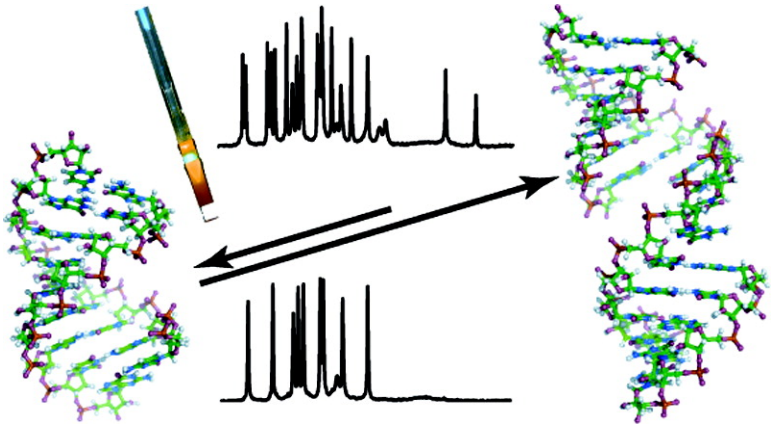
Article

Conformational Dynamics of Bistable RNAs Studied by Time-Resolved NMR Spectroscopy

Boris Frtig, Philipp Wenter, Luc Reymond, Christian Richter, Stefan Pitsch, and Harald Schwalbe

J. Am. Chem. Soc., **2007**, 129 (51), 16222-16229 • DOI: 10.1021/ja076739r



J|A|C|S
ARTICLES
Published on Web 11/30/2007

Conformational Dynamics of Bistable RNAs Studied by Time-Resolved NMR Spectroscopy

Boris Fürtig,[†] Philipp Wenter,[‡] Luc Reymond,[‡] Christian Richter,[†] Stefan Pitsch,[‡] and Harald Schwalbe*,[†]

Contribution from the Johann Wolfgang Goethe University, Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance, Max von Laue-Str. 7, 60438 Frankfurt, Germany, and École Polytechnique Fédérale de Lausanne Laboratory of Nucleic Acid Chemistry, EPFL-BCH 1015 Lausanne, Switzerland

Received September 6, 2007; E-mail: schwalbe@nmr.uni-frankfurt.de

Abstract: The structural transition between two alternate conformations of bistable RNAs has been characterized by time-resolved NMR spectroscopy. The mechanism, kinetics, and thermodynamics underlying the global structural transition of bistable RNAs were delineated. Both bistable RNA conformations and a partial unstructured RNA of identical sequence could be trapped using photolabile protecting groups. This trapping allowed for an investigation of the initial folding from an unfolded RNA to one of the preferred conformations of the bistable RNA and of the structural transitions involved. Folding of the secondary structure elements occurs rapidly, while the global structural transition of the bistable RNA occurs on a time scale of minutes and shows marked temperature dependence. Comparison of these results with bistable systems previously investigated leads to the prediction of activation enthalpies ($\Delta H^{\ddagger}$) associated with global structural transitions in RNA.

Introduction

In the 1970s, rapid temperature-jump experiments were developed in an effort to investigate the formation of secondary structure in DNA and RNA. Based on these experiments, it was concluded that RNA folding involves a process beginning from a completely unfolded state and proceeds toward a final, stable state represented by a single conformation.[1,2] In general, RNA folding is considered to follow a hierarchical pathway where rapid formation of a secondary structure precedes the slower distinct formation of tertiary elements.[3] However, folding studies of structural elements in the *Tetrahymena* ribozyme revealed that, at least for relatively large-size RNA, folding cannot be adequately represented by a purely hierarchical mechanism, but secondary and tertiary folding can be coupled.[4] Additionally, RNA folding may involve the formation of different functional states in the cell, and the structural transitions between those states are functionally relevant.[5] During transcription, the nascent RNA chain can adopt a number of different conformations that represent energetically favorable minima[6] on a rugged energy hypersurface.[7] During transcription, the initially sampled conformation(s) of growing RNA fragments may have little in common with the conformation of the full-length functional RNA form but may act as a kinetic trap.[8−11] *In vivo*, RNA structural transitions are often linked to the cellular functions of RNA.[12] One important example comprises RNA riboswitches, structural elements located in the 5′-untranslated region of bacterial mRNAs consisting of an aptamer domain that changes conformation upon ligand binding. This structural transition is coupled to structural rearrangements that regulate gene expression.[13−15] Further examples include RNA thermometers changing their conformation in response to temperature and act as translational regulators[16] and the leader RNA of HIV-1 that exists in two alternate conformations.[17] In HIV-1, the global structural transition is exogenously induced by binding to a viral NC protein which acts as an RNA chaperone.[18,19]

Here, we characterize the structural transitions of a 34-mer bistable RNA[20,21] that adopts an equilibrium between two

[†] Johann Wolfgang Goethe University.
[‡] École Polytechnique Fédérale de Lausanne Laboratory of Nucleic Acid Chemistry.

(1) Porschke, D. *Mol. Biol. Biochem. Biophys.* **1977**, *24*, 191−218.
(2) Gilbert, S. D.; Batey, R. T., *Chem. Biol.* **2006**, *13* (8), 805−7.
(3) Crothers, D. M.; Cole, P. E.; Hilbers, C. W.; Shulman, R. G. *J. Mol. Biol.* **1974**, *87* (1), 63−88.
(4) Wu, M.; Tinoco, I., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95* (20), 11555−60.
(5) Nagel, J. H.; Pleij, C. W. *Biochimie* **2002**, *84* (9), 913−23.
(6) Pan, T.; Sosnick, T. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 161−75.
(7) Leulliot, N.; Varani, G. *Biochemistry* **2001**, *40* (27), 7947−56.
(8) Treiber, D. K.; Rook, M. S.; Zarrinkar, P. P.; Williamson, J. R. *Science* **1998**, *279* (5358), 1943−6.
(9) Treiber, D. K.; Williamson, J. R. *Curr. Opin. Struct. Biol.* **1999**, *9* (3), 339−45.
(10) Treiber, D. K.; Williamson, J. R. *Curr. Opin. Struct. Biol.* **2001**, *11* (3), 309−14.
(11) Pan, T.; Sosnick, T. R. *Nat. Struct. Biol.* **1997**, *4* (11), 931−8.
(12) Al-Hashimi, H. M. *ChemBioChem* **2005**, *6* (9), 1506−19.
(13) Tucker, B. J.; Breaker, R. R. *Curr. Opin. Struct. Biol.* **2005**, *15* (3), 342−8.
(14) Schwalbe, H.; Buck, J.; Furtig, B.; Noeske, J.; Wohnert, J. *Angew. Chem., Int. Ed Engl.* **2007**, *46* (8), 1212−9.
(15) Nudler, E.; Mironov, A. S. *Trends Biochem. Sci.* **2004**, *29* (1), 11−7.
(16) Narberhaus, F.; Waldminghaus, T.; Chowdhury, S. *FEMS Microbiol. Rev.* **2006**, *30* (1), 3−16.
(17) Berkhout, B.; van Wamel, J. L. *RNA* **2000**, *6* (2), 282−95.
(18) Huthoff, H.; Berkhout, B. *RNA* **2001**, *7* (1), 143−57.
(19) Ooms, M.; Verhoef, K.; Southern, E.; Huthoff, H.; Berkhout, B. *Nucleic Acids Res.* **2004**, *32* (2), 819−27.

10.1021/ja076739r CCC: $37.00

distinct folds in solution. The sequence is engineered with the goal to mimic natural occurring bistable conformations. By the introduction of a single photoremovable protecting group[22,23] at different positions in the nucleotide sequence, either one of the two bistable conformations can exclusively be stabilized. *In situ* deprotection of the photoprotecting group by laser irradiation within the NMR spectrometer[24,25] allowed for investigation of the kinetics associated with the structural transition. Based on the results of these experiments, the refolding behaviors of secondary structural elements found in different bistable RNA systems were compared leading to more general rules concerning global RNA structural transitions.

Materials and Methods

Synthesis. The photolabile protected nucleotides were prepared as suitable building blocks for solid-phase RNA synthesis as previously described.[22,23] Incorporation of the modified building blocks into RNA occurred efficiently with coupling yields >98% and without side reactions under standard coupling and deprotection conditions. RNAs were assembled using 2′-O-TOM protected building blocks.[26,27]

The synthesis of selectively labeled sequences was conducted as described using canonical 2′-O-TOM protected ribonucleoside phosphoramidites and solid supports containing [$^{13}C_5$]-labeled ribose moieties.[28]

Real-Time NMR Experiments. Kinetic NMR experiments were conducted using a Bruker 800 MHz spectrometer equipped with a cryogenic *z*-gradient HCN probe. For triggering of the reaction, the laser setup was synchronized with the spectrometer via a TTL connection.[25] Kinetic traces were recorded in a pseudo 3D dataset (as previously described[29]). For each kinetic trace, one or two transients were averaged. Processing and analysis of the experimental data were performed using the software FELIX (Accelrys). Nonlinear fitting of the kinetic traces was achieved using the program Sigmaplot (Systat Software Inc.).

NMR Spectroscopy. 1D proton spectra: All spectra were measured at temperatures between 283 and 313 K using a Bruker 800 MHz spectrometer equipped with a cryogenic *z*-gradient HCN probe. Water suppression was achieved by application of either jump-return[30] or watergate[31] pulse trains. Proton pulses were applied on resonance at the water frequency with field strength of 29 kHz. Sufficient *S*/*N* was achieved by recording of a single transient using 0.1 mM RNA samples. For static measurements, 128 transients were averaged, the relaxation delay was set to 1.2 s, and the spectral width was 25 ppm.

Assignment and Secondary Structure Determination. For the assignment of imino-proton resonances, jump-return-NOESY spectra were recorded at 278 K in 9:1 H_2O/D_2O, 25 mM potassium phosphate buffer (pH 6.5) containing 50 mM NaCl. All jump-return-NOESY spectra were recorded using a Bruker 800 MHz spectrometer equipped with a cryogenic *z*-gradient HCN probe. Hard pulses were applied at the midpoint of all RNA proton resonances prior to mixing and at the water frequency with a field strength of 26 kHz. Spectra were recorded with a spectral width of 13 ppm and 25 ppm for the indirect and direct dimensions; 1024 pts and 4096 pts were recorded for t_1 and t_2, respectively. Gradients had a sinusoidal shape and were applied with a strength of 44 G/cm. The relaxation delay was 2 s, and the mixing time ranged between 150 and 250 ms. Other proton resonances were assigned by recording soft-watergate-NOESY spectra with similar parameters. Unambiguous resonance assignment of the selectively ^{13}C-labeled 27-mer RNA was performed using a combination of HCP,[32] HCP-TOCSY,[33] long-range HMBC (optimized here), and forward directed HCC-TOCSY-CCH E.COSY[34,35] experiments. The 3D HCP and HCP-TOCSY experiments were recorded using a Bruker 600 MHz spectrometer equipped with a cryogenic HCP *z*-gradient probe at 298 K. Proton pulses were applied with a field strength of 24 kHz on-resonance at the water frequency, and the carriers for ^{13}C and ^{31}P were set to 77 and 0 ppm, respectively. The spectrum was recorded with 4096, 128, and 64 pts for t_3, t_2, and t_1, respectively. The relaxation delay was set to 1.5 s, and 16 transients were recorded for each increment. In the HCP-TOCSY experiment, the DIPSI-3 sequence[36] was used for the CC-TOCSY with a field strength of 9.3 kHz for 13 ms. The HMBC spectrum was recorded using a Bruker 700 MHz spectrometer at 298K and in a fashion optimized for long-range J(C,H) couplings and with a low-pass J-filter to suppress unwanted one-bond correlations. No decoupling was applied during acquisition. Sinusoidal shaped gradient pulses of 27.5 G/cm and 22.055 G/cm were used for the selection of coherences. Proton pulses and ^{13}C pulses were applied with 23 kHz, 18.5 kHz on resonance at the water frequency and at 88.5 ppm, respectively. The spectrum was recorded with spectral width of 8 ppm and 10 ppm and 2048 pts for t_2 and 64 pts for t_1. Off-resonant carbon Q3 pulses (512 μs) were applied during carbon evolution with an offset of 5000 Hz in order to suppress ^1J(C1′,C2′) coupling. The 3D forward directed HCC-TOCSY-CCH E.COSY experiment was recorded using a DRX600 spectrometer at 298 K. The times for transfer were set to $\tau_L = 14$ ms for the CC-TOCSY and $\tau_L = 8$ ms for the CT delay to optimize the transfer amplitude, as discussed in Glaser et al.[37] The ^{13}C carrier was set to 75 ppm. All ^{1}H pulses were applied at the water frequency. Hard pulses were applied with a field strength of 29 and 19 kHz for ^{1}H and ^{13}C, respectively. GARP decoupling[38] during acquisition was applied with a field strength of 3.85 kHz, and the DIPSI-3 sequence during the CC-TOCSY transfer was applied with a field strength of 9.2 kHz. The spectrum was recorded with 108, 88, and 2048 pts for t_1, t_2, and t_3, respectively, resulting in an experimental duration of 36 h (relaxation delay of 1.4 s). Processing was conducted using XWINNMR 3.5, and the Fourier transformed spectrum was converted to MSI Felix2000 for further analysis. Scalar ^3J(H,H) coupling constants were also extracted from the latter experiment, and the scalar coupling constants were extracted as described.[39,40] The resolution in the direct dimension in which the ^3J(H,H) splitting was measured was 1.75 Hz per complex point and was zero-filled during processing to a resolution of 0.4 Hz per complex point. NOE-derived restraints for modeling the secondary structure were extracted from

(20) Flamm, C.; Hofacker, I. L.; Maurer-Stroh, S.; Stadler, P. F.; Zehl, M. *RNA* **2001**, *7* (2), 254−65.
(21) Hobartner, C.; Micura, R. *J. Mol. Biol.* **2003**, *325* (3), 421−31.
(22) Wenter, P.; Furtig, B.; Hainard, A.; Schwalbe, H.; Pitsch, S. *Angew. Chem., Int. Ed. Engl.* **2005**, *44* (17), 2600−2603.
(23) Wenter, P.; Furtig, B.; Hainard, A.; Schwalbe, H.; Pitsch, S. *ChemBioChem* **2006**, *7* (3), 417−420.
(24) Wirmer, J.; Kuhn, J.; Schwalbe, H. *Angew. Chem., Int. Ed. Engl.* **2001**, *40* (22), 4248−4251.
(25) Kuhn, T.; Schwalbe, H. *J. Am. Chem. Soc.* **2000**, *122* (26), 6169−6174.
(26) Pitsch, S.; Ackermann, D.; Denarie, C.; Meylan, F.; Meyyappan, M.; Muller, E.; Peer, A.; Porcher, S.; Reymond, L.; Stutz, A.; Wenter, P.; Wu, X. L. *Chimia* **2005**, *59* (11), 808−811.
(27) Pitsch, S.; Weiss, P. A.; Jenny, L.; Stutz, A.; Wu, X. L. *Helv. Chim. Acta* **2001**, *84* (12), 3773−3795.
(28) Wenter, P.; Reymond, L.; Auweter, S. D.; Allain, F. H.; Pitsch, S. *Nucleic Acids Res.* **2006**, *34* (11), e79.
(29) Furtig, B.; Buck, J.; Manoharan, V.; Bermel, W.; Jaschke, A.; Wenter, P.; Pitsch, S.; Schwalbe, H. *Biopolymers* **2007**, *86* (5−6), 360−83.
(30) Sklenar, V.; Bax, A. *J. Magn. Reson.* **1987**, *74* (3), 469−479.
(31) Liu, M. L.; Mao, X. A.; Ye, C. H.; Huang, H.; Nicholson, J. K.; Lindon, J. C. *J. Magn. Reson.* **1998**, *132* (1), 125−129.
(32) Marino, J. P.; Schwalbe, H.; Anklin, C.; Bermel, W.; Crothers, D. M.; Griesinger, C. *J. Am. Chem. Soc.* **1994**, *116* (14), 6472−6473.
(33) Marino, J. P.; Schwalbe, H.; Anklin, C.; Bermel, W.; Crothers, D. M.; Griesinger, C. *J. Biomol. NMR.* **1995**, *5* (1), 87−92.
(34) Glaser, S. J.; Schwalbe, H.; Marino, J. P.; Griesinger, C. *J. Magn. Reson. B* **1996**, *112* (2), 160−180.
(35) Schwalbe, H.; Marino, J. P.; Glaser, S. J.; Griesinger, C. *J. Am. Chem. Soc.* **1995**, *117* (27), 7251−7252.
(36) Shaka, A. J.; Lee, C. J.; Pines, A. *J. Magn. Reson.* **1988**, *77* (2), 274−293.
(37) Glaser, S. J.; Schwalbe, H.; Marino, J. P.; Griesinger, C. *J. Magn. Reson. B* **1996**, *112*, (2), 160−80.
(38) Shaka, A. J.; Barker, P. B.; Freeman, R. *J. Magn. Reson.* **1985**, *64* (3), 547−552.
(39) Schwalbe, H.; Samstag, W.; Engels, J. W.; Bermel, W.; Griesinger, C. *J. Biomol. NMR* **1993**, *3* (4), 479−486.
(40) Schwalbe, H.; Marino, J. P.; King, G. C.; Wechselberger, R.; Bermel, W.; Griesinger, C. *J. Biomol. NMR* **1994**, *4* (5), 631−644.

2D double half-filtered NOESY spectra.[41] All spectra were processed using Bruker XWINNMR 3.5, and the Fourier transformed spectra were converted to MSI Felix2000 for further analysis.

Modeling of the structures was achieved by structure calculations using CNX2000 from Accelrys with input of distance restraints from the NOESY experiments, dihedral restraints derived from scalar couplings, and base pair restraints derived from the detection of imino- and amino-resonances.

^{13}C Relaxation Experiments. NMR experiments were carried out using a 600 MHz spectrometer equipped with a 5 mm HCN *z*-gradient probe. ^{13}C-{^{1}H} steady-state NOE data were obtained using ^{13}C modifications of the pulse sequence invnoef3gpsi from the Bruker pulse sequence library. The carrier frequency was set to 89 ppm for C1′, and the spectral width was 6 ppm for the ribose moiety. 128 complex points were acquired in the indirect dimension. Off-resonant carbon Q3 pulses (512 μs) were applied during carbon evolution with an offset of 5000 Hz in order to suppress ^{1}J(C1′,C2′) coupling. Relaxation delays of 5 s were applied. NOE data were obtained with 16 scans for each t_1 increment. The NOE experiment was recorded interleaved, with alternating proton-presaturated and non-presaturated spectra. In the former, proton presaturation was applied for 3 s subsequent to a 2 s relaxation delay in the presaturated spectra, while in the latter a relaxation delay of 5 s was used.

Results

In solution, the investigated 34-mer RNA adopts a dynamic equilibrium between two conformations, referred to as A′ and B″. Conformation A′ consists of two hairpin stem-loop secondary structure elements connected by four single-stranded nucleotides (5′-G15U16U17G18-3′). The hairpin stem-loop structure at the 5′-end of the molecule consists of five base pairs and is closed by a YNMG-tetraloop, while at the 3′-end of the molecule a stem of six base pairs is also closed by a YNMG-tetraloop. In conformation B″, there is a single unpaired nucleotide (A1) and a nine-base-pair canonical stem that is closed by a sequence of seven nucleotides at the 5′-end, while at the 3′-end there is a stem-like structure comprising a single-stranded overhang of eight nucleotides.

The predicted difference in free enthalpy $\Delta G_{\text{theoretical}}$(310K) $= -0.5$ kcal/mol (mfold prediction[42]) results in a ratio of 1:2 of conformation A′ to conformation B″. This prediction is in agreement with the population derived from integration of well-resolved NMR signals stemming from either one of the two conformations and integration at different temperatures yielded values ranging from $\Delta G_{\text{experimental}}$(283K) $= 0.68$ kcal/mol to $\Delta G_{\text{experimental}}$(313K) $= 0.15$ kcal/mol.

NMR Determination of the Topology of Secondary Structure Elements. In order to correlate the kinetics of the refolding of the bistable RNAs with structure, the secondary structure of both conformations have to be known. The three-dimensional structures of all structural elements of conformation A′ (Figure 1) are well-known. In fact, the NMR analysis carried out here including chemical shift analysis fully support the proposition that the structural elements of the 34-mer can be represented by structures reported in the literature.[43,44] Based on these data and all our NMR observations we were able to model the conformation as a long helix capped by the respective loops

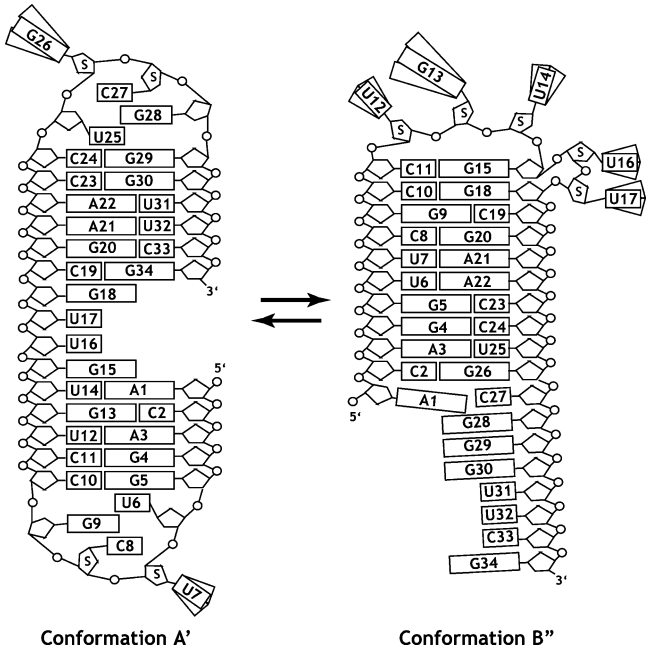


Figure 1. Secondary structure models as derived from NMR data. Left panel depicts the secondary structure model of conformation A′: the tetraloops U6-U7-C8-G9 and U25-G26-C27-G28 adopt the typical YNMG topology, whereas the stems adopt a coaxially stacked conformation. Right panel displays the secondary structure model of conformation B″, where the proposed heptaloop folds onto the helical stem, nucleotides C11 and G15 elongate the helix, and nucleotides U12-G13-U14 and U16-U17 form a triloop and a two-nucleotide bulge, respectively. South-type sugar conformations deviating from canonical A-form RNA are indicated. The nomenclature A′ and B″ was employed due to the fact that NMR resonance annotations can consequently be shortened.

(Figure 1). In contrast to this finding for conformation A, mfold predicted conformation B″ to contain an unknown structural element, namely a heptaloop (5′-C11U12G13U14G15U16U17-3′). In order to determine the structure of the heptaloop, a selectively labeled construct was prepared using chemical synthesis[28,45] that contained ^{13}C-labeled ribose moieties for the loop nucleotides and the two closing nucleotides.

Based on a combination of through-bond correlation and NOESY experiments (see Materials and Methods) (Figure 2), complete assignment of the loop resonances could be obtained, thereby facilitating structural determination of the heptaloop based on NOE and ^{3}J(H,H) coupling constant restraints. Interestingly, the structure adopted by nucleotides 11−17 is not a heptaloop but can best be described as the combination of a two-nucleotide bulge and a three-nucleotide loop (Figure 1). As evident from the NOE connectivities, the nucleobase of nucleotide G15 is stacked on the nucleobase G18. Nucleotides C11 and G15 adopt C3′-endo sugar conformations and extend the stem, whereas the ribose moieties of nucleotides G13, U14, U16, and U17 adopt C2′-endo conformations (indicated by S in Figure 1). Based on our data, nucleotide U12 is conformationally averaged. Although the loop is structurally restrained, {^{1}H}-^{13}C heteronuclear NOE measurements revealed that some nucleotides within this element are highly flexible. The high values for heteronuclear NOE, ranging from 1.4 to 1.6 for nucleotides U12, G13, U14 and U16, U17, indicate that these residues display motion. In contrast, nucleotides C11 and G15

(41) Otting, G.; Wuthrich, K. *Q. Rev. Biophys.* **1990**, *23* (1), 39−96.
(42) Zuker, M. *Nucleic Acids Res.* **2003**, *31* (13), 3406−15.
(43) Allain, F. H.; Varani, G. *J. Mol. Biol.* **1995**, *250* (3), 333−53.
(44) Ennifar, E.; Nikulin, A.; Tishchenko, S.; Serganov, A.; Nevskaya, N.; Garber, M.; Ehresmann, B.; Ehresmann, C.; Nikonov, S.; Dumas, P. *J. Mol. Biol.* **2000**, *304* (1), 35−42.
(45) Quant, S.; Wechselberger, R. W.; Wolter, M. A.; Worner, K. H.; Schell, P.; Engels, J. W.; Griesinger, C.; Schwalbe, H. *Tetrahedron Lett.* **1994**, *35* (36), 6649−6652.

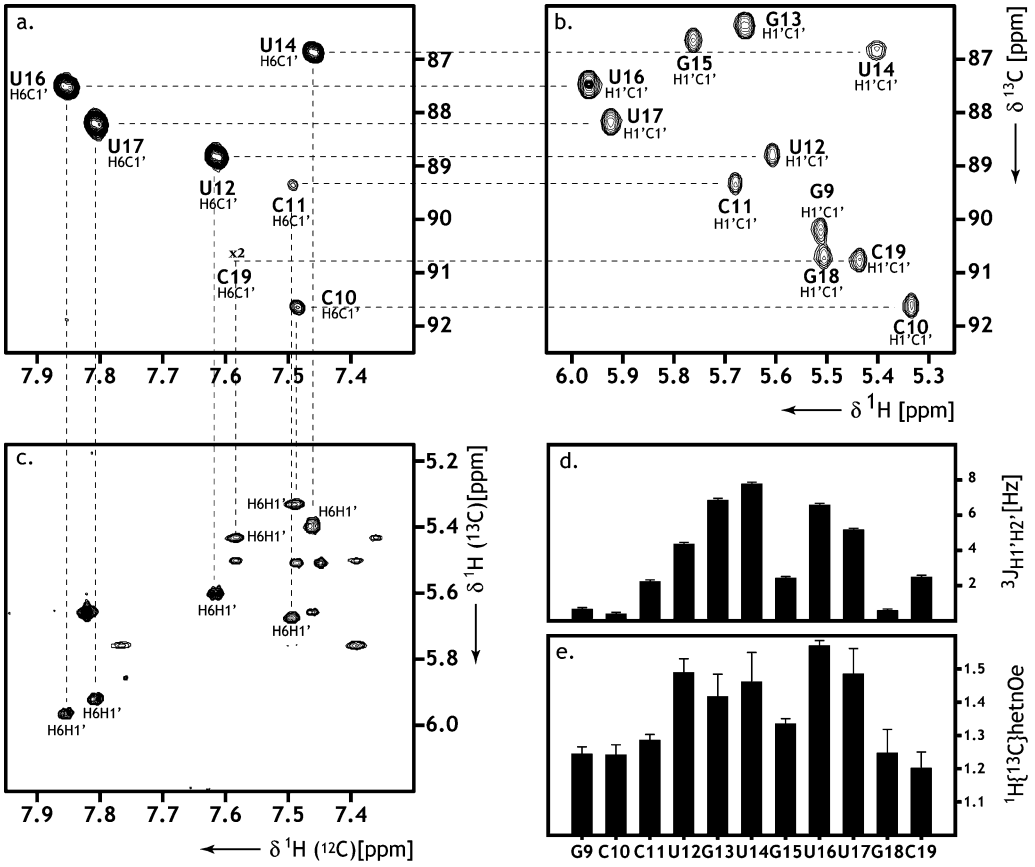


Figure 2. Representation of collected NMR data on the 27-mer reference molecule (secondary structure is depicted in Figure S1d), where the two closing base pairs and nucleotides spanning the heptaloop contain ^{13}C-labeled ribose moieties. Panels a, b, and c depict the H1′/C1′ regions of the long-range ^{1}H,^{13}C-HMBC, the ^{1}H,^{13}C-HSQC, and the ^{1}H(^{13}C),^{1}H-half filtered NOESY spectra, respectively. Resonance assignment is annotated. Panel d displays the measured ^{3}J(H1′,H2′) scalar coupling constants indicative of the conformation of the ribose moieties. In panel e, the ^{1}H{^{13}C}het-NOE data are displayed.

originally assigned as loop nucleotides exhibit significantly lower heteronuclear NOE values of 1.28 and 1.33, which are in the range of values determined for nucleotides G9, C10, and G18, with nucleotide C19 being part of the helical stem and involved in base pairing (as evident from independent NOESY measurements). Additionally, low-temperature NMR measurements revealed that at temperatures at or lower than 273 K, additional exchange-protected imino signals emerge that likely stem from exchange-protected residues in the heptaloop, where one signal likely corresponds to G15 involved in a weak base pairing interaction with C11.

Real-Time NMR Analysis. The investigated bistable RNA displays a dynamic equilibrium. Due to the slow rate of interconversion, the kinetics of this RNA refolding process could not be measured by standard NMR methods such as line shape analysis or exchange spectroscopy at equilibrium. Therefore, we prepared photocaged full-length RNA sequences that selectively destabilize one of the conformations due to modified base pairing properties.[22,23] To stabilize conformation A′, a photoprotecting group was introduced at sequence position U7; this nucleotide was selected since it does not interact with any other nucleotide in conformation A′ but interacts in conformation B″ by forming a Watson−Crick base pair with nucleotide A21. Similarly, conformation B″ was exclusively stabilized by introducing the same photoprotecting group at position G30, which is a "non Watson−Crick interacting" residue in conformation B″ but is involved in canonical G−C base pairs in conformation A′ formed between residues G30 and C23.

The predicted stabilization of the "caged" conformations was monitored by NMR spectroscopy. The imino-region of the 1D spectrum reveals that in either case the resonances indicative of the secondary structural elements are exclusively present (Figure S2). The prepared caged molecules can be "uncaged" using a single laser pulse (1−1.5 s of 4.5 W at 350 nm) which leads to deprotection of the photoprotected nucleobases. After uncaging, the sequences retain their normal base pairing potential and the molecules are able to adopt either one of the two possible conformations according to the thermodynamics of the system. The system therefore relaxes into its equilibrium form with interconversion between the two conformations. This process can be monitored by subsequent recording of NMR spectra (as described in the Materials and Methods section). The buildup and decay curves of signals originating from the different conformations can be integrated over time in order to obtain the folding profile and to extract the kinetic rates by fitting these traces to the kinetic equations of a two-state folding model. These experiments were conducted at different temperatures to extract information about the temperature dependence of refolding. The rates for interconversion of the two conformations are in the range $k_{A'\rightarrow B''} = 0.00003\ s^{-1}$ and $k_{B''\rightarrow A'} = 0.00008\ s^{-1}$ at 288 K and up to $k_{A'\rightarrow B''} = 0.013\ s^{-1}$ and $k_{B''\rightarrow A'} = 0.017\ s^{-1}$ at 313 K (Figure 3a, b). Comparison of the kinetic data with previously reported refolding rates of bistable 20-mer RNA

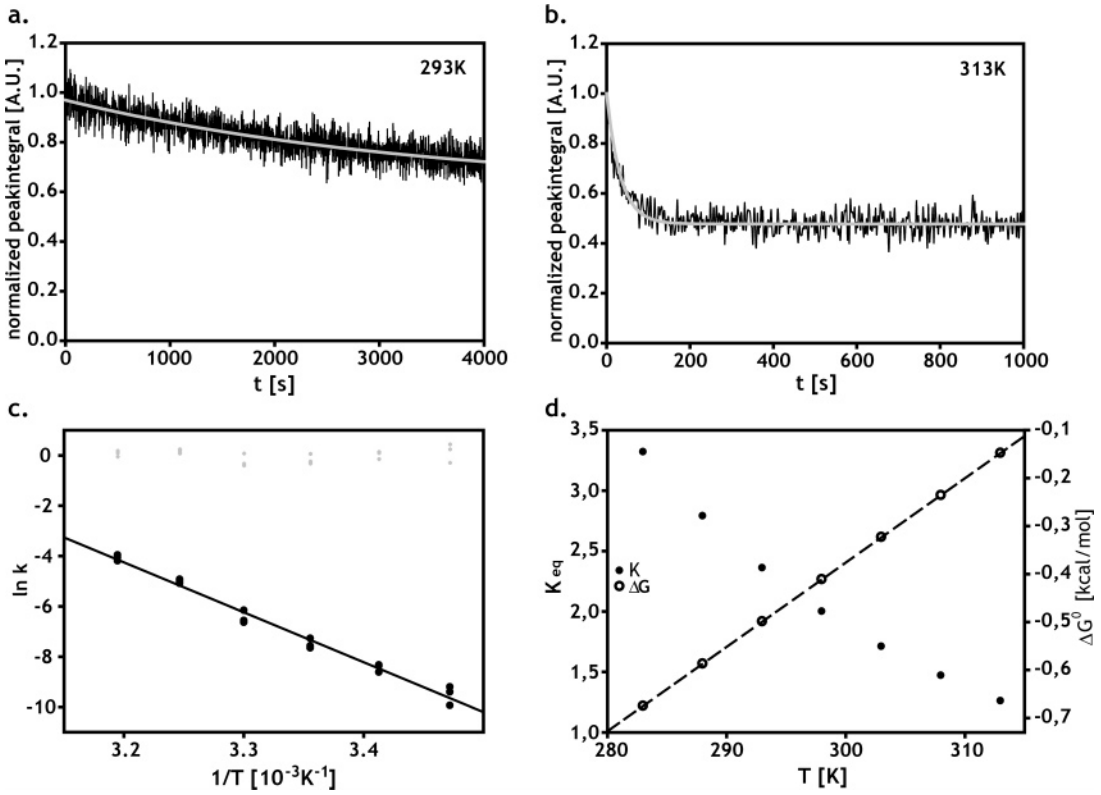


Figure 3. Panels a and b represent kinetic traces measured by real-time NMR of the relaxation of photoprotected conformation B″ (containing $G30^{hv}$) toward equilibrium measured at 293 and 313 K, respectively. Panel c depicts the Arrhenius analysis of the kinetics of this conformational switching process (black points) and the residuals of the linear fit of $\ln k = -(E_a/RT) + \ln A$ (gray points). Panel d shows the extracted equilibrium constants that can be employed to calculate the free enthalpy of the equilibrium ($\Delta G = -RT \ln K_{eq}$). The kinetics were measured at different temperatures in duplicate and triplicate to determine experimental errors.

systems reveals that the rates for the 34-mer bistable RNA are approximately 40 to 60 times smaller. Arrhenius analysis of the refolding rates leads to the determination of activation enthalpies of $\Delta H^{\ddagger}_{(A'\rightarrow B'')} = 44 \pm 2$ kcal/mol and $\Delta H^{\ddagger}_{(B''\rightarrow A')} = 39 \pm 1.2$ kcal/mol and to frequency factors of $A_{(A'\rightarrow B'')} = 1029.6 \pm 2\ s^{-1}$ and $A_{(B''\rightarrow A')} = 1025.7 \pm 1\ s^{-1}$. From the kinetic rates, the activation energies were determined to be ca. 10 kcal/mol higher compared to those of smaller RNA systems (Figure 3c). It should also be mentioned that in the present study we were able for the first time to cage both interconverting conformations which allows for cross-validation of the determined kinetic rates describing the relaxation toward equilibrium following uncaging.

From melting experiments carried out under the same conditions as those for the real-time NMR experiments, base pairing enthalpies were determined. For the specifically synthesized truncated hairpins (Figure S1, Supporting Information), representative secondary structure elements of the full-length conformation, values of $\Delta H_{truncatedA'(1)} = 49$ kcal/mol, $\Delta H_{truncatedA'(2)} = 56$ kcal/mol, [$\Sigma\Delta H_{truncatedA'} = 105$ kcal/mol], and $\Delta H_{truncatedB''} = 78$ kcal/mol were determined. Strikingly, these values are almost equal to twice the activation energies determined by kinetic analysis. This leads us to propose that toward the transition state the RNA has to break half of its stabilizing interactions in order to refold (Figure 4).

Folding of the Structural Elements Occurs Rapidly. The refolding rates of RNA secondary structure is 4 orders of magnitude smaller than the folding rates of RNA from an unstructured RNA chain as reported by several groups.[1] We were able to support these findings for the bistable RNA investigated here by designing and synthesizing a 34-mer sequence carrying a photoprotecting group at nucleotide G9 (Figure 1). As known from extensive structural studies carried out on UUCG tetraloops, the introduction of a photoprotecting group at the fourth nucleotide in the tetraloop leads to disruption of structurally essential base pairing between nucleotides G9

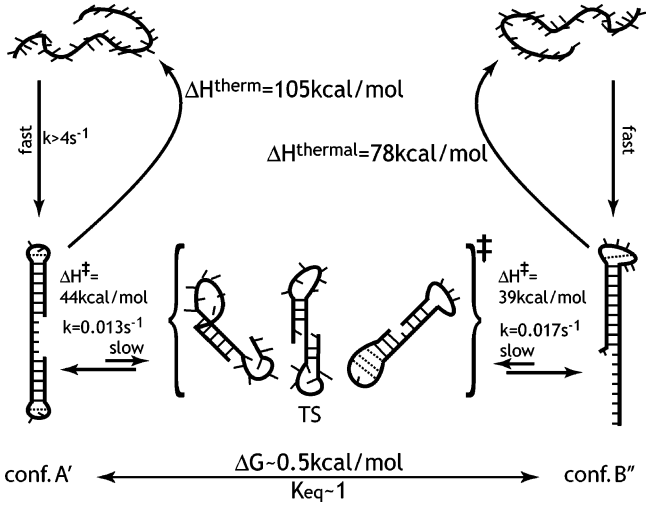


Figure 4. Schematic representation of the refolding transition: In order to conformationally switch between two stable ground states the helical RNA conformation has to pass a transition state (TS) that is represented by an ensemble of conformations characterized by the presence of half of the interactions present in the ground states. Therefore, the rate-limiting step in the RNA refolding reaction represents unfolding toward the transition state, while folding occurs very rapidly.

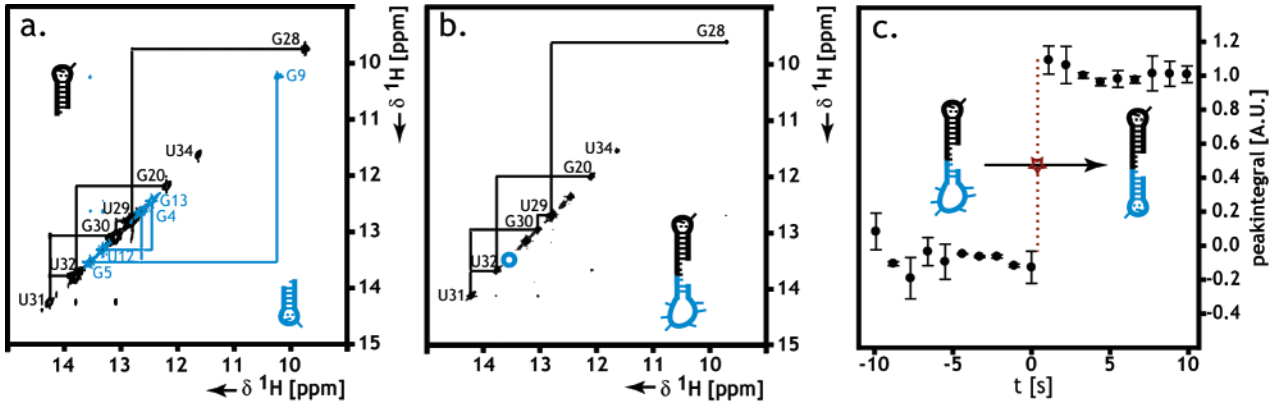


Figure 5. Panel a corresponds to an overlay of the imino region of the $^1H-^1H$ NOESY spectra of the two reference hairpins representing the structural elements present in conformation A′ (5′-ACAGGUUCGCCUGUG-3′ and 5′-GCGAACCUGCGGGUUCG-3′ in blue and black, respectively). Panel b shows the $^1H-^1H$ NOESY spectrum of molecule G9$^{h\nu}$ (5′-ACAGGUUCG$^{h\nu}$CCUGUGUUGCGAACCUGCGGGUUCG-3′) and shows the complete sequential NOESY walk just for one of the two stems (upper stem, black). The second stem (lower stem, blue) is partially unfolded due to incorporation of the photolabile protecting group at a sequence position critical for the formation of the UUCG tetraloop. Panel c shows the kinetic trace that reflects rapid kinetics of the formation of the proper sequence element (loop and stem); the peak of G5 (position indicated by blue circle in b) is present in the first spectrum recorded directly following laser-induced deprotection (indicated by the red star).

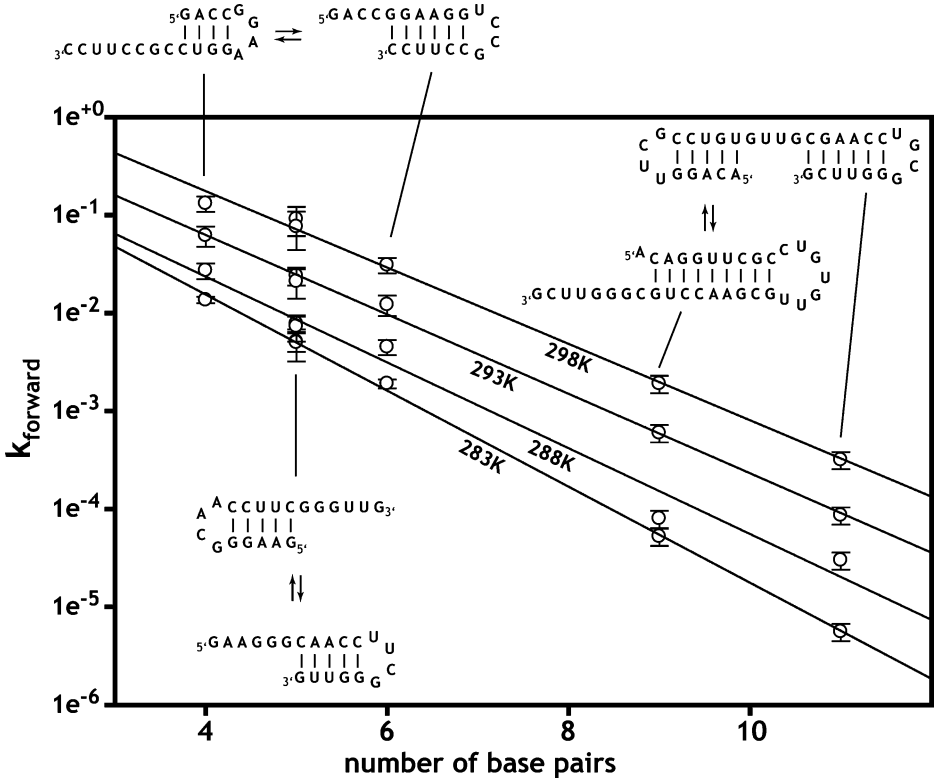


Figure 6. Depiction of measured refolding rates versus the number of base pairs present in the respective stem-loop structures. At all measured temperatures (283, 288, 293, and 298 K), an exponential decrease in the folding rate with increasing helical length was observed.

(fourth loop nucleotide) and U6 (first loop nucleotide), thereby destabilizing the UUCG loop. As a result, although the lower stem-loop structure (nucleotides A1 to U14) in conformation A′ opens, the RNA is not forced into the second bistable conformation B″ since the necessary interaction between G9 and C19 is blocked. The NOESY spectrum of the G9$^{h\nu}$ (Figure 5b) photoprotected molecule supports this proposition. The introduction of a protecting group at G9 leads to at least partial opening of the five base pairs containing stem 5′-A1C2A3G4G5-3′/5′-C10C11U12G13U14-3′ in conformation A′. This becomes obvious from a comparison of the NOESY spectra of reference hairpin 1 (black colored spectrum Figure 5a) and reference hairpin 2 (blue colored spectrum Figure 5a) with the NOESY spectrum of the G9$^{h\nu}$ molecule (Figure 5b). The connectivity of resonances is still present for that part of the molecule that is represented by reference hairpin 2 but is lost for resonances of the first stem-loop represented by reference hairpin 1. Although resonances of the 5′-end of the stem-loop structure are partially present in the 1D-^1H NMR spectra, a NOESY walk fails, indicating that although base pairs distant from the distorted loop may be present their stability is drastically reduced and therefore magnetization of imino resonances is lost due to increased chemical exchange with solvent water. Additionally, resonances of the tetraloop and those resonances indicative of apical base pairs, e.g., G5 (blue circle in Figure 5b), are absent. Apparently, a distorted partially unfolded conformation is

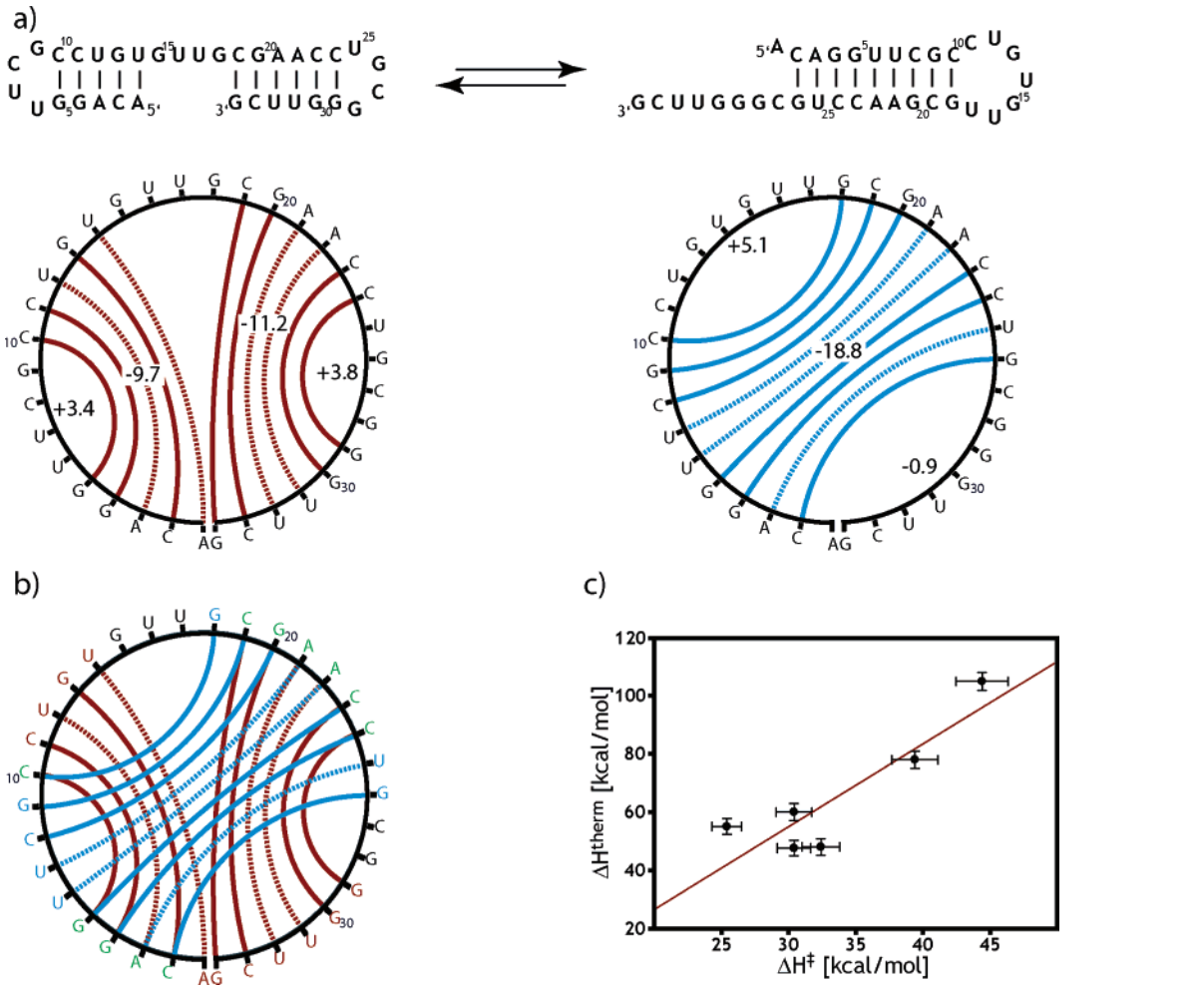


Figure 7. (a) Depiction of the bistable 34-mer used in this study as secondary structure elements (upper panels) and as circular plots (lower panels) as derived with m-fold. Solid lines indicate Watson−Crick G−C base pairs, and dotted lines, Watson−Crick A−U basepairs; annotated numbers represent the energy contribution to the free energy of the structural elements. (b) Overlay of the two circular plots shown in Figure 7a. Color coding indicates the nucleobases that are involved in pairing interactions in one of the conformations (red/blue for Conformation A′ and B″, respectively), in none (black), or in both conformations (green). (c) Comparison of energy parameters for the six different conformations populated by the three bistable RNA sequences. The diagram depicts the linear correlation between $\Delta H^{\ddagger}$ (derived from kinetic experiments) and $\Delta H^{\mathrm{therm}}$ (derived from thermal melting experiments).

present in the G9$^{h\nu}$ state. The kinetics of RNA folding from unfolded or partially unfolded states occurs too rapidly for it to be sampled by time-resolved NMR. As shown in Figure 5c, the peak of G5′ is completely absent before deprotection with a laser pulse and becomes fully present in the first spectrum recorded following laser application. The rates for folding from an unfolded RNA state are higher than those for conformational switching between different folded states. The lower limit for such folding from an unfolded state is estimated to be $k > 4$ s^{-1}, based on the assumption of single exponential behavior.

Folding Rates as Functions of Topologies and Base Pairs Involved in Secondary Structures. The folding rates of three different bistable RNAs were determined (Figure 6). A decrease in folding rate with increasing number of base pairing interactions per A-form helical stem can be observed. This exponential decrease can be related via the Eyring equation (eq 1)

$$k = \kappa \frac{k_B T}{h} \exp\left(-\frac{\Delta G^{\ddagger}}{RT}\right) \qquad (1)$$

to an increase in $\Delta G^{\ddagger}$, where κ, k_b, h, T, and R are known constants.

Comparison of activation energies E_a, derived from Arrhenius analysis of the kinetic data, shows a linear increase with increasing number of base pairs present in the initial structure. Comparison of the kinetic data with the results of prediction programs such as mfold yields a nearly constant dependence of the ratio $\Delta G^{\ddagger}/{}_{\mathrm{initial}}\Delta G$ of ca. 0.5. Taking into account that ${}_{\mathrm{initial}}\Delta G$ is a measure of the interactions of the base pairs involved in the initial conformation, we conclude that during these refolding processes a transition state has to be passed that retains half of the base pairs formed in the initial state. A comparative analysis of the activation enthalpies $\Delta H^{\ddagger}$ with the enthalpies measured in the melting experiments shows a similar behavior (Figure 7c). In both cases, a linear dependence with the number of base pairs is demonstrated. This suggests a refolding process in which an opening of base pairs toward the transition state has to occur as the initial and rate-determining step, followed by propagation in the opening of initial base pairs that is concomitant with the formation of new base pairs present in the final conformational state in the refolding process. The interactions present in the two conformations A′ and B″ and their change during the refolding process can also be visualized

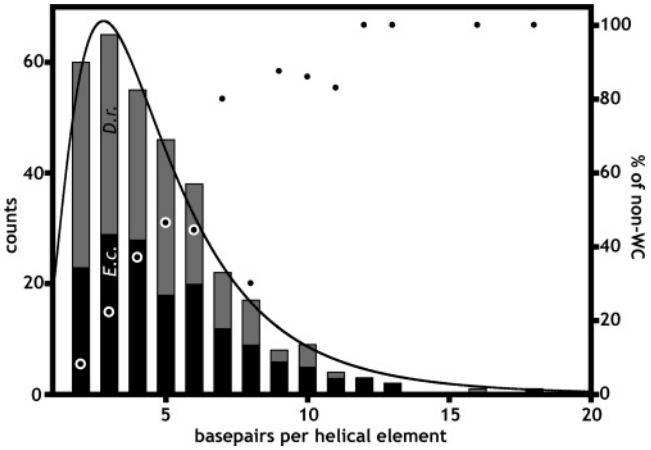


Figure 8. Evaluation of helix length in ribosomal RNAs. A histogram of the length of helical elements in 23S RNA of *E. coli* and *D. radiodurans* as found in secondary structure models J01695 and AE002087. Counted are helical segments that consist solely of Watson−Crick and G−U base pairs. Helices interrupted by noncanonical base pairs are counted as separate helical segments, the percentage of helical elements that contain G−U base pairs (non-WC) is highlighted in the histogram, and the length of the helical elements in the rRNA follows a log-normal distribution with a maximum at three Watson−Crick base pairs.

using circular plots (Figure 7a). In the bistable RNA, 11 and 7 nucleotides are involved in unidirectional interaction only present in either conformation A′ or B″, respectively (Figure 7b green), and 5 nucleotides are unconnected in either of the two conformations (Figure 7b black). In the transition state, the interactions for the bifurcated nucleotides (Figure 7b green) can be paired with their respective partners of either conformation A′ or B″, representing an ensemble with approximately half of the final interactions formed. Extrapolating this consideration to the refolding processes of RNAs in general, the refolding kinetics slows down exponentially with a linear increase in base pairs.

In this context, it is interesting to compare the length of helices under investigation with the length of helices found in ribosomal RNAs. We conducted an analysis on the secondary structure models for the 23S rRNA of *E. coli* (J01695) and *D. radiodurans* (AE002087)[46] and cross-checked these with the three-dimensional model of the large ribosomal subunit of *D. radiodurans*.[47] This analysis reveals that helices with canonical Watson−Crick base pairing patterns in the ribosome seldom exceed a length of 10 base pairs (Figure 8) and that most of the helices comprise 7 or fewer base pairs. Additionally, the number of noncanonical nucleotides increases with increasing helix length in the ribosome, perhaps by destabilization of the potential long helices. We are aware of the argument that states that the length of helices representing the major determinant in the refolding of a structural building block neglects further stabilizing effects such as coaxial stacking or simplifies the role of single nucleotide bulges as breaking elements in helices. We propose, however, that canonical base-paired stems longer than 10 base pairs are avoided in order not to generate kinetic traps that would have to refold to the required structures on a time scale of minutes or even hours. Although the addition of cosolvents such as counterions and denaturants, which are present in the *in vivo* environment of the RNA, may modulate the speed of folding and could contribute significantly to avoiding folding traps such as long "misfolded" helices, the observation that relatively few helices are present in the largest known RNA structure is still striking.

(46) Cannone, J. J.; Subramanian, S.; Schnare, M. N.; Collett, J. R.; D'Souza, L. M.; Du, Y.; Feng, B.; Lin, N.; Madabusi, L. V.; Muller, K. M.; Pande, N.; Shang, Z.; Yu, N.; Gutell, R. R. *BMC Bioinformatics* **2002**, *3*, 2.
(47) Harms, J.; Schluenzen, F.; Zarivach, R.; Bashan, A.; Gat, S.; Agmon, I.; Bartels, H.; Franceschi, F.; Yonath, A. *Cell* **2001**, *107* (5), 679−88.

Conclusions

We could show that folding of secondary structure elements from partially unfolded RNA chains is a rapid process and that, in contrast, the switching of a preexisting secondary structure toward an energetically similar secondary structure is accompanied by significantly slower rates of interconversion. The kinetics of these processes are mainly determined by the energetic contribution of the structural buildings blocks. Comparison of six different conformations of model systems with the secondary structure elements found in ribosomal RNA leads us to propose that natural RNAs avoid canonical helices longer than 10 base pairs per A-form helical element in order to avoid being kinetically trapped in non-native states with very significant barriers for a conformational switch into the most thermodynamically stable conformation.

Acknowledgment. We thank Julia Wirmer, Kai Schlepckow, Janina Buck, and Jens Wöhnert for discussions. Our work was supported by the DFG (SFB 579 "RNA-Ligand-Wechselwirkungen"), the "Studienstiftung des Deutschen Volkes" (B.F.), and the "Fonds der Chemischen Industrie" (H.S.). The Center for Biomolecular Magnetic Resonance (BMRZ) is supported by the state of Hesse.

Supporting Information Available: Figures displaying the NMR spectra of the full-length as well as of the reference RNA molecules and of the photoprotected sequences. This material is available free of charge via the Internet at http://pubs.acs.org.

JA076739R