require only that the average distance during a certain period of time remain within the set range and thus allow for exchange between conformations that might individually not satisfy the instantaneous restraint.

NMR spectra of AvD5 contain indications of conformational averaging in all three noncanonical regions: around the GA mismatch, in the bulge, and in the pentaloop. In the two former regions, this is evident from the resonance broadening in the spectra that leads to a scarcity of restraints. In the loop, it is reflected by atom−atom distance restraints that are impossible to satisfy by a single conformation. By omitting the contradictory restraints and using generous error limits for distances derived from broadened peaks, we could calculate a set of structures using the standard structure refinement protocol as implemented in CNS[30] and Xplor-NIH.[47,48] While satisfying the restraints, however, the set of structures does not give a well-resolved picture of the close contacts such as hydrogen bonds in the GA, bulge, and pentaloop regions. In addition, in the bulge region two different conformations, one with G26 bulged into the minor groove (G26 up, Figure 1B) and one with G26 bulged into the major groove (G26 down, Figure 1C) similar to those of the ScD5 and PlD5 structures, were suggested. This ambiguity and the indications of conformational exchange in the unpaired regions prompted us to proceed with a structure refinement protocol employing time-averaged distance restraints and an explicit solvent environment. For comparison, the simulations with time-averaged distance restraints were complemented by simulations with instantaneous distance restraints as well as by simulations at elevated temperature to increase the degree of sampling. In each case, we performed 10 ns of restrained molecular dynamics simulations starting from each of the two lowest-energy structures resulting from the XPLOR calculation [D5_xplor1 with G26 up (Figure 1B) and D5_xplor2 with G26 down (Figure 1C)] to investigate the structural ensemble of AvD5 focusing on the GA, bulge, and pentaloop regions.

MATERIALS AND METHODS

NMR Experiments. *Sample Preparation.* All chemicals were purchased from either Fluka-Sigma-Aldrich (Buchs, Switzerland) or Brunschwig Chemie (Basel, Switzerland) at puriss p.a. or biograde. DNA strands were from Microsynth (Balgach, Switzerland), with or without 2′-methoxy modifications at the two terminal nucleotides to prevent RNA 3′-heterogeneity.[49] The nucleotide 5′-triphosphates (NTPs) were obtained from GE Healthcare (Glattbrugg, Switzerland), with the exception of UTP, which was purchased from Acros-Organics (Geel, Belgium). ^{13}C- and ^{15}N-labeled NTPs were purchased from Silantes GmbH (München, Germany), while the partially deuterated NTPs were obtained from Cambridge Isotope Laboratories (Andover, MA). AvD5 was transcribed *in vitro* using in-house-prepared T7 polymerase and purified as described previously.[50] The transcription mix contained each NTP at 5 mM (either natural abundance isotopes, completely ^{15}N- and ^{13}C-labeled, or partially deuterated), 0.5 mM double-stranded DNA template (Microsynth), 0.1% Triton X-100, 40 mM Tris-HCl (pH 7.5), 40 mM DTT, 2 mM spermidine, and 35 mM $MgCl_2$. The RNA was purified on 18% polyacrylamide gels and recovered by electroelution using a Whatman Elutrap System. Samples were desalted using Vivaspin centrifugal concentrators (Sartorius Stedim biotech, Aubagne, France), lyophilized, and redissolved in 250 μL of 60 mM KCl and 10 μM EDTA in D_2O or a mixture of H_2O and D_2O (9:1) for NMR experiments. The pH (pD for samples in D_2O) was adjusted to 6.7−6.9.

NMR Spectroscopy. NMR spectra were recorded on a Bruker AV700 MHz spectrometer equipped with a CP-TXI *z*-axis pulsed field gradient CryoProbe and on an AV600 MHz spectrometer equipped with a TCI z-gradient CryoProbe. ^{31}P spectra were recorded on an AV500 spectrometer with a 5 mm QNP CryoProbe. Nonexchangeable protons were assigned from two-dimensional (2D) ^{1}H−^{1}H NOESY spectra in D_2O

dx.doi.org/10.1021/bi400784r | *Biochemistry* 2013, 52, 7099−7113

recorded at 290 or 300 K with a mixing time of 60, 120, or 250 ms. Exchangeable protons were assigned from ^{1}H−^{1}H NOESY spectra in a mixture of H_2O and D_2O (9:1) recorded at 275 K with a mixing time of 150 ms. Watson−Crick base pairs were confirmed in 2D J_{NN} HNN-COSY spectra at 275 K.[51] ^{13}C and ^{15}N resonances were assigned in ^1J and ^2J HSQC spectra. In addition, phase sensitive 2D w1,w2-^{15}N,^{13}C-filtered NOESY and TOCSY experiments with watergate H_2O suppression using nucleotide-specifically ^{15}N- and ^{13}C-labeled samples helped to reduce the extent of spectral overlap.[52] ^{1}H chemical shifts were directly referred to external DSS (0.2%, pH 7.5), and ^{13}C and ^{15}N were indirectly referred to ^{1}H of DSS.[53]

NMR data were processed using TopSpin 3.0 (Bruker) and analyzed using Sparky (http://www.cgl.ucsf.edu/home/sparky/).

Structure Calculation using Xplor-NIH.[47,48] Distance restraints were derived from ^{1}H−^{1}H NOESY spectra recorded in D_2O with mixing times of 60−250 ms and from ^{1}H−^{1}H NOESY spectra measured in a mixture of H_2O and D_2O (9:1) with a mixing time of 150 ms. In regions with little or no overlap, peaks were integrated in Sparky and distances calibrated according to the volumes of H5H6 or H1′H2′ peaks using the CALIBA macro of DYANA.[54] On the basis of this, all assigned peaks were categorized as strong (0.18−0.3 nm), medium (0.18−0.45 nm), weak (0.3−0.6 nm), and very weak (0.4−0.7 nm). Sugar puckering was inferred from 50 ms ^{1}H−^{1}H TOCSY spectra: G17 and U18 have strong H1′H2′ and H1′H3′ peaks and were restrained to C2′-endo (δ = 145°, ν_1 = 25°, ν_2 = −35°, ±15°). G1, G4, A16, U19, C25, and G26 were left unrestrained because of intermediate H1′H2′ peaks or overlaps with the HOD signal, and all other sugar rings were restrained to C3′-endo (δ = 85°, ν_1 = −25°, ν_2 = 37.3°, ±15°). Glycosidic angles were restrained to the *anti* conformation (χ = −160 ± 15°) based on intranucleotide H8/H6−H1′ cross-peaks for all residues except G4, G17, U18, U19, C25, and G26, which were left unrestrained. Backbone torsional angles were restrained to standard A-helical values (α = −68°, β = 178°, γ = 54°, ε = −153°, ζ = −71°, ±10°) in the Watson−Crick base-paired regions. In other regions, α and ζ were restrained to exclude the *trans* range (0 ± 120°) based on the fact that there are no ^{31}P resonance outliers. Hydrogen bonds in Watson−Crick base pairs were enforced by short distance restraints between the hydrogen and the acceptor atom as well as between the donor and the acceptor atoms and by planarity restraints.

Starting from an extended conformation, we calculated an initial set of 100 structures using CNS 1.21[30,55] by applying 40 ps torsion-angle dynamics at 20000 K, followed by slow cooling for 90 ps in torsion-angle space and for 30 ps in Cartesian space. Afterward, an 80 ps refinement using Xplor-NIH 2.24[47,48] was performed, slowly cooling the system from 3000 to 50 K.

The 20 lowest-energy structures (none of them had NOE violations larger than 0.03 nm or torsional angles that deviated by more than 5° from the restrained range around the standard A-helical values described above) were selected as the final ensemble.

Molecular Dynamics Simulations. *Simulation Setup.* Molecular dynamics (MD) simulations reported here were performed with the GROMOS biomolecular simulation package (http://www.gromos.net).[56−58] The GROMOS 53A6$_{OXY+A}$ force field[59,60] with partial charges on the nucleic acid bases assigned to reproduce the chloroform/water partition coefficients has been used (see Tables S1−S4 of the Supporting Information). The D5_xplor1 (Figure 1B) and D5_xplor2 (Figure 1C) NMR model structures were taken as the initial structures for the simulations. The MD simulations were performed using rectangular periodic boundary conditions and explicit solvation using the SPC water model.[61] Na^+ and Cl^- ions were added to the simulation box to neutralize the negative charge on the RNA backbone and to mimic the experimental ionic strength of 60 mM. NOE distance restraints were used either as instantaneous or as time-averaged distance restraints and were imposed with a force constant of 6000 kJ mol^{-1} nm^{-2}. The memory relaxation time used in the time-averaged NOE distance restraints was 20 ps.[62−64]

The systems were first relaxed by performing a steepest descent energy minimization with harmonic positional restraints on all solute atoms (force constant of 2.5 × 10^4 kJ mol^{-1} nm^{-2}). The initial atomic velocities were taken from a Maxwell distribution at 60 K, which was followed by a 250 ps equilibration, in which the positional restraints were gradually released reducing the force constant to 0.0 kJ mol^{-1} nm^{-2} and the temperature was increased from 60 to 293 K. The leapfrog algorithm was used to integrate Newton's equations of motion using a time step of 2 fs. The center of mass motion was stopped every 2 ps. All bond lengths were constrained to their ideal values using the SHAKE algorithm with a relative geometric precision of 10^{-4}.[65] The temperature and pressure were maintained at 293 K and 1 atm, respectively, using weak coupling (τ_T = 0.1 ps, and τ_P = 0.5 ps) and an isothermal compressibility of 4.575 × 10^{-4} (kJ mol^{-1} nm^{-3})$^{-1}$.[66] Solute and solvent degrees of freedom were separately coupled to the heat bath. A reaction field approach[67] was used to treat the electrostatics by employing a triple-range cutoff scheme, with cutoffs of 0.8 and 1.4 nm, a five-step update of intermediate-range nonbonded forces, and a continuum dielectric permittivity of 61.[68]

From the equilibrated structures, the following MD simulations were started: two 10 ns MD simulations in which the NOE distance restraints were imposed as instantaneous restraints (D5_xplor1_IR and D5_xplor2_IR) and two 10 ns MD simulations in which the NOE distance restraints were imposed as time-averaged restraints (D5_xplor1_TAR and D5_xplor2_TAR). In addition, two 10 ns MD simulations with time-averaged NOE distance restraints were performed at 323 K to enhance the sampling in the AvD5 bulge region (D5_xplor1_TAR_323 and D5_xplor2_TAR_323).

Trajectory Analysis. The trajectory configurations were saved every 1 ps. Atom-positional root-mean-square deviations (rmsds) with respect to the energy-minimized initial structures were calculated as follows. The heavy atoms of the sugar−phosphate backbone were used to perform a superposition of centers of mass and a rotational least-squares fit superposition[69] of the trajectory structures onto the reference one. The rmsd values of the whole AvD5 and of its GA, loop, and bulge regions were calculated for all the heavy atoms. Additionally, atom-positional root-mean-square fluctuations (rmsfs) were calculated for all the heavy atoms of the AvD5 backbone and bases and are, for each nucleotide, reported as averages over the heavy atoms of the base and as averages over the corresponding backbone part. The trajectory configurations were also analyzed in terms of Watson−Crick hydrogen bonds and with respect to temporal evolution of hydrogen bonding in the GA, bulge, and pentaloop regions of AvD5. The criteria used in the hydrogen bond analysis were 0.25 nm as the upper bound of the H···A

(A, acceptor) distance and 135° as the lower bound of the D···H···A angle (D, donor). The RNA backbone geometry was analyzed in terms of torsional angles presented in Figure S3 of the Supporting Information. A joint conformational cluster analysis focusing on the AvD5 GA mismatch, bulge, and pentaloop regions was conducted on the trajectories generated at 293 K (D5_xplor1_IR, D5_xplor1_TAR, D5_xplor2_IR, and D5_xplor2_TAR) using the structures at 2 ps intervals and atom-positional rmsd similarity criteria of 0.15, 0.12, and 0.1 nm for the GA, bulge, and loop regions, respectively. A translational superposition of centers of mass and a rotational least-squares fit for every pair of configurations were performed using the AvD5 backbone heavy atoms. Interproton distance bounds derived from the NOE cross-peak intensities were compared with the average interproton distances calculated from the trajectories using $\langle r^{-6}\rangle^{-1/6}$ averaging. The H···H distances involving aliphatic H atoms were calculated by constructing virtual atoms based on standard geometries.[70] The results are presented as distance bound violations, that is, as a difference between the distances averaged over the simulation and the corresponding NMR-derived upper distance bounds. In this study, a total number of 668 NOEs assigned on the basis of the interpretation of the spectra were used (see Tables S5 and S6 of the Supporting Information). Of these, the NOEs between atoms in the three nonhelical regions, loop, bulge, and catalytic triad, are visualized in Figure S4 of the Supporting Information. Additionally, all possible hydrogen−hydrogen r^{-6} averaged distances of <0.45 nm in the AvD5 RNA hairpin were calculated. This resulted in approximately twice as many NOEs versus the number that was actually assigned. The absence of predicted NOEs from the list of observed NOEs is mainly due to the extensive overlap of the intra- and intersugar proton−proton peaks as well as to other causes such as spin diffusion or extinction of the NOE signals due to rotational tumbling.

Visualization of the AvD5 RNA hairpin was conducted using MolMol,[71] and RNA helical parameters were analyzed using Curves+.[72]

RESULTS

Atom-Positional Root-Mean-Square Deviations from the NMR Model Structures. The time series of the atom-positional rmsds from the energy-minimized starting structures for the heavy atoms of the entire AvD5 RNA hairpin and of its GA, bulge, and pentaloop regions are displayed in Figure 2 for the four simulations performed at 293 K. For the entire AvD5 RNA hairpin, the rmsd values vary between ~0.2 and 0.9 nm and show no drift after 2 ns (Figure 2A). The peaks observed in the case of the D5_xplor2_IR simulation can be attributed to a bending of the molecule across the flexible bulge. Such long-range motions are also responsible for like fluctuations in rmsd values in the other simulations. Because the distance information from NOE intensities is rather short-range, these data do not seem to inhibit the global dynamics of AvD5. As expected, local rmsd fluctuations in the smaller bulge, loop, and GA segments are smaller and of higher frequency (Figure 2B−D). The rmsd values of GA regions for the simulations initiated from the NMR model D5_xplor1 (D5_xplor1_TAR and D5_xplor1_IR) show an increase in the first 4 ns of simulation time when they reach the same values as in the simulations initiated from NMR model D5_xplor2 (D5_xplor2_TAR and D5_xplor2_IR) (Figure 2B). This slow gradual increase reflects the conversion of G4 from the looped out to a stacked conformation.

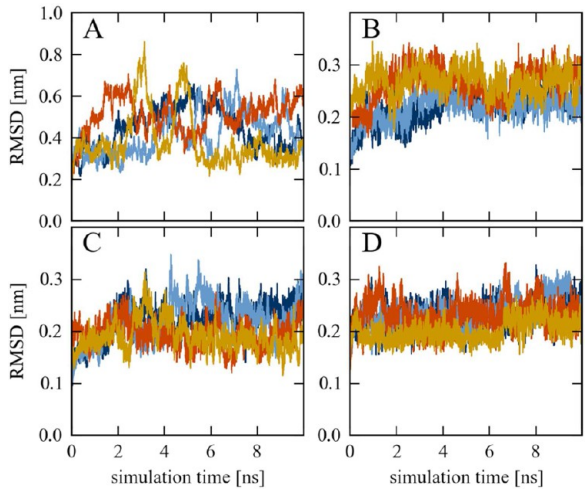


Figure 2. Time evolution of the root mean-square deviations (rmsds) with respect to starting structures D5_xplor1 and D5_xplor2 in the four MD simulations at 293 K: D5_xplor1_TAR (dark blue), D5_xplor1_IR (light blue), D5_xplor2_TAR (orange), and D5_xplor2_IR (yellow). The rmsds for the heavy atoms of (A) the complete hairpin, (B) the GA mismatch region (residues 3−5 and 31−33), (C) the bulge (residues 9−11 and 23−27), and (D) the loop (residues 15−19) are shown.

Atom-Positional Root-Mean-Square Fluctuations. Atom-positional rmsfs are highest at the hairpin's ends and in the loop region but are also elevated around the bulge (Figure 3). In all four simulations performed at 293 K, the fluctuations are larger in the sugar−phosphate backbone than in the nucleobases with the exception only of G4, G17, G26, and U27 in some of the simulations, reflecting the solvent exposure of these four bases. Interestingly, larger fluctuations can be observed in the MD simulations in which NOE distances were imposed as instantaneous restraints (Figure 3C,D) than in the MD simulations where NOE distances were imposed as time-averaged restraints (Figure 3A,B). This indicates that instantaneous distance restraints, while forcing the agreement with the experimental data, may artificially increase the mobility of unrestrained parts of the AvD5 RNA hairpin. Additionally, MD simulations initiated from NMR model structure D5_xplor1 with base G26 bulged up into the minor groove (Figure 3A,C) show fluctuations larger than those initiated from NMR model structure D5_xplor2 in which base G26 is turned down into the major groove (Figure 3B,D).

NOE Analysis. The initial structures do not contain violations of >0.03 nm that was the cutoff value for allowed NOE violations in the XPLOR calculations (Table S6 of the Supporting Information). As expected, all four MD simulations agree very well with the experimental NOE upper distance bounds that were imposed as either instantaneous or time-averaged distance restraints (Figure S5 of the Supporting Information). However, despite the NOE distance restraining a number of small NOE violations can be observed in the MD simulations. The strongest ones and the only ones exceeding 0.03 nm are found in the region of the pentaloop where the higher proton density makes spin diffusion more probable. The most pronounced violation is observed for the G17:H1′−C20:H1′ proton pair (entry 568 in Table S6 of the Supporting Information) in all four simulations with the largest value of 0.078 nm occurring in the D5_xplor1_IR simulation. This is not surprising because the NOE cross-peak of the G17:H1′−

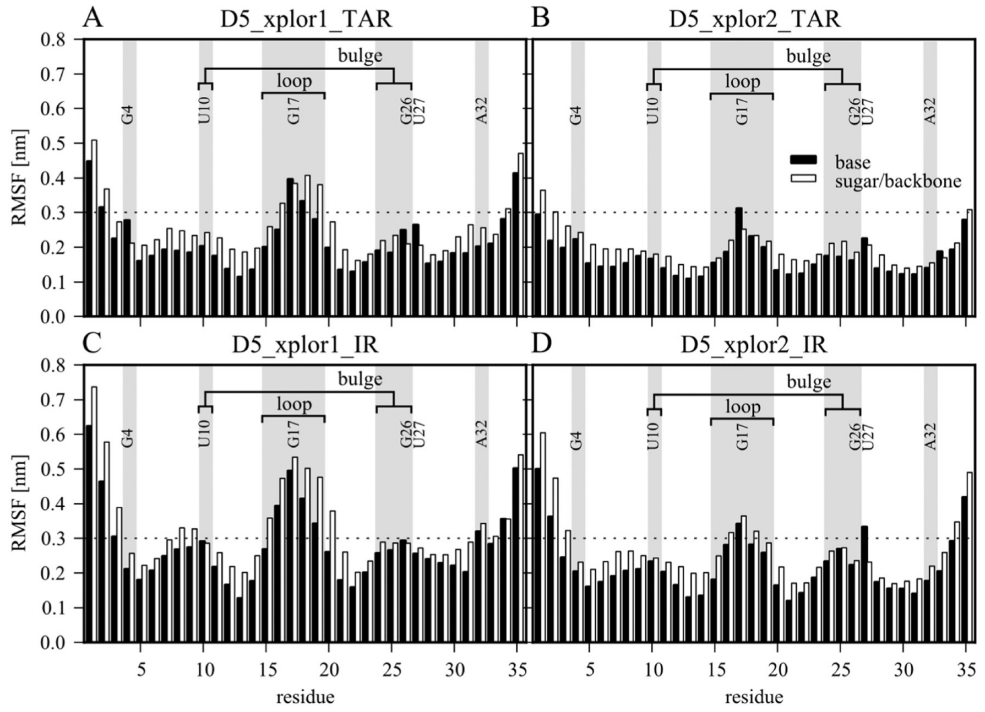


Figure 3. Atom-positional root mean-square fluctuations (rmsfs) during the last 8 ns of the four MD simulations at 293 K: D5_xplor1_TAR (A), D5_xplor2_TAR (B), D5_xplor1_IR (C), and D5_xplor2_IR (D). Average rmsfs over all heavy atoms of the base (black) and of the sugar−phosphate backbone (white) of each residue are displayed. The positions of the GA mismatch, the bulge, and the loop are shaded in gray, and key residues mentioned in the text are indicated. To avoid the bias from the initial structure, the first 2 ns of each trajectory was excluded from the rmsf analysis.

C20:H1′ proton pair is weak and was used in XPLOR calculations with a large error margin because spin diffusion could not be excluded. When time-averaged distance restraints are applied (Figure S5B,E of the Supporting Information) instead of instantaneous restraints (Figure S5C,F of the Supporting Information), many of the NOE violations decrease or completely disappear. Irrespective of the initial structure used, violations in the bulge are small, indicating that the orientation of base G26 cannot be unambiguously determined from the measured NMR data. The only noticeable NOE violation in the helical region of AvD5 comes from the C12:H5″−C12:H5 proton pair (entry 76 in Table S6 of the Supporting Information). The violations of 0.017 and 0.018 nm in D5_xplor1_IR and D5_xplor2_IR, respectively, disappear when using time-averaged distance restraints because of the concerted changing of the C12 α (−75° to −225°) and γ (65−180°) torsional angles that shortens the average C12:H5″−C12:H5 distance from 0.6 to 0.45 nm (Figure 4A−C). The small violation of the G1:H1−C34:H2′ proton pair in simulations with instantaneous restraints (D5_xplor1_IR and D5_xplor2_IR) is alleviated by using time-averaged distance restraints, too (Figure 4D). Time-averaged distance restraining allows for distance fluctuations that reduce bound violations compared to instantaneous distance restraining due to the nonlinear dependence of the NOE on the atom−atom distances.

End-to-End Distance. The analysis of the end-to-end distance of the molecule presented in Figure 5 reveals some compaction of the RNA hairpin when the time-averaged distance restraints are applied. The average end-to-end distance, which we define as the distance between the center of geometry of the G2-C34 base pair (we use the second base pair to avoid effects from fraying of the first base pair) and the phosphorus atom of G17, fluctuates around the initial value of 5.6 nm in the simulations with instantaneous distance restraints, while in the simulations with time-averaged distance restraints, it decreases to 5.0 nm. The AvD5 end-to-end distance is strongly influenced by bending of the hairpin in the bulge region, which occurs for instance after 6 ns in the case of the D5_xplor1_IR simulation. This prompted us to further investigate the compaction of the AvD5 RNA by calculating the base-pair parameters for the helical regions of AvD5.

Helical Base-Pair Parameters. To gain some quantitative description of the overall structural differences between the RNA hairpins from the four MD simulations, we analyzed the simulated trajectories in terms of the helical base-pair parameters using CURVES+.[72] In Table 1, the average values for the base-pair inclination and inter-base-pair roll calculated over the last 8 ns of the trajectories are reported. The compaction of the RNA hairpins simulated using time-averaged distance restraints observed in the analysis of end-to-end distance is reflected in an increase in the inter-base-pair roll and base-pair inclination. In the XPLOR starting structures the inter-base-pair roll and inclination are significantly lower than the standard A-helical values. In all MD simulations, the values of both parameters increase toward standard A-helical values. The increase is, however, much more pronounced in the simulations with time-averaged distance restraints.

Sugar and Base Flipping in the Loop and Bulge. As is evident from time series of the pseudorotation angle in Figure 6A−D, conformational flipping of the ribose sugar is more frequent when time-averaged distance restraints are applied. This reflects the ability of time-averaged distance restraints to compensate for short restraint violations that might arise during

dx.doi.org/10.1021/bi400784r | Biochemistry 2013, 52, 7099−7113

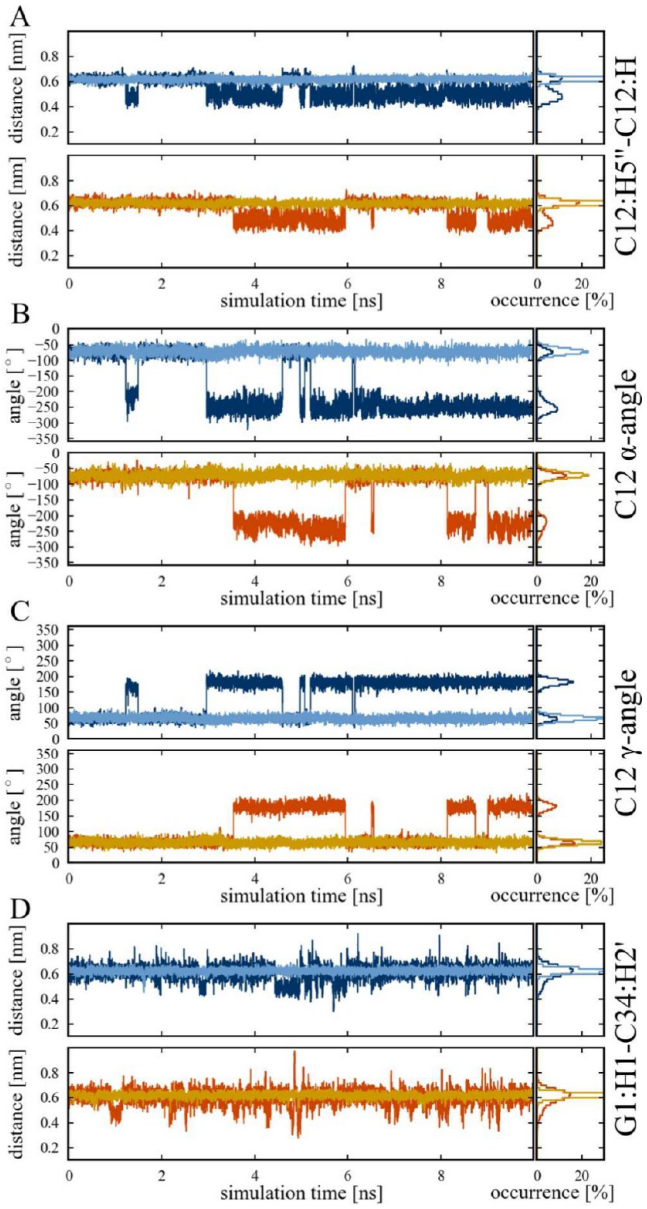


Figure 4. Distance and torsional-angle time series: (A) C12:H5″−C12:H5 distance, (B) α torsional angle in C12, (C) γ torsional angle in C12, and (D) G1:H1−C34:H2′ distance. Dark blue indicates the D5_xplor1_TAR simulation, light blue the D5_xplor1_IR simulation, orange the D5_xplor2_TAR simulation, and yellow the D5_xplor2_IR simulation.

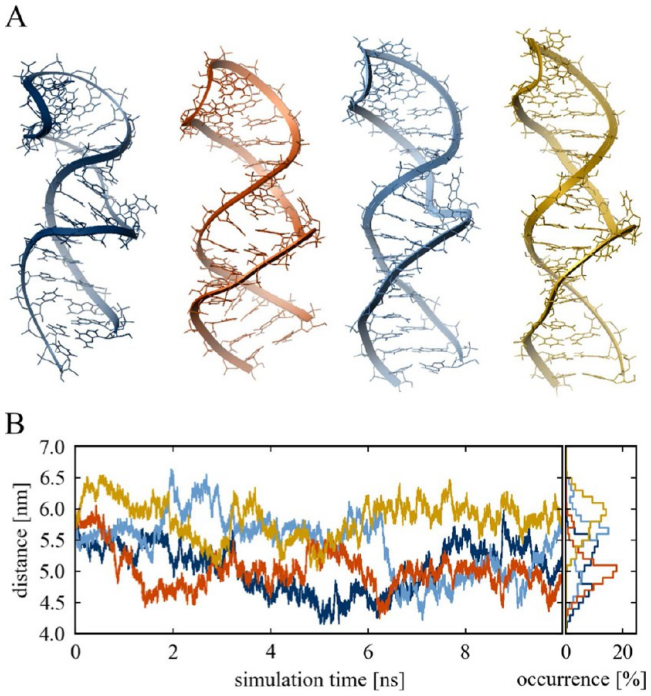


Figure 5. (A) Depicted from left to right are central members of the most populated conformational clusters of simulations D5_xplor1_-TAR, D5_xplor2_TAR, D5_xplor1_IR, and D5_xplor2_IR, respectively. A backbone heavy atom-positional rmsd of 0.2 nm has been used as a similarity criterion in the clustering. To avoid the bias from the initial structure, the first 2 ns of each trajectory was excluded from the clustering analysis. (B) Time traces (left) and distributions (right) of the end-to-end distance (distance between the center of geometry of the G2-C34 base pair and the phosphorus atom of G17) during MD simulations D5_xplor1_TAR (dark blue), D5_xplor1_IR (light blue), D5_xplor2_TAR (orange), and D5_xplor2_IR (yellow).

Table 1. Inter-Base-Pair Roll and Base-Pair Inclination (degrees) As Determined with Curves+[72] Using the Last 8 ns of the Simulations[a]

	inter-base-pair roll	inclination
D5_xplor1	3.9	5.6
D5_xplor2	4.4	9.0
D5_xplor1_IR	9.2	14.8
D5_xplor2_IR	6.8	13.0
D5_xplor1_TAR	14.3	19.9
D5_xplor2_TAR	12.5	22.1
idealized A-helix	12.6	22.6

[a]Reported are the average values for the two helical regions involving residues 1−8 and 28−35 and residues 11−14 and 20−23. A-Helical values were derived from an idealized dsRNA structure created by the fiber utility of 3DNA[81] that is based on Arnott's work on fiber A-DNA from calf thymus.[82]

the transition between two states. Changes from one sugar conformation to another are mainly observed in and close to the unpaired regions of the loop (Figure 6B,C) and bulge (Figure 6D) as well as from the GA region to the ends of the chain (data not shown). The sugar rings of some residues make only very short (picosecond) excursions away from a main conformation (Figure 6B), while others stay longer (nanoseconds) in different conformations (Figure 6C). Even though in the helical base-paired regions fluctuations are less pronounced, characteristic short sugar flips with a duration of a few picoseconds from C3′- to C2′-endo pucker are observed with time-averaged distance restraints (Figure 6A).

^{1}H−^{1}H TOCSY spectra indicate that both G17 and U18 in the AvD5 loop are predominantly in the C2′-endo conformation. In the simulations reported here, the U18 ribose shows short picosecond excursions toward the C3′-endo conformation (Figure 6B), while the G17 ribose explores different conformations on longer time scales (Figure 6C). Although the G17:H8−G17:H1′ cross-peak in NOESY spectra is exceptionally strong (corresponding to a distance of 0.25 nm), which is usually a sign of a nucleotide in a *syn* conformation, a pure *syn* conformation would not allow for the experimentally suggested G17:H8−G17:H3′ distance of 0.32 nm, indicating that conformational averaging takes place in

dx.doi.org/10.1021/bi400784r | *Biochemistry* 2013, 52, 7099−7113

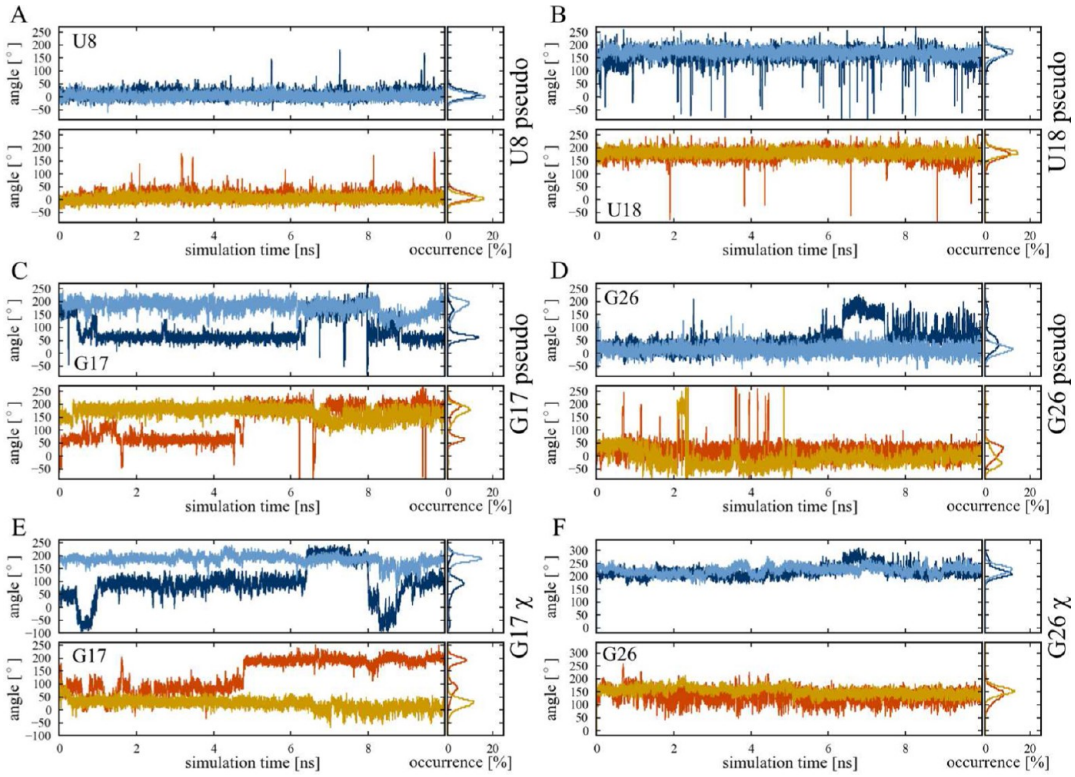


Figure 6. Time series of sugar pseudorotation (pseudo) and glycosidic (χ) angles in MD simulations initiated from structures D5_xplor1 [D5_xplor1_TAR (dark blue) and D5_xplor1_IR (light blue)] and D5_xplor2 [D5_xplor2_TAR (orange) and D5_xplor2_IR (yellow)].

the loop region. For this reason, no restraints of G17:H8 to any of its own sugar protons were included in the XPLOR calculations or in the MD simulations. While in the MD simulations with time-averaged distance restraints the G17 sugar pucker and glycosidic angle χ switch in a correlated manner and G17 is completely solvent-exposed 50−60% of the time, they remain rather constant in the MD simulations with instantaneous distance restraints (Figure 6C,E). In the latter case, G17 remains stacked throughout the simulations and the G17:H8−G17:H3′ distance is always longer than the NOESY experiment suggests (0.5−0.6 nm). Also, in the time-averaged simulations, this distance remains similarly long; however, occasionally conformations are sampled for which the distance decreases to 0.3 nm (Figure S6 of the Supporting Information, left panels).

In the structure refinement process leading to the ScD5 and PlD5 solution structures[12,13] in which base G26 is bulged down into the major groove, the conformation of G26 has been restrained to *syn* (70 ± 30°) on the basis of the intensity of the G26:H8−G26:H1′ NOE cross-peak. In the case of AvD5, the G26 χ angle was left unrestrained because the peak intensity, while stronger than in regular *anti*-oriented bases, indicated a distance of 0.3 nm, which is not as short as would be expected for a pure *syn* conformation. Interestingly, test calculations with Xplor-NIH in which the G26 χ angle of AvD5 was restrained to *syn* again resulted in both "G26 up" and "G26 down" structures indicating that the restraint of the G26 χ angle is not the reason for the lack of heterogeneity in ScD5 and PlD5. As is evident from Figure 6F in all four MD simulations at 293 K reported here, the G26 χ angle remains in the *anti* range, i.e., around 220° in simulations initiated from D5_xplor1 and around 150° in simulations initiated from D5_xplor2, resulting in G26:H8−G26:H1′ distances of 0.38 and 0.29 nm, respectively, which are in agreement with the experimental upper bound of 0.45 nm. Although the experimental evidence from ^{1}H−^{1}H TOCSY spectra suggests that the G26 sugar does not stay purely in a C2′- or C3′-endo conformation, our MD simulations depict it mainly in the C3′-endo conformation (Figure 6D).

Hydrogen Bonding. In Figure 7, the overall occurrence of Watson−Crick hydrogen bonds in the helical regions of the AvD5 hairpin is presented. The frequency of Watson−Crick hydrogen bonds at the end of the hairpin (base pairs G1-U35 and G2-C34) and close to the unpaired GA, bulge, and loop

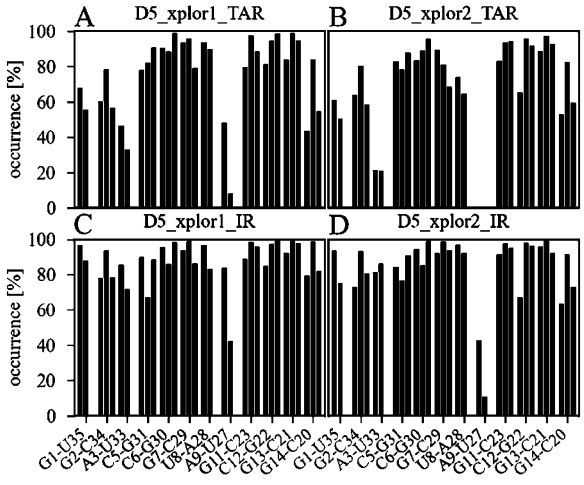


Figure 7. Occurrence of Watson−Crick base-pair hydrogen bonds in the helical regions of AvD5 during MD simulations D5_xplor1_TAR (A), D5_xplor2_TAR (B), D5_xplor1_IR (C), and D5_xplor2_IR (D).

dx.doi.org/10.1021/bi400784r | *Biochemistry* 2013, 52, 7099−7113

regions (A3-U33, C5-G31, A9-U27, and G14-C20) is higher in the instantaneously restrained MD simulations than in MD simulations in which time-averaged distance restraints have been used. In particular, for the A9-U27 base pair, the frequency of Watson−Crick hydrogen bonds strongly depends on the type of distance restraints applied as well as on the starting structure and thus on the conformation of G26. Fraying at the closing base pair (G1-U35) that can also be observed is not surprising considering that the closing base pair is a wobble one. In addition, the fraying of the closing base pair is in line with the severe broadening of G1 and U35 imino proton resonances in ^{1}H NMR spectra already at 275 K. Time series of non-Watson−Crick hydrogen bonds that occur in the non-helical regions of AvD5 are presented in Figure 8 for the loop, bulge, and GA regions.

In the pentaloop, the most populated hydrogen bonds (U19:H3−A16:N3, A16:H2′−U19:O2, U15:H3−C20:O2, and U15:H3−U19:O4) are present in all four simulations, except for the G17:H22−G17:O2P hydrogen bond, which occurs only in the D5_xplor2_IR simulation (Figure 8A). Interestingly, the time series of the hydrogen bonds in the loop show that dissolution of U15:H3−U19:O4, A16:H62−U15:O4, A16:H61−U18:O4, and U19:H3−U15:O2 hydrogen bonds is associated with the appearance of U19H3−A16:N3, A16:H2′−U19:O2, and U15:H3−C20:O2 hydrogen bonds indicating structural transitions between two different hydrogen bonding patterns of the AvD5 pentaloop.

In the bulge region and in particular in the region around the GA mismatch, the hydrogen bonding patterns of different simulations exhibit less overlap (Figure 8B). Prominent hydrogen bonds characteristic of the bulge region that occur in the D5_xplor2_TAR simulation as well as in the D5_xplor1_IR and D5_xplor2_IR simulations are the C25:H41−U10:O4, U10:H3−C25:N3, and U10:H3−C25:O2 bonds pairing C25 and U10. Hydrogen bonds formed by G26, A9:H61−G26:N7 and U10:H3−G26:O6, correspond to the G26 up conformation. In the G26 down conformation, the functional groups of the G26 base interact with the major groove of the helix between the catalytic triad and the bulge forming hydrogen bonds with C29 and A28 (data not shown). The orientation of the G26 down conformation is more similar to that in PlD5 [Protein Data Bank (PDB) entry 2F88[13] (Figure S2B of the Supporting Information)] than in ScD5 [PDB entry 1R2P[12] (Figure S2A of the Supporting Information)], where the Watson−Crick side of G26 points away from the helix.

In the region around the GA mismatch, the most prominent non-Watson−Crick hydrogen bonds are evident in the D5_xplor2_TAR simulation. The appearance of the G4:H21−A32:N1 hydrogen bond is related to the dissolution of the hydrogen bonding network formed by the A32:H61−G4:O6, G4:H21−U33:O2, G4:H1−U33:O2, and A3:H61−U33:O2 bonds, indicating structural transitions in the D5_xplor2_TAR simulation (Figure 8C).

Conformational Clustering Analysis. Structural transitions in the loop, bulge, and GA region indicated by hydrogen bonding analysis stimulated us to characterize the dominant conformations of these regions by performing a joined conformational clustering analysis of the D5_xplor1_IR, D5_xplor2_IR, D5_xplor1_TAR, and D5_xplor2_TAR trajectories using the atom-positional rmsd as the similarity criterion. The rmsd similarity cutoffs of 0.1, 0.12, and 0.15 nm have been used in the clustering analysis of the AvD5 loop, bulge, and GA region, respectively, to ensure that the first three clusters include a similar number (between 70 and 80% of the total population) of the trajectory configurations. To avoid the bias from the initial structure, the first 2 ns of each trajectory were excluded from the analysis.

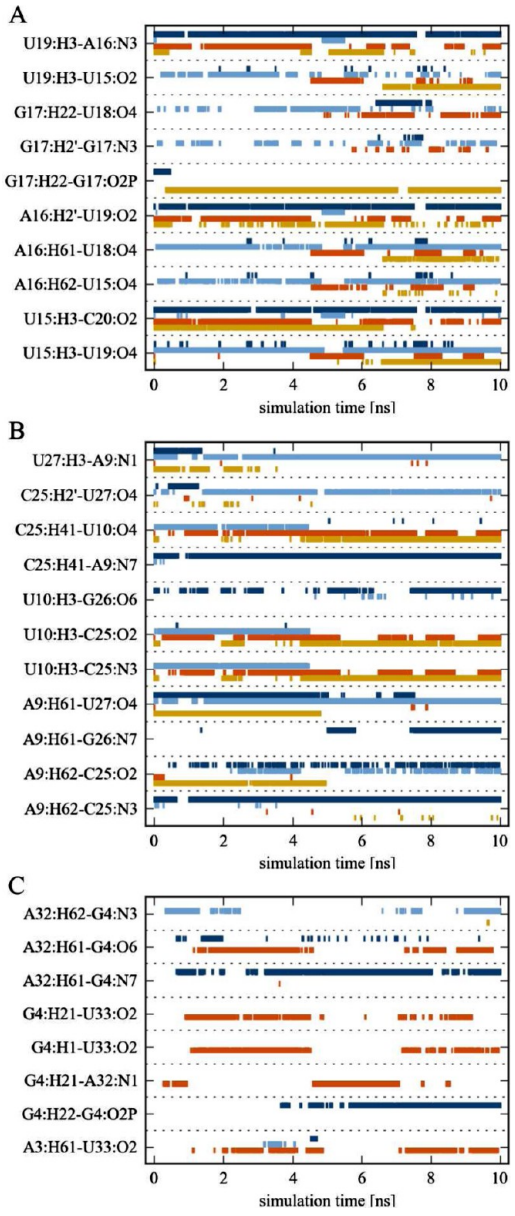


Figure 8. Time series of hydrogen bonds (A) in the loop (G14, U15, A16, G17, U18, U19, and C20), (B) in the bulge (A9, U10, A24, C25, G26, and U27), and (C) in the catalytic triad (A3, G4, C5, G31, A32, and U33) that are present for more than 20% in at least one of the four MD simulations. Hydrogen bonds are defined by a distance between the H and the acceptor of <0.25 nm and a donor···H···acceptor angle of >135°. Hydrogen bonds defining Watson−Crick base pairs G14-C20, A3-U33, and C5-G31 and hydrogen bonds between sugar−phosphate backbone atoms of adjacent residues were omitted for the sake of clarity. Dark blue for the D5_xplor1_TAR simulation, light blue for the D5_xplor1_IR simulation, orange for the D5_xplor2_-TAR simulation, and yellow for the D5_xplor2_IR simulation.

In the loop, the results of the conformational clustering when all the atoms of residues G14, U15, A16, G17, U18, U19, and C20 are included are largely determined by the orientation of the solvent-exposed G17 base. To analyze the subtle differences

dx.doi.org/10.1021/bi400784r | Biochemistry 2013, 52, 7099−7113

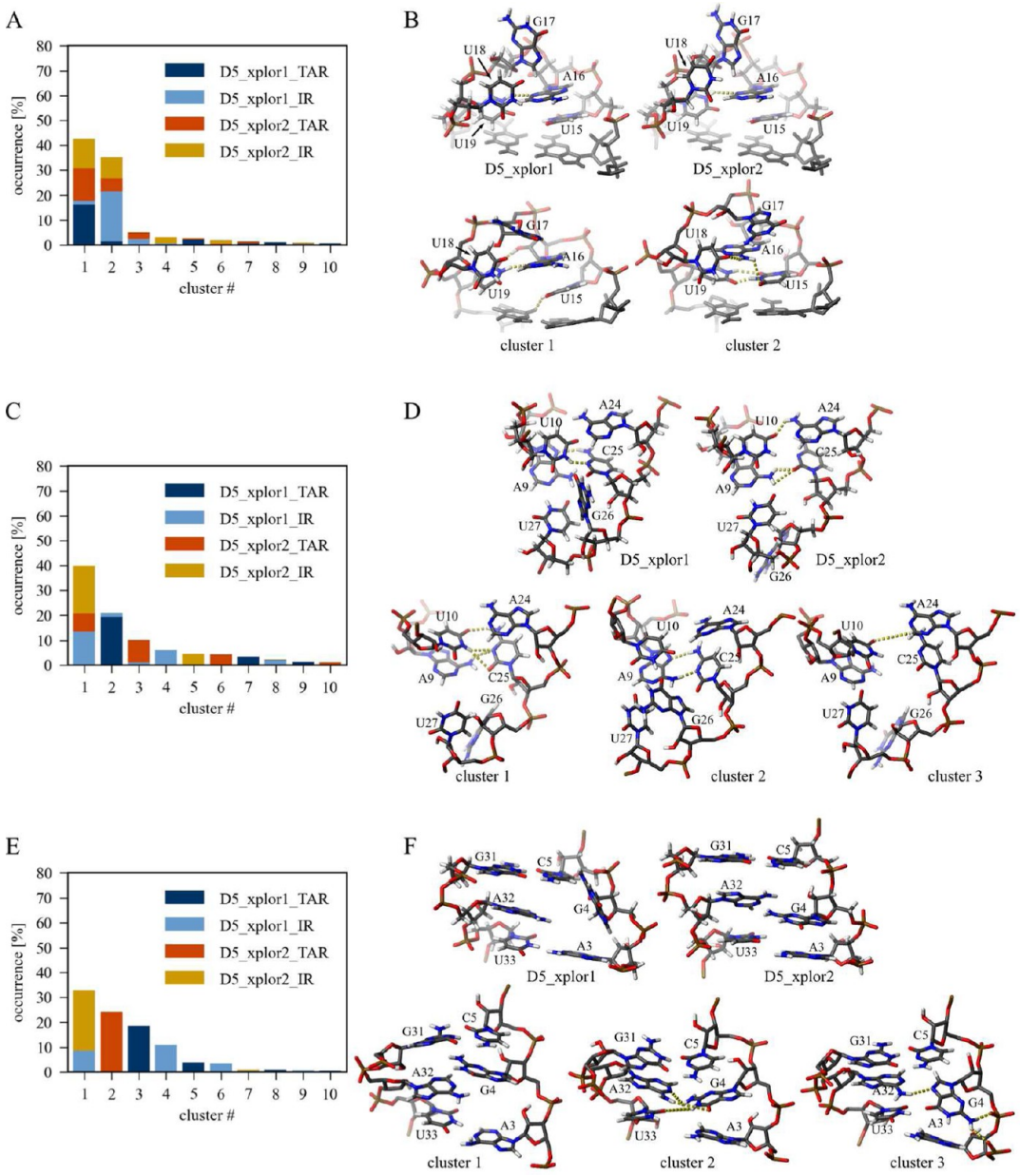


Figure 9. Results of joined conformational clustering analysis of the loop region (panels A and B, residues G14−C20, excluding G17), the bulge region (panels C and D, residues A9−G11 and C23−C25), and the catalytic triad region (panels E and F, residues A3−C5 and G31−U33) using D5_xplor1_TAR (dark blue), D5_xplor1_IR (light blue), D5_xplor2_TAR (orange), and D5_xplor2_IR (yellow) trajectories and a heavy atom-positional rmsd similarity cutoff of 0.1 nm. Panels A, C, and E show the frequency of the first 10 conformational clusters. In panels B, D, and F, conformations of the initial D5_xplor1 and D5_xplor2 structures and of the central member structures of the major clusters are depicted. Hydrogen bonds that occur more than 20% of the time (see also Tables S7−S9 of the Supporting Information) are indicated, excluding for the sake of clarity those defining G14-C20, A3-U33, and C5-G31 Watson−Crick base pairs and those between sugar−phosphate backbone atoms of adjacent residues. To avoid bias from the initial structure, the first 2 ns of each trajectory was excluded from the clustering analysis.

in the conformations of the loop in the four simulations, G17 was thus excluded from the clustering and hydrogen bonding analysis reported here. As shown in Figure 9A, two almost equally populated clusters, which both contain configurations sampled in all four simulations, emerge at the core of the loop. The analysis of hydrogen bonds of the ensemble of configurations belonging to cluster 1 and cluster 2 (Table S7 of the Supporting Information) agrees with the general hydrogen bonding analysis reported above (Figure 8A). Three dominant non-Watson−Crick hydrogen bonds, U19:H3−A16:N3, A16:H2′−U19:O2, and U15:H3−C20:O2, characteristic of cluster 1 are, in cluster 2, replaced by a network of four hydrogen bonds pairing U19 and U15 (U15:H3−U19:O4 and U19:H3−U15:O2), A16 and U18 (A16:H61−U18:O4), and A16 and U15 (A16:H62−U15:O4) (Figure 9B).

The results of the conformational clustering analysis in the AvD5 bulge are largely determined by the orientation of the G26 base. Because the interconversion between G26 up and G26 down conformations has not been observed in any of the MD simulations at 293 K, we have excluded G26 from the clustering analysis that thus involved residues A9, U10, G11, C23, A24, and C25. This allowed us to inspect the conformational overlap of the bulge region of the simulated ensembles independent of the G26 orientation. As is evident

dx.doi.org/10.1021/bi400784r | *Biochemistry* 2013, 52, 7099−7113

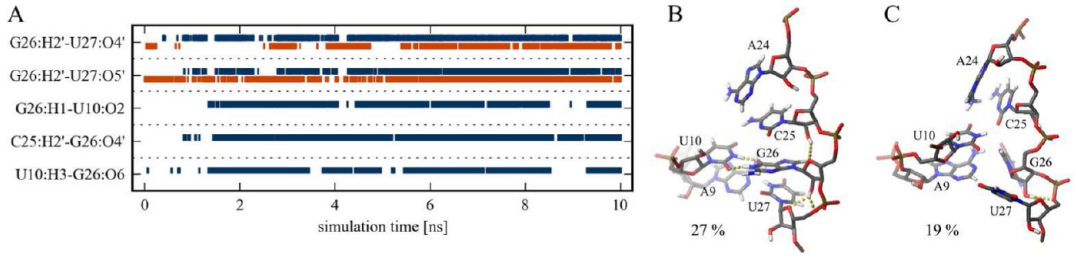


Figure 10. (A) Time series of hydrogen bonds involving G26 in the MD simulations at 323 K: dark blue for the D5_xplor1_TAR_323 simulation and orange for the D5_xplor2_TAR_323 simulation. (B) Central member structure of the most populated cluster in the simulation initiated from the D5_xplor1 structure and its frequency of occurrence. A heavy atom-positional rmsd similarity cutoff of 0.12 nm has been used. (C) Central member structure of the most populated cluster in the simulation initiated from D5_xplor2 and its frequency of occurrence. A heavy atom-positional rmsd similarity cutoff of 0.15 nm has been used. Hydrogen bonds formed by G26, whose time series are shown in panel A, are indicated.

from Figure 9C, the ensembles of D5_xplor1_IR, D5_xplor2_-IR, and D5_xplor2_TAR simulations share a significant part of the configurational space. Three characteristic hydrogen bonds already observed before, C25:H41−U10:O4, U10:H3−C25:O2, and U10:H3−C25:N3 (Figure 8B), occur in the bulge of cluster 1 and indicate the clear presence of a U10-C25 base pair (Table S8 of the Supporting Information). On the other hand, in cluster 2, which is dominated by configurations of the D5_xplor1_TAR simulation, a prominent hydrogen bonding interaction is formed between C25 and the Hoogsteen edge of A9 involving A9:H62−C25:N3 and C25:H41−A9:N7 hydrogen bonds. Cluster 3 is populated mainly with trajectory configurations of the D5_xplor2_TAR simulation. Here C25 is turned slightly away from the opposite RNA strand and forms a hydrogen bond with U10 (C25:H41−U10:O4) (Table S8 of the Supporting Information). While in the two XPLOR structures U10 is rather close to bonding with A24, our simulations suggest the formation of a U10-C25 base pair. This base pair has also been observed in the recently re-refined structures of ScD5 and PlD5,[38] where it alternates with the U10-A24 base pair. The Hoogsteen−Watson−Crick interaction between A9 and C25, which we also observe and which occurs only in cluster 2 (Figure 9D), might be characteristic of the G26 up conformation and would therefore not be observed in the refined structures of ScD5 and PlD5.

Clustering in the region around the G4-A32 mismatch leads to a conformational overlap only in the case of the simulations performed with instantaneous distance restraints (Figure 9E). The central member structure of cluster 1 features G4 stacked into the helix above A32, with no stable hydrogen bonding (Figure 9F). In cluster 2, populated only by configurations from the D5_xplor2_TAR simulation, G4 forms hydrogen bonds with U33 (G4:H1−U33:O2) and with A32 (G4:H21−A32:N1 and A32:H61−G4:O6) (Table S9 of the Supporting Information). Cluster 3 is dominated by configurations of the D5_xplor1_TAR simulation, which is the only simulation in which G4 adopts a *syn* conformation characterized by A32:H61−G4:N7 and G4:H22−G4:O2P hydrogen bonds (Table S9 of the Supporting Information).

Enhanced Conformational Sampling. In the MD simulations at 293 K reported so far, stable AvD5 conformations with an orientation of G26 either up into the minor groove or down into the major groove were found, and no transition between the two bulge conformations has been observed, which may be due to high energy barriers and a complex transition pathway separating the two states.

To enhance the conformational sampling of the bulge region of AvD5, we have performed two additional 10 ns MD simulations using a slightly elevated temperature of 323 K. The simulations were started from the D5_xplor1 and D5_xplor2 initial structures, and the experimental NOE upper bounds were imposed as time-averaged distance restraints to bias the sampling. The conformational space available to the bulge has been inspected using conformational clustering and hydrogen bond analysis with a focus on the orientation of G26.

While G26 remains in the major groove in the simulation initiated from the D5_xplor2 structure and does not form stable hydrogen bonds via base moieties, it undergoes a striking conformational change already within the first 2 ns of the simulation initiated from the D5_xplor1 structure (Figure 10). From its initial position in the minor groove, G26 moves toward the opposite strand and forms a GU wobble pair with U10 via U10:H3−G26:O6 and G26:H1−U10:O2 hydrogen bonds (Figure 10A). The observed change in the orientation of the G26 nucleobase suggests that the G26 up conformation of the AvD5 bulge region is rather unfavorable and may be an artifact of the initial structure refinement protocol.

Using an elevated temperature in the simulation also affected the conformational sampling in the AvD5 loop and GA mismatch region. In the loop, the enhanced sampling increases the number of conformational fluctuations of G17, generating an ensemble that is in agreement with the seemingly incompatible experimental NMR data discussed in the paragraph on the sugar and base flipping in the loop and bulge. In particular, in the case of the D5_xplor2_TAR_323 simulation, some of the conformations agree with the experimentally determined distance bound of 0.25 nm for the G17:H8−G17:H1′ bond and the others with the experimentally determined distance bound of 0.32 nm for the G17:H8−G17:H3′ bond (Figure S6 of the Supporting Information). This supports the observation from the clustering analysis that the conformational ensemble of the loop is dominated by more than one major conformation and cannot be described by a single structure.

In the GA mismatch region, we observe the adoption of a sheared GA base pair characterized by A32:H62−G4:N3 and G4:H22−A32:N7 hydrogen bonds (Figure S7 of the Supporting Information). At 293 K, this base pair was observed only in the D5_xplor1_IR simulation.

■ DISCUSSION

In this work, we describe structures of a bacterial group II intron domain 5 (AvD5) using solution NMR data and explicit solvent molecular dynamics simulations. To investigate the influence of averaging in structure refinement, the simulations were performed using instantaneous and time-averaged NOE-

derived distance restraints. In total, 60 ns of restrained MD simulations have been accumulated allowing a characterization of the structure and motions of the AvD5 hairpin with the emphasis on its nonhelical loop and bulge regions and on the GA mismatch in the catalytic triad.

UAGUU Loop Structure and Comparison with Other RNA Loops. In contrast to the abundant and well-characterized GAAA tetraloops in ScD5 and PlD5, the UAGUU loop in AvD5 is no common or conserved motif. Nonetheless, considering the conserved domain architecture of group II introns, it can be expected that this loop performs a structural role similar to that of the GAAA tetraloops in ScD5 and PlD5[73] forming a tertiary contact to domain 1 (D1). The minimal consensus sequence of the D5 loop in bacterial class E group II introns is NANNU (N stands for A, G, C, or U).[74] A survey of pentaloops with this consensus sequence using the RLooM database (http://rloom.mpimp-golm.mpg.de/[75]) reveals a handful of very different loop conformations, yet none of them have the structural features observed in the AvD5 pentaloop. On the other hand, the very conserved CAGUGC loop of the iron responsive element (IRE) hairpin shares remarkable structural features with the pentaloop of AvD5.[76] In the CAGUGC loop of IRE, C1 and G5 are base-paired and C6 is bulged out. The C1-G5 base pair in IRE corresponds to the U15-U19 base pair in the AvD5 loop (Figure 8B, cluster 2). The imino proton resonance of the C1-G5 base pair in IRE is broadened by exchange with the solvent, indicating that this base pair is not stably formed,[76] similar to the U15-U19 base pair in AvD5 where U19 can alternatively form hydrogen bonds also with A16. The sequence of the two loops is identical for A2, G3, and U4 (IRE) and A16, G17, and U18 (AvD5). In both loops, the adenine (A2 in IRE and A16 in AvD5) is stacked on its 5′-neighbor and is less flexible than the following G and U nucleobases.[76−78]

NMR and fluorescence experiments indicate that residues G3 and U4 in the IRE loop (G17 and U18 in AvD5) are more disordered than the other loop residues.[76,79] Their positions are therefore the least well-defined in the IRE loop structure.[76] Similarities to our AvD5 loop residues G17 and U18 are nonetheless apparent already from comparison of NMR observables. $^{3}J_{H1',H2'}$ coupling constants indicate that sugars of both G3 and U4 in the IRE loop are interconverting between C2′-endo and C3′-endo conformations or predominantly in the C2′-endo conformation,[76,78] which we also observe for G17 and U18 in AvD5.

Both loops also share an intense H8−H1′ peak for the central G residue (G3 in the IRE loop and G17 in AvD5), which is a strong sign of a *syn* conformation. However, while in the IRE loop the H8−H2′ cross-peak is only half as strong as the H8−H1′ cross-peak, which is expected for a *syn* conformation,[76] in AvD5 the H8−H2′ peak of G17 is as intense as the H8−H1′ peak. In addition, the seemingly inconsistent G17:H8−G17:H3′ and G17:H8−G17:H1′ NOE cross-peaks discussed above indicate that conformational averaging around the glycosidic angle is more pronounced in AvD5 than in the IRE loop.

The dynamic behavior of G4 (IRE) and G17 (AvD5) makes them an interesting subject for MD simulations. G4 has been studied in stochastic MD simulations of the IRE hairpin by Hall and Williams[79] and can be compared to G17 in our simulations. Hall and Williams observe the sugar of G4 (IRE) to repucker on a nanosecond time scale similar to that of G17 in AvD5.[79] The G4 glycosidic angle in the IRE hairpin loop remains mostly *syn*, with 100−500 ps excursions to the *anti* range, while in AvD5 simulations with time-averaged distance restraints, *anti* and *syn* ranges of the G17 glycosidic angle are more equally populated. Note that in the simulations with instantaneous distance restraints no interconversion between *syn* and *anti* has been observed: the G17 glycosidic angle remains *anti* in the simulations initiated from the D5_xplor1 structure and *syn* in the simulations initiated from the D5_xplor2 structure. Consequently, also the conformational averaging that is necessary to satisfy the incompatibly short measured sugar-to-base distances (G17:H8−H1′, −H2′, and −H3′, described in the previous paragraph) cannot be observed in the instantaneously restrained simulations. In the time-averaged simulations, on the other hand, each restraint is individually satisfied in at least one sampled conformation, even though the population satisfying the G17:H8−G17:H3′ distance is very small at 293 K. It appears that this minor conformation is sampled to a larger extent under our experimental conditions and is responsible for the observed sugar-to-base peak intensities in G17. The enhanced sampling at 323 K increases the size of the population of this conformation in the simulation started from the D5_xplor2 structure.

In light of the flexibility of the AvD5 pentaloop indicated by NMR experiments and MD simulations reported here, an adaptive rearrangement of the loop structure can occur when it binds to its receptor upon folding of the intron. The bound conformation of the IRE loop[80] reveals a conformational change in which both A2 and G3 bulge out to interact with the protein while the previously solvent-exposed U4 is stacked onto the closing base pair. It is tempting to hypothesize that binding of AvD5 to its receptor could induce a similar conformational change in the AvD5 pentaloop.

Orientation of G26 and Conformation of the Bulge. Structure determination of the flexible bulge region of AvD5 presents a challenge to both NMR experiments and MD simulations. Distance restraints derived from the NMR experiments in the bulge are too sparse to determine a unique conformation of this region. Unlike the previously reported structures of ScD5[12] and PlD5,[13] the initial structure refinement using the standard simulated annealing protocol in the gas phase resulted in two equally possible major conformations of the bulge, one with G26 oriented up to the minor groove (D5_xplor1) and one with G26 oriented down into the major groove (D5_xplor2). Both structures are compatible with all NOE bounds. Testing of the two structures against the published NOE bounds of ScD5[12] and PlD5[13] bulge regions reveals two NOE violations that are larger than 0.1 nm and involve the G26:H8−C25:H3′ bond, a NOE distance bound that is not present in the AvD5 experimental data set because of heavy spectral overlap, and the G26:H8−C25:H6 bond, for which in AvD5 a larger upper distance bound was used. Together with the fact that we did not restrain the glycosidic angle of G26 to the *syn* range and no restraints are present in our data set for the hydrogen bonds of the A9-U27 base pair, this offers an explanation for the G26 orientational ambiguity in AvD5, which has not been observed in other D5 structures.

To resolve the structure of the AvD5 bulge region and to account for the average nature of the measured NMR data, we performed end-stage refinement of the RNA AvD5 hairpin using molecular dynamics simulations with time-averaged distance restraints. While no transition between the two possible bulge conformations could be observed in MD

simulations at 293 K, in the simulation at 323 K initiated from the D5_xplor1 structure there is a rearrangement of the bulge region. A rapid conversion of G26 to a stacked-in conformation base-paired with U10 suggests that the solvent-exposed position in the minor groove was an unfavorable, artificial configuration. Note that the G26-U10 base pair adopted in this simulation is formed in all D5 crystal structures.[23,24,27,28] However, all NMR studies so far clearly indicate that this base pair is not stably formed in solution.[12,13] The observed rearrangement should therefore be interpreted as a movement away from a less favorable conformation in the minor groove rather than as a conversion to the crystal structure conformation. In contrast, in the simulations initiated from the D5_xplor2 structure, in which G26 is turned down to the major groove in agreement with the previous solution NMR model structures, the G26 region was not much affected by the elevated temperature.

Detailed analysis of the G26:H8−G26:H1′ distance additionally supports the preference of the AvD5 bulge for the D5_xplor2 conformation. Inspection of the NOE cross-peak intensity of the G26:H8−G26:H1′ bond reveals a distance close to 0.3 nm that is the borderline distance at which the signal is characterized as having a very strong (<0.3 nm) or strong (<0.45 nm) intensity. To allow for the uncertainty of the distance measurement because of the partial overlap of the G26:H8−G26:H1′ peak in NOESY spectra and to avoid over-restraining, the upper distance bound was set to 0.45 nm. The analysis of the D5_xplor1_IR and D5_xplor1_TAR simulations shows that in these simulations the G26:H8−G26:H1′ distance remains close to 0.38 nm whereas it fluctuates around a value of 0.28−0.29 nm in D5_xplor2_IR and D5_xplor2_TAR simulations. The simulations starting from the D5_xplor2 structure with G26 in the major groove thus agree better with the measured NMR data.

Unlike in our study, a recent re-refinement of the PlD5 and ScD5 NMR structures[38] included also dihedral-angle restraints. The glycosidic angle of G26 was restrained to the *syn* range ($\chi = 70 \pm 30°$) on the basis of the strong G26:H8−G26:H1′ NOE cross-peak. Interestingly, despite the restraining, Henriksen et al.[38] describe a frequent flipping of G26 to an *anti* conformation, which could only be prevented by considerably increasing the weight of the dihedral-angle restraint. Being unrestrained in our simulations initiated from D5_xplor2, where it protrudes into the major groove like in PlD5 and ScD5, G26 prefers a conformation with a χ of 150°, which is still in the general *anti* range. Although this value is far outside the restrained range used in the structure calculations of PlD5 and ScD5 ($\chi = 70 \pm 30°$),[12,13,38] it nonetheless results in a short (<0.3 nm) G26:H8−G26:H1′ distance that satisfies not only the upper bound used in our study (0.45 nm) but also those found in the NOE distance bound lists of ScD5 (0.45 nm, PDB entry 1R2P) and PlD5 (0.5 nm, PDB entry 2F88). In cases such as these, we would very much caution against the use of restraints for the glycosidic angle.

The example of the D5 bulge shows how very small variations in restraints can be responsible for a conformational ambiguity that cannot be resolved even after MD simulation for several nanoseconds because the barrier between the two conformations is too high. Nonetheless, extended MD simulations do provide us with a more comprehensive picture for critically evaluating the two conformations and identifying one solution as being more favorable.

Conformational Variability in the Catalytic Triad. Clustering and hydrogen bond analysis of the four restrained MD simulations at 293 K indicate many different conformations with G4 in *syn* and *anti* conformations making various interactions with A32 and with nucleotides of the neighboring base pairs (Figures 8C and 9E,F).

However, there is very little overlap between conformations sampled in the different simulations, indicating that the simulation time of 10 ns is too short to observe convergence. With an increase in the temperature to 323 K, the sampling was enhanced and a sheared GA mismatch in two simulations with time-averaged distance restraints has been formed (Figure S7 of the Supporting Information). This base pair is consistent with experimental distance bounds and might represent a dominant conformation in the AvD5 conformational ensemble present in solution.

Such conformational variability in the catalytic triad region as in AvD5 is not observed in the ScD5 or PlD5 solution structures,[12,13] which contain a G-U base pair instead of a G-A base pair. In the crystal structures of the *O. iheyensis* intron, only the G and not its opposite base is involved in a tertiary interaction and its backbone is part of the coordination sphere of a Mg(II) ion involved in catalysis.[24,27,28] While flexibility in this region would definitely influence the catalytic propensity, it is not expected that the dynamics in the complex of the folded intron, when G4 is involved in a tertiary contact, would be the same as in solution.

Time-Averaged Distance Restraints Improve the Sampling and Accelerate Relaxation toward Standard A-Form Helical Base-Pair Parameters. Experimental measurements of NMR observables provide time- and ensemble-averaged values. Forcing a single structure to satisfy all the experimental data simultaneously can thus lead to a representation of the system that is too rigid and very likely unrealistic. This problem can be avoided by introducing time-averaged restraining in the NMR structure refinement protocols. In MD simulations reported here, 668 NOE upper distance bounds were imposed as either instantaneous or time-averaged distance restraints to bias the sampling of the AvD5 conformational space. The comparison of the trajectories generated using time-averaged distance restraints with the trajectories generated using conventional instantaneous distance restraints indicates that the former increase the internal mobility of the AvD5 hairpin leading to a larger number of conformational transitions as illustrated by the analysis of the time series of angles and distances depicted in Figures 4 and 6.

In addition, inspection of the persistence of Watson−Crick base pairing, global structure elongation, and helical parameters indicates that the trajectory structures stay much closer to the initial XPLOR structures when generated using instantaneous distance restraints. All the base pairs, except those belonging to the stable helical regions (C6-U8 and G11-G13), occur to a greater extent in the simulations with instantaneous distance restraints (Figure 7). Furthermore, the length of the hairpin measured from G1 to G17 in the simulations with instantaneous distance restraints remains the same as in the initial XPLOR structures (6 nm), while it decreases by more than 1 nm in the simulations using time-averaged distance restraints (Figure 5). Because NOE restraints alone are too short-range for defining global features of an elongated structure like an RNA hairpin, the observed global structural changes originate from subtle effects that may occur at each nucleotide step. In the case of AvD5, we identified two helix

parameters that reflect the observed global compaction. Inter-base-pair roll and base-pair inclination are unusually small in the XPLOR structures, which might be a result of the absence of solvent and proper electrostatics in the initial structure refinement. In the course of MD simulations reported here, these values move toward the standard A-helical values, in particular when time-averaged distance restraints are used (Table 1).

CONCLUSION

Through the example of AvD5, we show the impact of electrostatics and an explicit solvent on the conformational ensemble of an RNA hairpin. Locally, the molecular dynamics simulations described in this work reproduce fast conformational exchange processes suggested by experimental data that were not obvious in the models derived by conventional single-structure simulated annealing refinement. Importantly, these exchange processes are observed only when time-averaged instead of instantaneous distance restraints are employed. In addition, the simulated helical parameter values do agree better with those for ideal A-form helices when time averaging is applied. The use of time-averaged distance restraining in the refinement of NMR structure models leads to a more realistic sampling of the conformational ensemble.

ASSOCIATED CONTENT

Supporting Information

Conformations of the bulge in crystal and solution structures (Figure S1), sequence and solution structure of ScD5 and PlD5 (Figure S2), representation of torsional angles in a polynucleotide chain (Figure S3), visualization of NOE restraints in the three nonhelical regions of AvD5 (Figure S4), NOE distance bound violations in starting structures D5_xplor1 and D5_xplor2 and in MD simulations D5_xplor1_TAR, D5_xplor1_IR, D5_xplor2_TAR, and D5_xplor2_IR (Figure S5), time series of particular sugar−base distances in loop residue G17 at 293 and 323 K (Figure S6), time series of hydrogen bonds in the GA mismatch region in MD simulations at 323 K (Figure S7), atom names, masses, GROMOS integer atom codes describing the Lennard-Jones parameters, and partial charges for adenine, guanine, cytosine, and uracil (Tables S1−S4, respectively), number of experimentally collected NOE restraints sorted by type and region (Table S5), NOE distances and distance bounds (Table S6), occurence of hydrogen bonds in the two most populated conformational clusters in the loop region (Table S7), occurrence of hydrogen bonds in the three most populated conformational clusters in the bulge region (Table S8), and occurrence of hydrogen bonds in the three most populated conformational clusters in the GA mismatch region (Table S9). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, CH-8093 Zurich, Switzerland. Phone: +41 44 633 4239. E-mail: jozi@igc.phys.chem.ethz.ch.

Funding

This work was financially supported by the National Center of Competence in Research (NCCR) in Structural Biology, the University of Zurich, the SERI (COST Action CM1105), Grants 200020-137827 (to W.F.v.G.) and 200020-143750 (to R.K.O.S.) of the Swiss National Science Foundation, and Grants 228076 (to W.F.v.G.) and 259092 (to R.K.O.S.) of the European Research Council, which is gratefully acknowledged.

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

MD, molecular dynamics; NMR, nuclear magnetic resonance; IR, instantaneous restraining; TAR, time-averaged restraining.

REFERENCES

(1) Strobel, S. A., and Cochrane, J. C. (2007) RNA catalysis: Ribozymes, ribosomes, and riboswitches. *Curr. Opin. Chem. Biol. 11*, 636−643.

(2) Scott, W. G., Martick, M., and Chi, Y.-I. (2009) Structure and function of regulatory RNA elements: Ribozymes that regulate gene expression. *Biochim. Biophys. Acta 1789*, 634−641.

(3) Guo, P. (2010) The emerging field of RNA nanotechnology. *Nat. Nanotechnol. 5*, 833−842.

(4) Burnett, J. C., and Rossi, J. J. (2012) RNA-based therapeutics: Current progress and future prospects. *Chem. Biol. 19*, 60−71.

(5) Butcher, S. E., and Pyle, A. M. (2011) The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc. Chem. Res. 44*, 1302−1311.

(6) Davis, J. H., Tonelli, M., Scott, L. G., Jaeger, L., Williamson, J. R., and Butcher, S. E. (2005) RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex. *J. Mol. Biol. 351*, 371−382.

(7) Zhang, Q., Stelzer, A. C., Fisher, C. K., and Al-Hashimi, H. M. (2007) Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature 450*, 1263−1267.

(8) Haller, A., Soulière, M. F., and Micura, R. (2011) The dynamic nature of RNA as key to understanding riboswitch mechanisms. *Acc. Chem. Res. 44*, 1339−1348.

(9) Frank, A. T., Stelzer, A. C., Al-Hashimi, H. M., and Andricioaei, I. (2009) Constructing RNA dynamical ensembles by combining MD and motionally decoupled NMR RDCs: New insights into RNA dynamics and adaptive ligand recognition. *Nucleic Acids Res. 37*, 3670−3679.

(10) Lehmann, K., and Schmidt, U. (2003) Group II introns: Structure and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol. Biol. 38*, 249−303.

(11) Michels, W. J., and Pyle, A. M. (1995) Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: Elucidation of reaction mechanism and structure/function relationships. *Biochemistry 34*, 2965−2977.

(12) Sigel, R. K. O., Sashital, D. G., Abramovitz, D. L., Palmer, A. G., III, Butcher, S. E., and Pyle, A. M. (2004) Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif. *Nat. Struct. Mol. Biol. 11*, 187−192.

(13) Seetharaman, M., Eldho, N. V., Padgett, R. A., and Dayie, K. T. (2006) Structure of a self-splicing group II intron catalytic effector domain 5: Parallels with spliceosomal U6 RNA. *RNA 12*, 235−247.

(14) Michel, F., and Ferat, J.-L. (1995) Structure and activities of group II introns. *Annu. Rev. Biochem. 64*, 435−461.

(15) Michel, F., Umesono, K., and Ozeki, H. (1989) Comparative and functional anatomy of group II catalytic introns: A review. *Gene 82*, 5−30.

(16) Boulanger, S. C., Belcher, S. M., Schmidt, U., Dib-Hajj, S. D., Schmidt, T., and Perlman, P. S. (1995) Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. *Mol. Cell. Biol. 15*, 4479−4488.

(17) Abramovitz, D. L., Friedman, R. A., and Pyle, A. M. (1996) Catalytic role of 2′-hydroxyl groups within a group II intron active site. *Science 271*, 1410−1413.

(18) Schmidt, U., Podar, M., Stahl, U., and Perlman, P. S. (1996) Mutations of the two-nucleotide bulge of D5 of a group II intron block

splicing in vitro and in vivo: Phenotypes and suppressor mutations. *RNA 2*, 1161−1172.
(19) Boudvillain, M., and Pyle, A. M. (1998) Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: A NAIM analysis. *EMBO J. 17*, 7091−7104.
(20) Costa, M., and Michel, F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J. 14*, 1276−1285.
(21) Shukla, G. C., and Padgett, R. A. (2002) A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. *Mol. Cell 9*, 1145−1150.
(22) Keating, K. S., Toor, N., Perlman, P. S., and Pyle, A. M. (2010) A structural analysis of the group II intron active site and implications for the spliceosome. *RNA 16*, 1−9.
(23) Zhang, L., and Doudna, J. A. (2002) Structural insights into group II intron catalysis and branch-site selection. *Science 295*, 2084−2088.
(24) Toor, N., Keating, K. S., Taylor, S. D., and Pyle, A. M. (2008) Crystal structure of a self-spliced group II intron. *Science 320*, 77−82.
(25) Toor, N., Rajashankar, K., Keating, K. S., and Pyle, A. M. (2008) Structural basis for exon recognition by a group II intron. *Nat. Struct. Mol. Biol. 15*, 1221−1222.
(26) Wang, J. (2010) Inclusion of weak high-resolution X-ray data for improvement of a group II intron structure. *Acta Crystallogr. D66*, 988−1000.
(27) Chan, R. T., Robart, A. R., Rajashankar, K. R., Pyle, A. M., and Toor, N. (2012) Crystal structure of a group II intron in the pre-catalytic state. *Nat. Struct. Mol. Biol. 19*, 555−557.
(28) Marcia, M., and Pyle, A. M. (2012) Visualizing group II intron catalysis through the stages of splicing. *Cell 151*, 497−507.
(29) Eldho, N. V., and Dayie, K. T. (2007) Internal bulge and tetraloop of the catalytic domain 5 of a group II intron ribozyme are flexible: Implications for catalysis. *J. Mol. Biol. 365*, 930−944.
(30) Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905−921.
(31) Chang, A. T., and Nikonowicz, E. P. (2012) Solution nuclear magnetic resonance analyses of the anticodon arms of proteinogenic and nonproteinogenic tRNAGly. *Biochemistry 51*, 3662−3674.
(32) Levengood, J. D., Rollins, C., Mishler, C. H. J., Johnson, C. A., Miner, G., Rajan, P., Znosko, B. M., and Tolbert, B. S. (2012) Solution structure of the HIV-1 exon splicing silencer 3. *J. Mol. Biol. 415*, 680−698.
(33) Skov, J., Gaudin, M., Podbevsek, P., Olsthoorn, R. C. L., and Petersen, M. (2012) The subgenomic promoter of brome mosaic virus folds into a stem-loop structure capped by a pseudo-triloop that is structurally similar to the triloop of the genomic promoter. *RNA 18*, 992−1000.
(34) Zhao, Q., Huang, H.-C., Nagaswamy, U., Xia, Y., Gao, X., and Fox, G. E. (2012) UNAC tetraloops: To what extent do they mimic GNRA tetraloops? *Biopolymers 97*, 617−628.
(35) Padrta, P., Stefl, R., Kralik, L., Zidek, L., and Sklenar, V. (2002) Refinement of d(GCGAAGC) hairpin structure using one- and two-bond residual dipolar couplings. *J. Biomol. NMR 24*, 1−14.
(36) Stefl, R., Oberstrass, F. C., Hood, J. L., Jourdan, M., Zimmermann, M., Skrisovska, L., Maris, C., Peng, L., Hofr, C., Emeson, R. B., and Allain, F. H. T. (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell 143*, 225−237.
(37) Lerman, Y. V., Kennedy, S. D., Shankar, N., Parisien, M., Major, F., and Turner, D. H. (2011) NMR structure of a 4 × 4 nucleotide RNA internal loop from an R2 retrotransposon: Identification of a three purine-purine sheared pair motif and comparison to MC-SYM predictions. *RNA 17*, 1664−1677.
(38) Henriksen, N., Davis, D., and Cheatham, T., III (2012) Molecular dynamics re-refinement of two different small RNA loop structures using the original NMR data suggest a common structure. *J. Biomol. NMR*, 1−19.
(39) Kruschel, D., and Zagrovic, B. (2009) Conformational averaging in structural biology: Issues, challenges and computational solutions. *Mol. BioSyst. 5*, 1606−1616.
(40) Dolenc, J., Missimer, J. H., Steinmetz, M. O., and van Gunsteren, W. F. (2010) Methods of NMR structure refinement: Molecular dynamics simulations improve the agreement with measured NMR data of a C-terminal peptide of GCN4-p1. *J. Biomol. NMR 47*, 221−235.
(41) Bonvin, A. M. J. J., Boelens, R., and Kaptein, R. (1994) Time-averaged and ensemble-averaged direct NOE restraints. *J. Biomol. NMR 4*, 143−149.
(42) Markwick, P. R. L., and Nilges, M. (2012) Computational approaches to the interpretation of NMR data for studying protein dynamics. *Chem. Phys. 396*, 124−134.
(43) Bonvin, A. M. J. J., and Brunger, A. T. (1995) Conformational variability of solution nuclear-magnetic-resonance structures. *J. Mol. Biol. 250*, 80−93.
(44) Daura, X., Antes, I., van Gunsteren, W. F., Thiel, W., and Mark, A. E. (1999) The effect of motional averaging on the calculation of NMR-derived structural properties. *Proteins: Struct., Funct., Genet. 36*, 542−555.
(45) Pearlman, D. A., and Kollman, P. A. (1991) Are time-averaged restraints necessary for nuclear-magnetic-resonance refinement: A model study for DNA. *J. Mol. Biol. 220*, 457−479.
(46) Isaacs, R. J., and Spielmann, H. P. (2004) Insight into G-T mismatch recognition using molecular dynamics with time-averaged restraints derived from NMR spectroscopy. *J. Am. Chem. Soc. 126*, 583−590.
(47) Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. (2006) Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc. 48*, 47−62.
(48) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson. 160*, 65−73.
(49) Kao, C., Zheng, M., and Rüdisser, S. (1999) A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA 5*, 1268−1272.
(50) Gallo, S., Furler, M., and Sigel, R. K. O. (2005) In vitro transcription and purification of RNAs of different size. *Chimia 59*, 812−816.
(51) Luy, B., and Marino, J. P. (2000) Direct evidence for Watson-Crick base pairs in a dynamic region of RNA structure. *J. Am. Chem. Soc. 122*, 8095−8096.
(52) Zwahlen, C., Legault, P., Vincent, S. J. F., Greenblatt, J., Konrat, R., and Kay, L. E. (1997) Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage λ N-Peptide/boxB RNA complex. *J. Am. Chem. Soc. 119*, 6711−6721.
(53) Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wüthrich, K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Mol. Biol. 280*, 933−952.
(54) Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol. 273*, 283−298.
(55) Brünger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat. Protoc. 2*, 2728−2733.
(56) Scott, W. R. P., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Krüger, P., and van Gunsteren, W. F. (1999) The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A 103*, 3596−3607.
(57) Christen, M., Hünenberger, P. H., Bakowies, D., Baron, R., Bürgi, R., Geerke, D. P., Heinz, T. N., Kastenholz, M. A., Kräutler, V., Oostenbrink, C., Peter, C., Trzesniak, D., and van Gunsteren, W. F. (2005) The GROMOS software for biomolecular simulation: GROMOS05. *J. Comput. Chem. 26*, 1719−1751.

(58) Schmid, N., Allison, J. R., Dolenc, J., Eichenberger, A. P., Kunz, A.-P. E., and van Gunsteren, W. F. (2011) Biomolecular structure refinement using the GROMOS simulation software. *J. Biomol. NMR 51*, 265−281.
(59) Horta, B. A. C., Fuchs, P. F. J., van Gunsteren, W. F., and Hünenberger, P. H. (2011) New interaction parameters for oxygen compounds in the GROMOS force field: Improved pure-liquid and solvation properties for alcohols, ethers, aldehydes, ketones, carboxylic acids, and esters. *J. Chem. Theory Comput. 7*, 1016−1031.
(60) Horta, B. A. C., Lin, Z., Huang, W., Riniker, S., van Gunsteren, W. F., and Hünenberger, P. H. (2012) Reoptimized interaction parameters for the peptide-backbone model compound N-methyl-acetamide in the GROMOS force field: Influence on the folding properties of two β-peptides in methanol. *J. Comput. Chem. 33*, 1907−1917.
(61) Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981) Interaction models for water in relation to protein hydration. In *Intermolecular Forces* (Pullman, B., Ed.) pp 331−342, Reidel, Dordrecht, The Netherlands.
(62) Torda, A. E., Scheek, R. M., and van Gunsteren, W. F. (1989) Time-dependent distance restraints in molecular dynamics simulations. *Chem. Phys. Lett. 157*, 289−294.
(63) Nanzer, A. P., Huber, T., Torda, A. E., and van Gunsteren, W. F. (1996) Molecular dynamics simulation using weak-coupling NOE distance restraining. *J. Biomol. NMR 8*, 285−291.
(64) Nanzer, A. P., van Gunsteren, W. F., and Torda, A. E. (1995) Parametrization of time-averaged distance restraints in MD simulations. *J. Biomol. NMR 6*, 313−320.
(65) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical-integration of cartesian equations of motion of a system with constraints: Molecular-dynamics of n-alkanes. *J. Comput. Phys. 23*, 327−341.
(66) Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular-dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684−3690.
(67) Tironi, I. G., Sperb, R., Smith, P. E., and van Gunsteren, W. F. (1995) A generalized reaction field method for molecular-dynamics simulations. *J. Chem. Phys. 102*, 5451−5459.
(68) Heinz, T. N., van Gunsteren, W. F., and Hünenberger, P. H. (2001) Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys. 115*, 1125−1136.
(69) Kearsley, S. K. (1989) On the orthogonal transformation used for structural comparisons. *Acta Crystallogr. A45*, 208−210.
(70) van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P., Mark, A. E., Scott, W. R. P., and Tironi, I. G. (1996) *Biomolecular simulation: The GROMOS96 manual and userguide*, Hochschuleverlag AG an der ETH, Zurich.
(71) Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 29−32 , 51−55.
(72) Lavery, R., Moakher, M., Maddocks, J. H., Petkeviciute, D., and Zakrzewska, K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res. 37*, 5917−5929.
(73) Reiter, N. J., Chan, C. W., and Mondragon, A. (2011) Emerging structural themes in large RNA molecules. *Curr. Opin. Struct. Biol. 21*, 319−326.
(74) Candales, M. A., Duong, A., Hood, K. S., Li, T., Neufeld, R. A. E., Sun, R., McNeil, B. A., Wu, L., Jarding, A. M., and Zimmerly, S. (2012) Database for bacterial group II introns. *Nucleic Acids Res. 40*, D187−D190.
(75) Schudoma, C., May, P., Nikiforova, V., and Walther, D. (2010) Sequence-structure relationships in RNA loops: Establishing the basis for loop homology modeling. *Nucleic Acids Res. 38*, 970−980.
(76) Addess, K. J., Basilion, J. P., Klausner, R. D., Rouault, T. A., and Pardi, A. (1997) Structure and dynamics of the iron responsive element RNA: Implications for binding of the RNA by iron regulatory binding proteins. *J. Mol. Biol. 274*, 72−83.
(77) Hall, K. B., and Tang, C. (1998) ^{13}C relaxation and dynamics of the purine bases in the iron responsive element RNA hairpin. *Biochemistry 37*, 9323−9332.
(78) Laing, L. G., and Hall, K. B. (1996) A model of the iron responsive element RNA hairpin loop structure determined from NMR and thermodynamic data. *Biochemistry 35*, 13586−13596.
(79) Hall, K. B., and Williams, D. J. (2004) Dynamics of the IRE RNA hairpin loop probed by 2-aminopurine fluorescence and stochastic dynamics simulations. *RNA 10*, 34−47.
(80) Walden, W. E., Selezneva, A. I., Dupuy, J., Volbeda, A., Fontecilla-Camps, J. C., Theil, E. C., and Volz, K. (2006) Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. *Science 314*, 1903−1908.
(81) Lu, X.-J., and Olson, W. K. (2008) 3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc. 3*, 1213−1227.
(82) Arnott, S. (1999) Polynucleotide secondary structures: An historical perspective. In *Oxford Handbook of Nucleic Acid Structure* (Neidle, S., Ed.) pp 1−38, Oxford University Press, Oxford, U.K.