



Bistable Secondary Structures of Small RNAs and Their Structural Probing by Comparative Imino Proton NMR Spectroscopy

Claudia Höbartner and Ronald Micura*

*Institute of Organic Chemistry
Leopold Franzens University
University of Innsbruck
Innrain 52a, A-6020 Innsbruck
Austria*

We investigate 25–34 nucleotide RNA sequences, that have been rationally designed to adopt two different secondary structures that are in thermodynamic equilibrium. Experimental evidence for the co-existence of the two conformers results from the $\text{NH}\cdots\text{N}$ ^1H NMR spectra. When compared to the $\text{NH}\cdots\text{N}$ ^1H NMR spectra of appropriate reference sequences the equilibrium position is easily quantifiable even without the assignment of the individual NH resonances. The reference sequences represent several Watson–Crick base-paired double helical segments, each encountered in either of the two conformers of the bistable target sequence. In addition, we rationalize the influence of nucleotide mutations on the equilibrium position of one of the bistable RNA sequences. The approach further allows a detailed thermodynamic analysis and the evaluation of secondary structure predictions for multistable RNAs obtained by computational methods.

© 2003 Elsevier Science Ltd. All rights reserved

Keywords: RNA folding; secondary structure; conformational equilibria; multistable RNAs; imino protons

*Corresponding author

Introduction

The biological function of ribonucleic acids is intimately determined by the three-dimensional structure adopted, the RNA fold.¹ Especially for larger RNAs, the pathway to reach the native fold usually proceeds *via* intermediates that represent local minima in the RNA folding free energy landscape.^{2,3} When these intermediates are separated by large energy barriers they constitute folding traps and, therefore, the timescale of the folding process may take up to minutes.^{4,5}

It is only recently that a single RNA sequence of about 150 nucleotides in length has been described to co-exist in two stable folds, each associated with a distinct ribozyme activity.⁶ Also for shorter RNA sequences (20–40 nt), the co-existence of conformers with different defined secondary structures has been proposed and is supported by various computational approaches.^{7,8} However, concerning the experimental verification of such RNA conformational equilibria there is an obvious gap. A reason for this might be that the commonly

used biochemical techniques, such as chemical and enzymatic structural probing or gel shift assays under non-denaturing conditions, which are used to identify alternative folds of large RNAs, are not well suited to investigate secondary structure equilibria of shorter RNAs. The conformers of a short RNA sequence are expected to exist in a dynamic equilibrium with a significantly low energy barrier for their interconversion although this process involves breakage and formation of a great number of different base-pairs. The timescale for the exchange of conformers is not slow enough to allow investigation by gel shift assays or UV-melting experiments; however, the timescale is slow enough to analyze this kind of equilibria by ^1H NMR spectroscopy. We recently demonstrated this for a series of 18–20 nt RNAs that comprise two competing stem–loop motifs.⁹

The NMR resonances for the hydrogen-bonded protons of Watson–Crick base-pairs (“imino protons”) directly reflect the double helical segments within a folded RNA. Therefore, imino proton spectra are in principle sufficient for a direct verification of secondary structure models provided the assignment of the NH resonances is available.¹⁰ The latter usually affords isotope-labeled RNA samples and more advanced NMR techniques.¹¹

Abbreviations used: trFRET, time-resolved fluorescence resonance energy transfer.

E-mail address of the corresponding author: ronald.micura@uibk.ac.at

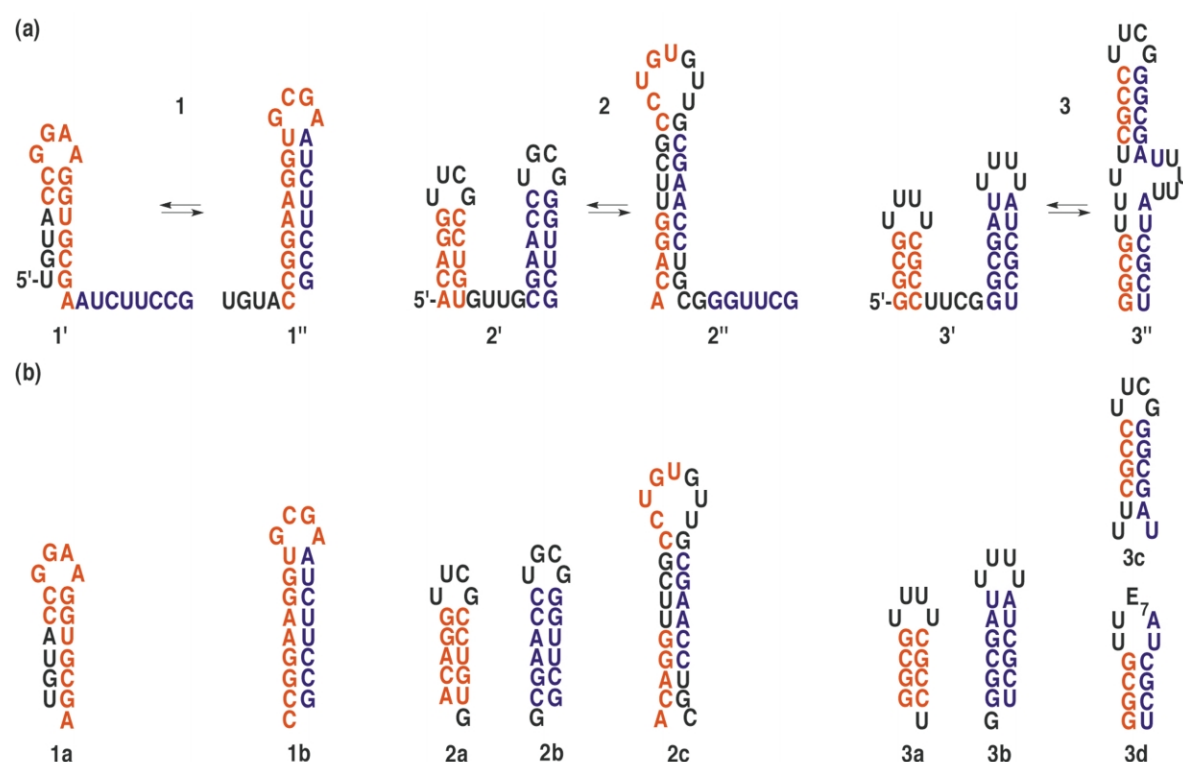


Figure 1. General approach for the investigation of small bistable RNAs. (a) Three examples for rationally designed RNA sequences (1–3) with two competing secondary structures. (b) Truncated sequences (1a/1b, 2a/2b/2c, or 3a/3b/3c/3d) of the corresponding parent RNAs represent references that are only able to adopt a single defined secondary structure. These references allow the assignment of the equilibrium position of 1, 2, or 3 by means of ^1H NMR spectroscopy (compare Figures 2–6) and a detailed thermodynamic analysis (compare Figure 7 and Table 1). E_7 in reference 3d represents a heptaethylene glycol phosphate linker.¹⁶

However, in certain cases the simple comparison of the imino proton spectrum of an RNA with those of truncated but sequence-homologous double helical references may work as a fast and reliable tool for the assessment of a secondary structure model. We first applied such an approach in order to assign the equilibrium of the very short bistable RNAs mentioned above.⁹ This approach is successful because of the possibility to receive a fingerprint set of imino resonances for a single double helical section within the complex pattern of resonances according to the whole target secondary structure. As this kind of structural probing involves NMR spectroscopy and is on the basis of a comparison with the (predicted) individual substructures, we employ the term “structural probing by comparative imino proton NMR spectroscopy”.

Here, we present the extension of our studies on small RNAs that are rationally designed to populate two defined secondary structures simultaneously (Figure 1). The experimental verification of the conformational equilibria is on the basis of comparative imino proton NMR spectroscopy and includes a detailed thermodynamic analysis of the competing secondary structure segments. Moreover, a comparison of the experimental results with the output of algorithms developed for secondary structure prediction of multistable RNAs is discussed.

Results

Bistable RNAs of minimum sequence length

The starting point for our investigations was the challenge to look for a bistable RNA sequence system that is minimal with respect to sequence length and minimal with respect to the number of secondary structure segments. Therefore, our original sequence design followed the objective to combine two common secondary structure motifs in a competing manner within one sequence.

For this purpose, GNRA or UNCG hairpin motifs seemed to be promising.¹² In particular, the GNRA hairpins, 1a and 1b, both comprise an identical sequence section, namely rCCGGAAG GUGCGA, as the characteristic sequence feature, which allows the design of a single sequence construct 1 that is equivalent to the two individual secondary structures 1' and 1'' (Figures 1 and 2). The reference sequences, 1a and 1b, were subject to UV-melting experiments. The melting points were 68.8 °C and 70.8 °C, respectively, in aqueous buffer containing 150 mM NaCl, 10 mM Na_2HPO_4 (pH 7). The thermodynamic analysis of the melting profiles showed that the pairing stability of the two hairpins differed by only 0.5 kcal/mol (1 cal = 4.184 J).^{13,14} This difference suggested a 70:30 equilibrium for 1 in favor of the GNRA-hairpin

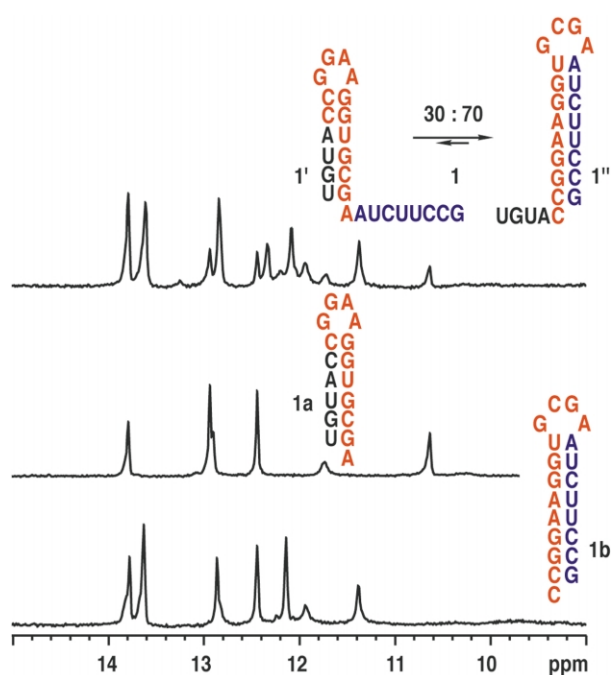


Figure 2. Example of a bistable RNA sequence that is minimal with respect to length and the number of double helical segments. The imino proton NMR spectrum of **1** shows signals according to both references **1a** and **1b**, and ascertains a 30:70 equilibrium of **1'** and **1''** at 298 K, in 25 mM sodium arsenate buffer (pH 7.4), $\text{H}_2\text{O}/^2\text{H}_2\text{O} = 9:1$ (v/v).

1'. Importantly, the melting profile of **1** showed a perfect sigmoid shape and therefore provided no evidence for the possibility of a conformational equilibrium. Also the gel shift assays performed under non-denaturing conditions did not disclose a bistable nature of the secondary structure.

Evidence for sequence **1** to co-exist in two different defined hairpin structures results from ^1H NMR spectroscopy (Figure 2). At 298 K and in 25 mM arsenate buffer (pH 7.4), the NH-spectrum of **1** (0.3 mM) shows the characteristics of slow exchange and nicely reflects both sets of NH resonances as observed for the individual references **1a** and **1b**. The equilibrium ratio for **1** is quantifiable as 70:30 in favor of the GNRA hairpin **1''**.

At this point, we also mention that the acquisition of NH spectra at lower temperature (15 °C and 1 °C) resulted in broader signals for sequence **1** (data not shown). Thereby, a change in the equilibrium position could not be observed or was within the range of accuracy ($\sim \pm 5\%$). The same temperature behavior was observed for one of the bistable sequences described below (**5**), so that we considered 298 K preferable for the major portion of the NMR measurements.

The sequence set of **1**, **1a**, and **1b** formed the basis for a series of even shorter bistable RNAs that we have investigated and documented in the context of the influence of methylation on RNA

secondary structure formation.⁹ In the following, we focus on bistable RNAs comprising more complex secondary structures.

Bistable RNAs: a double *versus* single hairpin motif

Sequence **2** was designed to adopt a double hairpin structure **2'** that competes with a single hairpin consisting of a longer stem and a larger loop **2''** (Figure 3). In comparable manner to the preceding example, truncated homologous sequences represent appropriate references for the two alternative folds of the 34-mer **2**. The 15-mer **2a** and the 17-mer **2b**, each adopts exclusively one defined hairpin. Their conformations correspond to conformation **2'** of parent sequence **2**. In contrast, the 27-mer **2c** represents the reference hairpin corresponding to the alternative conformation **2''** of sequence **2**.

The NH spectra of the references **2a** and **2b** display four and five sharp resonances in good accordance with a five and a six base-pair stem as the terminal base-pairs are usually not detectable at room temperature (298 K) due to fraying of the double helix ends. Moreover, the broadened signals at 10.25 ppm (**2a**) and 9.75 ppm (**2b**) are indicative for the U-G pairing interaction within the UNCG loops.¹⁵ Usually, one additional resonance is indicative for UNCG loops. However, for the buffer conditions used in our study, this resonance was observed only at lower temperatures (for an example see Supplementary Material). The NH spectrum of reference **2c** displays seven sharp signals at 298 K, one of them with double intensity, and is therefore also in accordance with the nine base-pair stem.

To make a direct comparison with the NH spectrum of parent sequence **2** easy, we carried out a numeric addition of the individual reference spectra, each equally weighted. Indeed, the resulting trace of the numeric addition matches almost perfectly the NH spectrum of parent sequence **2** and therefore confirms a 50:50 ratio of the secondary structure equilibrium between **2'** and **2''** (Figure 3). Only a single resonance (marked by an asterisk in Figure 3) is shifted by 0.2 ppm. According to reference **2c** and due to a chemical shift of about 13.5 ppm this resonance is most likely assigned to base-pair A3-U25. This base-pair close to the helix end may be influenced by the different length of the dangling nucleotide overhangs when reference **2c** is compared with the parent sequence in conformation **2''**.

Bistable RNAs: a double hairpin *versus* single hairpin-bulge motif

Sequence **3** was designed to adopt a double hairpin structure **3'**, which competes with a single hairpin comprising a large internal bulge **3''** (Figure 4). Whereas for conformer **3'** the reference sequences are easily defined as **3a** and **3b**, this appears to be

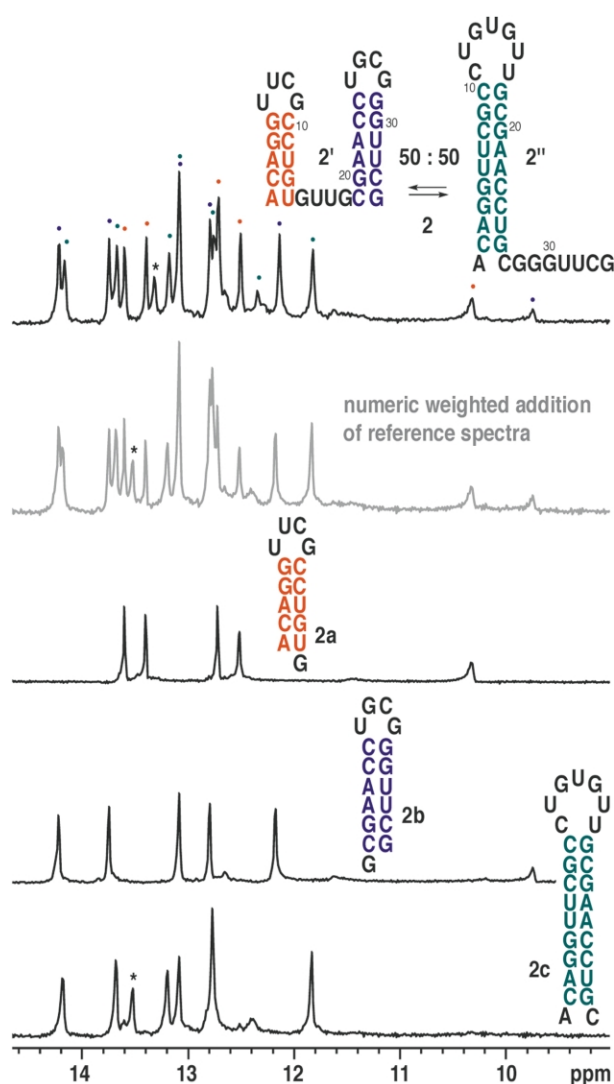


Figure 3. The bistable RNA sequence 2 comprises a double hairpin secondary structure competing with a single hairpin structure. The references allow the stem-wise assignment of the NH resonances (indicated by colored dots). The numeric weighted addition of the individual reference spectra (2a–c) permits the direct quantification of the equilibrium position (trace in gray). A 50:50 equilibrium of 2' and 2'' is ascertained at 298 K, in 25 mM sodium arsenate buffer (pH 7.4), $\text{H}_2\text{O}/^2\text{H}_2\text{O} = 9:1$ (v/v). Only a single resonance is shifted compared to the references (marked with an asterisk).

more difficult for conformer 3''. In this case, only the nucleotide sequence for the reference of the inner stem-loop motif is unambiguously defined as sequence 3c. The other reference that should represent the terminal double helix of 3'' is chosen as tethered oligoribonucleotide 3d with a heptaethylene glycol phosphate linkage.¹⁶ This reference has proven appropriate for a comparative NH NMR analysis. The trace resulting from the numeric addition of the reference spectra closely resembles the NH spectrum of sequence 3 if the reference spectra are equally weighted. Most indicative for the 50:50 population of 3'/3'' are the

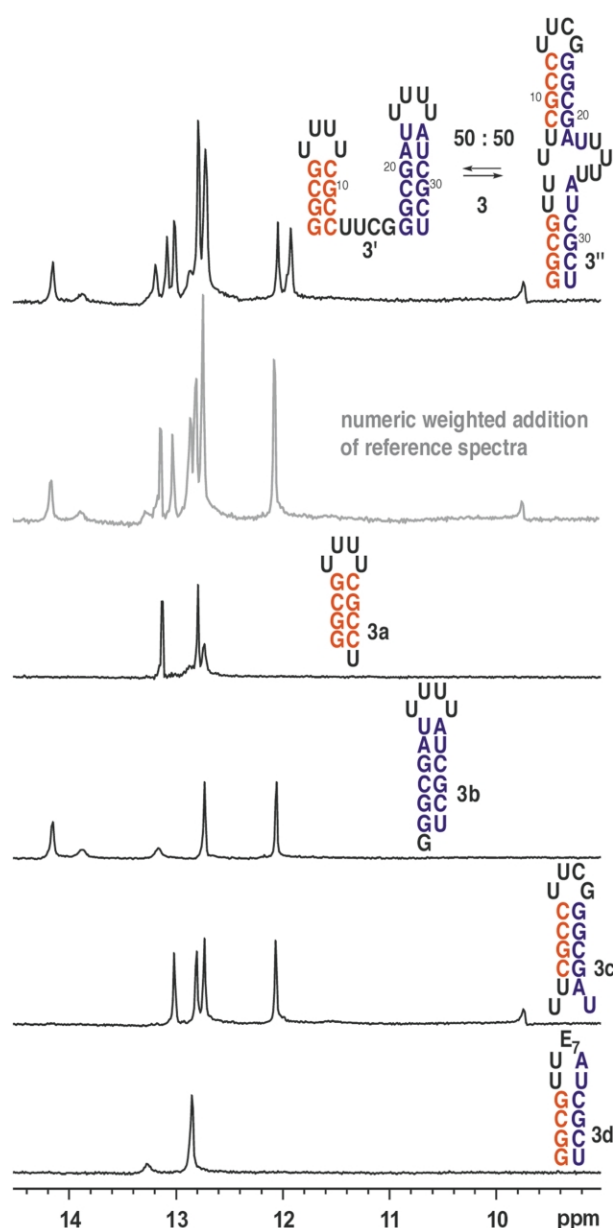


Figure 4. The bistable RNA sequence 3 comprises a double hairpin secondary structure competing with a hairpin-bulge structure. The numeric weighted addition of the individual reference spectra (3a–d) permits the direct quantification of the equilibrium position (trace in gray). Minor deviations of the numeric trace from the spectrum of 3 may result from the impact of the bulge onto the double helical segments, that cannot be mimicked by the references. No shift difference is more than 0.2 ppm. A 50:50 equilibrium of 3' and 3'' is unequivocally ascertained at 298 K, in 25 mM sodium arsenate buffer (pH 7.4), $\text{H}_2\text{O}/^2\text{H}_2\text{O} = 9:1$ (v/v).

resonances for A21–U28 and U22–A27 around 14 ppm (3', compare reference 3b) and the resonance at 9.75 ppm typical for the U–G interaction within a UNCG loop (3'', compare reference 3c). The G–C imino resonances of 3 observed between 12.0 ppm and 13.3 ppm are also in good accordance with the calculated trace; the chemical shift

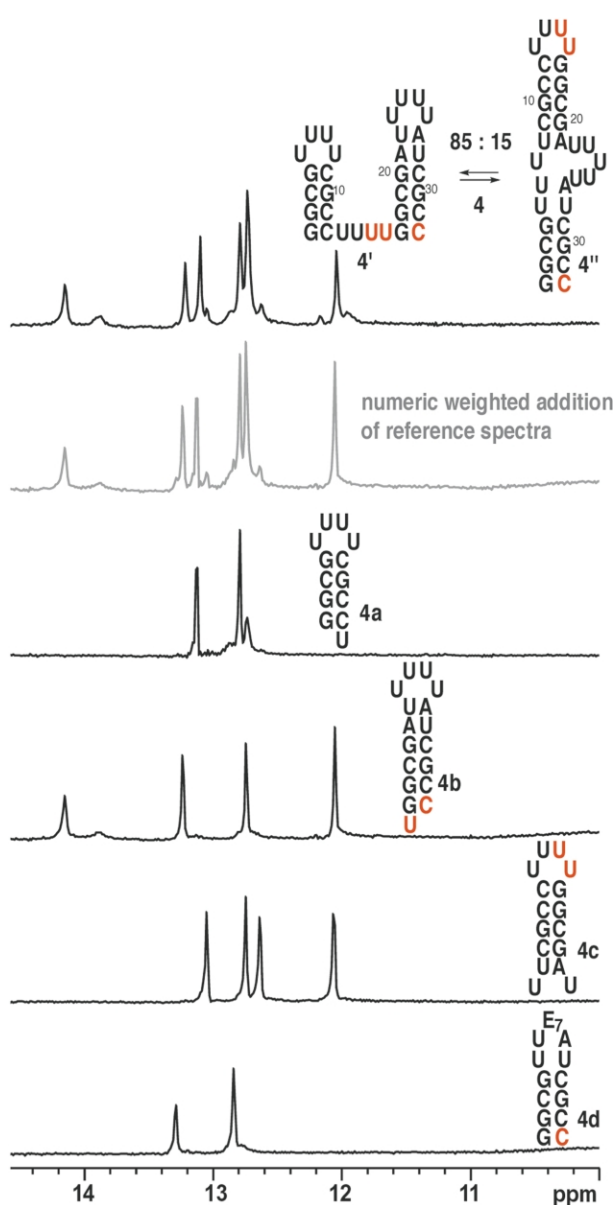


Figure 5. The bistable RNA sequence 4 comprises minor variations in sequence compared with sequence 3 (shown in red). Importantly, the original U_{NC}G loop in 3' is weakened by replacing this section with UUUU in 4'. Comparative imino proton NMR spectroscopy ascertains a 85:15 equilibrium of 4' and 4'' at 298 K, in 25 mM sodium arsenate buffer (pH 7.4), H₂O/²H₂O = 9:1 (v/v).

difference of single signals compared with the corresponding ones of the references does not exceed 0.2 ppm. Such a difference is encountered, e.g. for one of the G·C imino resonances at 12.0 ppm (3b, 3c). Tentatively, the shifted signal can be assigned to the base-pair C9·G20 close to the internal bulge, which may induce structural stress that cannot be mimicked by reference 3c.

Sequence variations and thermodynamics

The nucleotide sequences of 4 and 5 provide minor variations compared to sequence 3 by retain-

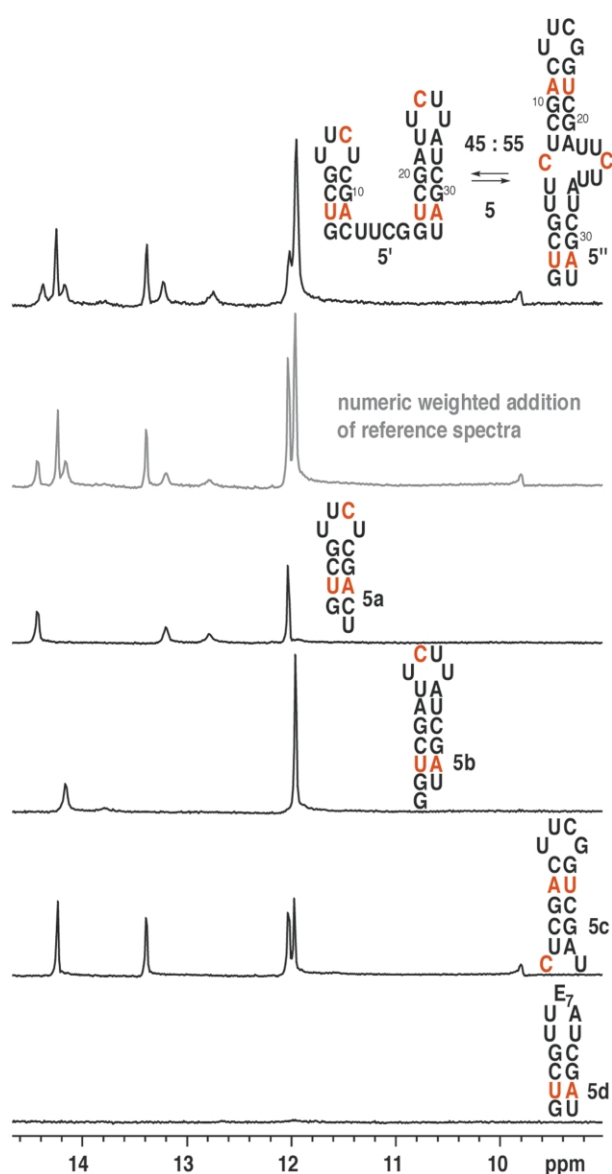


Figure 6. The bistable RNA sequence 5 comprises variations in sequence compared with sequence 3 (shown in red); one G·C base-pair in each stem is replaced by A·U. Comparative imino proton NMR spectroscopy ascertains a 45:55 equilibrium of 5' and 5'' at 298 K, in 25 mM sodium arsenate buffer (pH 7.4), H₂O/²H₂O = 9:1 (v/v). Reference 5d does not show any imino proton resonance. Evidence for a predominant double helical arrangement at 298 K results from UV-melting profiles of 5 ($t_m = 41.9^\circ\text{C}$) and 5d ($t_m = 40.6^\circ\text{C}$; compare Figure 7 and Table 1).

ing the double hairpin *versus* single hairpin-bulge motif of 3. Comparative imino proton NMR analysis as presented above ascertains an 85:15 equilibrium for 4'/4'' (Figure 5) and a 45:55 equilibrium for 5'/5'' at 298 K (Figure 6). The sequence set of 5 was also measured at 10 °C. Broadening of the signals was observed for the bistable sequence but not for the individual reference sequences (Supplementary Material). A change in

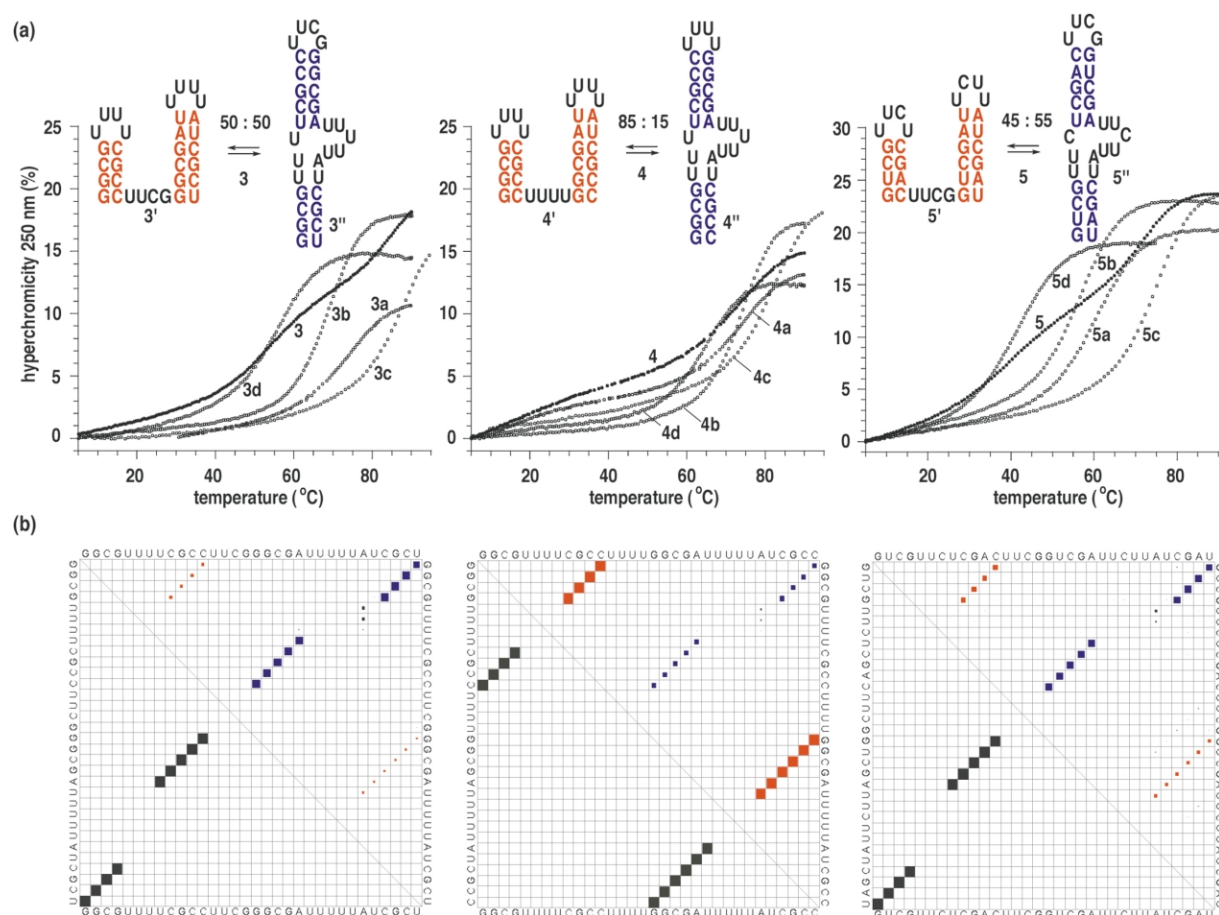


Figure 7. (a) UV-melting profiles of the bistable sequences 3–5, and the corresponding references (2 μ M, 150 mM NaCl, 10 mM Na_2HPO_4 , pH 7.0). (b) Dotplots for the sequences of 3–5, generated by the RNAfold program (Vienna RNA package). The lower left triangle shows the base-pairs of the minimum free energy structure. The upper right triangle represents the frequencies of each base-pair within the ensemble of all possible structures. Thereby, the area of the squares is proportional to pairing probability (see also Table 1).

the equilibrium position could not be observed or was within the range of accuracy ($\sim \pm 5\%$).

With sequence 4 we intended to weaken the extra-stable UUCG hairpin substructure encountered in 3'' by replacing the loop section with UUUU in 4'. The decrease in pairing strength can be estimated on the basis of the UV-melting analysis of the corresponding reference sequences and is reflected in the t_m and ΔG values of 3c ($t_m = 89.2^\circ\text{C}$, $\Delta G^{298\text{ K}} = -9.7\text{ kcal/mol}$) and 4c ($t_m = 82.9^\circ\text{C}$, $\Delta G^{298\text{ K}} = -8.8\text{ kcal/mol}$; Figure 7, Table 1).

In contrast, we increased the pairing strength of the terminal double helix in 4'' compared to 3'. This was achieved by replacement of the very last nucleotide U32 with C32, resulting in formation of a G-C instead of a G-U base-pair. Likewise, the increase in pairing strength can be estimated on the basis of UV-melting studies of the references and is reflected in the t_m and ΔG values of 3d ($t_m = 56.3^\circ\text{C}$, $\Delta G^{298\text{ K}} = -3.7\text{ kcal/mol}$) and 4d ($t_m = 65.8^\circ\text{C}$, $\Delta G^{298\text{ K}} = -5.9\text{ kcal/mol}$).

Importantly, the exchange of U32 to C32 must enhance the free energy of pairing of the competing secondary structure 4' (compare 3b

$t_m = 69.8^\circ\text{C}$, $\Delta G^{298\text{ K}} = -6.9\text{ kcal/mol}$; and 4b $t_m = 75.5^\circ\text{C}$, $\Delta G^{298\text{ K}} = -7.2\text{ kcal/mol}$), whereas the UUCG to UUUU replacement cannot have a primary effect on the free energy of 4'. This tetranucleotide section is unpaired in 4' and represents the linkage between the two individual hairpins.

In comparison with sequence 3, we replaced one of the G-C base-pairs of each stem by an A-U base-pair in sequence 5. This exchange decreases the stability of each double helical segment significantly, which is reflected in lower melting temperatures t_m and lower free energy values $\Delta G^{298\text{ K}}$ of all four reference sequences (Table 1). Concerning the terminal helix of conformer 5'' and the corresponding reference 5d, no imino proton resonance is detected at 298 K. However, a t_m value of 41.9°C and a $\Delta G^{298\text{ K}}$ of -1.8 kcal/mol is observed for reference 5d and provides evidence that the terminal double helix in 5'' is also predominantly formed at room temperature.

The significantly lower stability of the terminal double helical segment in 5'' compared to the other three double helical segments in 5'/5'' is also reflected in the biphasic melting profile of parent sequence 5. The lower transition of 40.6°C nicely

Table 1. Thermodynamic characterization of conformation equilibria of multistable RNA sequences and corresponding reference sequences

Sequence no.	$t_m^{a,f}$ (°C)	Thermodynamic data of double helix formation ^{a,b}			Conformation equilibrium constant ^c		Calculated free energy ^d $\Delta G^{298\text{ K-calc.}}$ (kcal mol ⁻¹)	Predicted conformation equilibrium constant ^e	
		ΔH° (kcal mol ⁻¹)	ΔS° (cal mol ⁻¹ K ⁻¹)	$\Delta G^{298\text{ K}}$ (kcal mol ⁻¹)	$K^{298\text{ K-NMR}}$	Ratio		$K^{298\text{ K-calc.}}$	Ratio
3	85.3, 55.0				1.0	50:50		11.5	8:92
3'							-16.48 ^g		
3''							-17.93		
3a	74.9	-47.5	-137.4	-6.5			-7.96		
3b	69.8	-54.7	-160.3	-6.9			-8.52		
3c	89.2	-58.0	-162.0	-9.7			-13.08		
3d	56.3	-40.8	-124.6	-3.7			-4.81		
4	74.6				0.18	85:15		0.18	85:15
4'							-17.85		
4''							-16.82 ^g		
4a	74.9	-47.5	-137.4	-6.5			-7.96		
4b	75.5	-52.3	-151.5	-7.2			-9.89		
4c	82.9	-56.7	-160.6	-8.8			-10.26		
4d	65.8	-51.6	-153.1	-5.9			-6.52		
5	72.0, 40.6				1.22	45:55		3.80	21:79
5'							-12.42 ^g		
5''							-13.21		
5a	61.0	-42.8	-128.4	-4.5			-5.76		
5b	56.9	-47.0	-142.9	-4.4			-6.66		
5c	76.6	-48.1	-138.8	-6.7			-10.67		
5d	41.9	-36.4	-116.0	-1.8			-2.95		

^a 2 μ M, 150 mM NaCl, 10 mM Na₂HPO₄ (pH 7.0).^b ΔH° and ΔS° were derived from α/T -plots by curve fitting;^{13,14} estimated error in ΔH° , $\Delta S^\circ \leq 10\%$, in $\Delta G^\circ \leq 5\%$.^c 0.3 mM, 25 mM sodium arsenate buffer (pH 7.4), 298 K, H₂O/²H₂O 9:1 (v/v); estimated error in determination of population $\leq 5\%$.^d Minimal free energy (mfe) calculated with RNAfold and RNAsubopt (Vienna RNA Package, version 1.4[†]; ^{37,47}). The temperature parameter was set 298 K ($-T$ 25); the nucleotide “dangling end”-parameter ($-d2$) was applied for calculation of the partition function ($-p$). The software uses the nearest-neighbour parameters derived from measurements in 1 M NaCl as reviewed by Turner.⁴⁸^e The equilibrium constant $K^{298\text{ K-calc.}}$ was obtained from the difference of the calculated free energy (RNAsubopt) of the individual conformers # and #'' ($\Delta\Delta G_{\#,\#''} = -RT \ln K$).^f t_m values were determined as the maximum of differentiated melting curves; estimated error in $t_m \pm 0.4^\circ\text{C}$.^g Free energy of the first suboptimal conformational alternative.

corresponds to the melting temperature of reference **5d**; the sigmoid second transition at 72 °C, however, does not reveal the individual melting transitions of the references **5a** ($t_m = 61.0^\circ\text{C}$), **5b** ($t_m = 56.9^\circ\text{C}$), and **5c** ($t_m = 76.6^\circ\text{C}$). This clearly shows the limitations of melting experiments for the interpretation of bistable RNA sequences. The melting profiles of sequence **3** and the corresponding reference set **3a–d** resemble those discussed for **5**, however, all melting transitions are shifted to somewhat higher temperatures (Figure 7, Table 1).

The melting profile of the bistable sequence **4** shows a single sigmoid transition at 74.6 °C, which corresponds well to the melting of the references **4a** ($t_m = 74.9^\circ\text{C}$) and **4b** ($t_m = 75.5^\circ\text{C}$). These references represent the dominating population **4'** (85%); the existence of the minor population, as proved by comparative imino proton NMR spectroscopy, is not disclosed by the melting profile.

All melting profiles were measured in 150 mM NaCl, 10 mM Na₂HPO₄ (pH 7). Additionally, we tested the condition of 150 mM NaCl, 2 mM MgCl₂, 10 mM Na₂HPO₄ (pH 7), for sequence **4**.

The melting profiles were comparable in shape and provided reversible cooling and heating profiles. The t_m values were higher compared to those recorded under non-magnesium conditions (Supplementary Material).

Discussion

Bistable RNAs and RNA switches

The sequences 1–5 presented in the preceding section demonstrate that small bistable RNAs with different defined secondary structures are easily designable on a rational basis. Bistable RNAs with co-existing conformers in equilibrium are closely related to the concept of RNA switches. An RNA switch mutually adapts one of the two possible conformations depending on the influence of an extrinsic stimulator such as a change in temperature, in pH value, or in Mg²⁺ concentration. Moreover, switching can be induced by binding of a ligand or a protein, or by interaction with another nucleic acid. RNA switches are found to play a

fundamental role in various biological processes.¹⁷ For instance, they are involved in replication of RNA in viroids and viruses,^{18–21} in transcription antitermination of aminoacyl-tRNA synthetase genes,²² in rRNA during decoding,²³ in translational control of phages,^{24,25} and in programmed cell death.^{26,27} Despite the growing number of examples for RNA switches the examples for well investigated bistable RNAs with co-existing conformers are still rare. At this point, we stress the studies by Millar and co-workers on time-resolved fluorescence resonance energy transfer (trFRET) for the analysis of nucleic acids. The fluorescence decay in terms of donor–acceptor distance distributions can resolve different conformers (on the tertiary structure level) providing information on the global structure and flexibility of each species as well as the equilibrium populations. This was demonstrated for the docking equilibrium of the hairpin ribozyme.^{28,29}

Secondary structure predictions for multistable RNAs

Several research groups have focused on computational approaches to explore the conformational energy landscapes and the statistical thermodynamics of RNA secondary structure folding.^{7,8,30–35} We provide a direct comparison of our experimental results with predictions obtained by algorithms (implied in the Vienna RNA package)[†] that cope with the existence of multiple secondary structures. A suboptimal folding algorithm (RNAsubopt) generates a list of all suboptimal secondary structures of a given sequence within a desired energy range above the minimum free energy (ground state).^{36,37} The partition function algorithm of the RNAfold program yields the base-pairing probabilities in a Boltzmann-weighted ensemble of all possible structures. The graphical output for the sequences 3–5 is displayed in Figure 7(b). The dotplots depicted represent the base-pair probability matrix. The lower left triangle shows only base-pairs contained in the minimum free energy structure. The upper right triangle represents the frequencies of each base-pair within thermodynamic equilibrium. Thereby, the area of the squares is proportional to pairing probability. Both algorithms recognize the bistable nature of the RNA sequences 3–5. However, the predicted pairing frequencies of the base-pairs contained in the two dominating conformations may deviate from the experimentally observed populations of the competing secondary structures (compare Table 1).

Kinetics

At the moment we have little knowledge about the kinetics of the bistable sequences presented.

However, the timescale for the interconversion of the conformers may be estimated from an example given by LeCuyer & Crothers. They determined the kinetics of an RNA conformational switch of about 60 nt in length. Thereby, short complementary oligonucleotides drove the two secondary structure alternatives of the spliced leader RNA of *Leptomonas collosoma* from one to the other on a timescale of a fraction of a second.³⁸

Structural probing by comparative imino proton spectroscopy

Determining RNA secondary and tertiary structure is done commonly on the basis of enzymatic and chemical footprinting methods.^{39,40} These methods need only minor sample amounts for analysis and provide a powerful tool for structural probing, especially of large RNAs when NMR and crystallographic analysis become difficult. However, there are also limitations, as the specificities of the enzymes and the reactivities of the probes are not always well understood. The size of an enzyme may limit its access to a susceptible site; the binding of an enzyme can alter the structure; and because a kinetic effect is being measured, a conformation present in small concentration but with high reactivity will be greatly over-represented in the concentrations of products formed. Therefore, interpretation of the data may be very complicated. Interconverting small RNA secondary structures are not accessible by enzymatic and chemical footprinting methods.

However, interconverting small RNA structures exchange slowly enough on the NMR timescale (> milliseconds). The imino proton resonances between 10 ppm and 15 ppm directly reflect the Watson–Crick base-paired double helical regions and are, in principle, sufficient to verify a secondary structure model.^{10,11,41–44} The assignment of the resonances is a prerequisite and usually affords ¹⁵N-labeled RNA samples and advanced NMR methods.^{10,11} We explored a different approach with a pronounced chemical emphasis. The great thermodynamic stability of the secondary structure as compared to tertiary structure of an RNA allows a segmentation into the individual Watson–Crick base-paired double helices: the separately synthesized reference oligonucleotides are easily defined as hairpins for terminal stem–loop segments, but also for internal double helical tracks of a larger RNA, reference oligonucleotides can be defined by the double helical sequence homologues with non-nucleotide linker units such as hexa- or heptaethylene glycol phosphates.¹⁶ The spectra of the individual references provide a characteristic set of imino resonances (fingerprint) and enable a stem-wise assignment of resonances within the complex NH spectrum of the complete RNA. This approach works very reliably for the assignment and quantification of the equilibrium position of bistable RNAs but should be expandable for the verification of secondary structure

[†] <http://www.tbi.univie.ac.at/~ivo/RNA/>

models of larger RNAs as well. For small RNAs, labeling is not needed and one-dimensional NMR spectroscopic methods are sufficient.

In principle, it might be possible to predict chemical shifts of interior base-pairs on the basis of a databank comprising the imino proton spectra of all base-pair combinations within a stem of defined length. However, it should be considered that the short reference sequences can also be commercially ordered so that our fast and direct method is not limited to a synthetic laboratory and represents an alternative to chemical and enzymatic structural probing of small RNAs.

With the present study we aimed at a detailed investigation of bistable RNAs including the rational sequence design, the direct quantification of the equilibrium position and the influence of nucleotide mutations. In ongoing studies we focus on the kinetics of bistable RNAs and on verification of secondary structure predictions by comparative imino NMR spectroscopy for RNAs up to 150-mers.

Materials and Methods

RNA synthesis and purification

5'-O-DMT-2'-O-TOM-protected nucleoside cyanoethylphosphoramidites were obtained from Xeragon, Zurich, and GlenResearch, Sterling, VA.

All oligoribonucleotides were synthesized on CPG supports on a Pharmacia Gene Assembler Plus following slightly modified DNA standard methods. Detritylation was achieved with dichloroacetic acid/1,2-dichloroethane (4:96, v/v) for two minutes. Coupling, 120 μ l of 0.1 M phosphoramidites in acetonitrile was activated by 360 μ l of 0.35 M benzyl thiotetrazole in acetonitrile. Capping was achieved with 1:1 (v/v) A/B for one minute (A, $\text{Ac}_2\text{O}/\text{sym-collidine}/\text{acetonitrile}$), 2:3:5, by vol.; B, 0.5 M 4-(dimethylamino)pyrimidine in acetonitrile). Oxidation was done with 10 mM I_2 in acetonitrile/*sym-collidine*/water (10:1:5, by vol.) for one minute. Amidite solutions, tetrazole solutions, and acetonitrile were dried over activated molecular sieves overnight. All sequences were synthesized trityl-off.

Deprotection and cleavage of the oligonucleotides from solid support were achieved with MeNH_2 in EtOH (8 M, 0.75 ml) and MeNH_2 in water (40%, 0.75 ml) for five to nine hours at room temperature; the solution was then evaporated to dryness. Removal of the 2'-O-silyl ethers was afforded by treatment with tetrabutylammonium fluoride trihydrate ($\text{TBAF}\cdot 3\text{H}_2\text{O}$) in THF (1 M, 0.90 ml) for at least 12 hours at room temperature. The reaction was quenched by the addition of Tris-HCl (1 M, pH 7.4, 0.95 ml). The volume of the solution was reduced to 1 ml and directly applied on a Sephadex G 10 column (30 cm \times 1.5 cm) monitored by UV-detection at 270 nm. The product was eluted with water and evaporated to dryness.

All oligoribonucleotides were purified by ion-exchange chromatography on a semipreparative Dionex DNAPac column at 80 °C. Flow rate: 2 ml/minute: eluant A, 25 mM Tris-HCl, 6 M urea, in water (pH 8.0); eluant B, 25 mM Tris-HCl, 0.5 M NaOCl_4 , 6 M urea, in water (pH 8.0); detection at 265 nm; gradient I (for checking

the purity of crude products after deprotection): 45 minutes 0–60% B in A; gradient II (for purification): 10–15% B in A within 30 minutes. Fractions containing the purified oligonucleotide were desalted by loading onto a C18 SepPak cartridge (Waters/Millipore), followed by elution with 0.1–0.2 M $(\text{Et}_3\text{NH})\text{HCO}_3$, water, and then water/ CH_3CN (6:4, v/v). Combined fractions containing the oligonucleotide were lyophilized to dryness.

The masses expected have been confirmed for the oligoribonucleotides by MALDI-TOF or ESI mass spectrometry.

Thermal denaturation studies

Absorbance *versus* temperature profiles were recorded at 250 nm, 260 nm, 265 nm, and 270 nm on a Cary-1 spectrophotometer equipped with a multiple cell holder and a Peltier temperature-control device. Each sequence was measured at least at two different concentrations (2 μ M and 40 μ M). All sequences were measured in buffer solutions of 10 mM Na_2HPO_4 at pH 7.0 containing 150 mM NaCl. Sequence 4 was additionally measured at a salt concentration of 150 mM NaCl and 2 mM MgCl_2 . Data were collected after a complete cooling and heating cycle at a rate of 0.7 deg. C/minute. Melting transitions were reversible for all sequences and essentially the same with respect to the four different wavelengths.

Oligonucleotides were lyophilized to dryness, dissolved in the corresponding buffer from stock solutions and subsequently degassed. A layer of silicon oil was placed on the surface of the solution.

Values of ΔH° and of ΔS° for monomolecular melting transitions were derived from a two-state van't Hoff analysis by fitting the shape of the individual α *versus* temperature curve as described.^{13,14} Errors for ΔH and ΔS arising from non-infinite cooperativity of two-state transitions and from the assumption of a temperature-independent enthalpy are typically 10–15%. Additional error is introduced when free energies are extrapolated far away from the melting transitions; errors for ΔG° are typically 3–5%.

NMR spectroscopy

^1H NMR NH-spectra were acquired on a Varian Unity 500, applying a selective excitation refocusing sequence employing selective pulses shaped according to the G4 (excitation;⁴⁵ 2.62 ms, rf amplitude 1.74 kHz) or REBURP (refocusing;⁴⁶ 1.4 ms, rf amplitude 4.47 kHz) profile, respectively. Both shaped pulses were centered at 13 ppm.

NMR sample preparation

$c = 0.2\text{--}0.3$ mM oligonucleotide; 25 mM sodium arsenate buffer (pH 7.4), 298 K, $\text{H}_2\text{O}/^2\text{H}_2\text{O} = 9:1$ (v/v). The oligoribonucleotides (triethyl ammonium salts) were lyophilized together with the corresponding amount of arsenate buffer (Sigma), redissolved in water, lyophilized again and finally dissolved in $\text{H}_2\text{O}/^2\text{H}_2\text{O}$. All samples were heated to 90 °C for one minute, then rapidly cooled to room temperature and equilibrated for one hour before measurements. The samples were stored at 4 °C and remeasured after a week. Thereby, some of the reference sequences revealed duplex formation to a minor extent.

Comment on the different salt conditions used in the secondary structure calculations, the NMR and the melting experiments

Most of the algorithms for RNA secondary structure free energy predictions implement the “nearest-neighbor” parameters which have been carefully measured and continuously refined by Turner and co-workers over the last decade.⁴⁸ The parameter set is predominantly on the basis of buffer solutions containing 1 M NaCl. For NMR measurements, such high salt concentrations are not applicable because of the high viscosity of the sample solution. For UV-melting experiments we chose 150 mM NaCl to be close to physiological salt conditions, and moreover, to obtain sigmoid melting profiles in a temperature range that allows a reliable thermodynamic analysis; in 1 M NaCl solutions the hairpin references 3c and 4c are expected to melt at temperatures higher than 90 °C.

Acknowledgements

Financial support from the Austrian Science Fund (P15042) is gratefully acknowledged. We thank Mag. G. Hölzl (Innsbruck), Dr K. Breuker (Ithaca, N.Y., Innsbruck) for mass analyses, Dr C. Eichmüller (Innsbruck), Professor R. Konrat (Vienna) for sharing their NMR expertise, and Professor B. Kräutler (Innsbruck) for continuous support. Special thanks to C. Flamm and I. Hofacker (Vienna) for a stimulating discussion on RNA secondary structure prediction.

References

1. Batey, R. T., Rambo, R. P. & Doudna, J. A. (1999). Tertiary motifs in RNA structure and folding. *Angew. Chem., Int. Ed. Engl.* **38**, 2326–2343.
2. Pan, J., Thirumalai, D. & Woodson, S. A. (1997). Folding of RNA involves parallel pathways. *J. Mol. Biol.* **273**, 7–13.
3. Chadalavada, D. M., Senchak, S. E. & Bevilacqua, P. C. (2002). The folding pathway of genomic hepatitis delta virus ribozyme is dominated by slow folding of the pseudoknots. *J. Mol. Biol.* **317**, 559–575.
4. Treiber, D. K., Rook, M. S., Zarrinkar, P. P. & Williamson, J. R. (1998). Kinetic intermediates trapped by native interactions in RNA folding. *Science*, **279**, 1943–1946.
5. Nagel, J. H. A., Gulyaev, A. P., Gerdes, K. & Pleij, C. W. A. (1999). Metastable structures and refolding kinetics in *hok* mRNA of plasmid R1. *RNA*, **5**, 1408–1419.
6. Schultes, E. A. & Bartel, D. P. (2000). One sequence, two ribozymes: implications for the emergence of new ribozyme folds. *Science*, **289**, 448–452.
7. Flamm, C., Hofacker, I. L., Maurer-Stroh, S., Stadler, P. F. & Zehl, M. (2001). Design of multistable RNA molecules. *RNA*, **7**, 254–265.
8. Tostesen, E., Chen, S.-J. & Dill, K. A. (2001). RNA folding transitions and cooperativity. *J. Phys. Chem. B*, **105**, 1618–1630.
9. Höbartner, C., Ebert, M.-O., Jaun, B. & Micura, R. (2002). On RNA two-state conformation equilibria and the impact of nucleobase methylation. *Angew. Chem., Int. Ed. Engl.* **41**, 605–609.
10. Wu, M. & Tinoco, I., Jr (1998). RNA folding causes secondary structure rearrangement. *Proc. Natl Acad. Sci. USA*, **95**, 11555–11560.
11. Tanaka, Y., Hori, T., Tagaya, M., Sakamoto, T., Kurihara, Y., Katahira, M. & Uesugi, S. (2002). Imino proton NMR analysis of HDV ribozymes: nested double pseudoknot structure and Mg²⁺ ion-binding site close to the catalytic core in solution. *Nucl. Acids Res.* **30**, 766–774.
12. Varani, G. (1995). Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 397–404.
13. Marky, L. & Breslauer, K. (1987). Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers*, **26**, 1601–1620.
14. Xia, T., Mathews, D. H. & Turner, D. H. (1999). Thermodynamics of RNA secondary structure formation. In *Comprehensive Natural Product Chemistry* (Söll, D., Nishimura, S. & Moore, P., eds), vol. 8, pp. 21–47, Elsevier, Oxford.
15. Williams, D. J., Boots, J. L. & Hall, K. B. (2001). Thermodynamics of 2'-ribose substitutions in UNCG tetraloops. *RNA*, **7**, 44–53.
16. Pils, W. & Micura, R. (2000). Flexible non-nucleotide linkers as loop replacements in short double helical RNAs. *Nucl. Acids Res.* **28**, 1859–1863.
17. de Smit, M. H. (1998). Translational control by mRNA structure in eubacteria: molecular biology and physical chemistry. In *RNA Structure and Function* (Simons, R. W. & Grunberg-Manago, M., eds), pp. 495–540, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
18. Repsilber, D., Wiese, S., Rachen, M., Schroder, A. W., Riesner, D. & Steger, G. (1999). Formation of metastable RNA structures by sequential folding during transcription: time-resolved structural analysis of potato spindle tuber viroid (–)-stranded RNA by temperature-gradient gel electrophoresis. *RNA*, **5**, 574–584.
19. Olsthoorn, R. C., Mertens, S., Brederode, F. T. & Bol, J. F. (1999). A conformational switch at the 3'-end of a plant virus regulates viral replication. *EMBO J.* **18**, 4856–4864.
20. Huthoff, H. & Berkhout, B. (2001). Mutations in the TAR hairpin affect the equilibrium between alternative conformations of the HIV-1 leader RNA. *Nucl. Acids Res.* **29**, 2594–2600.
21. Dirac, A. M. G., Huthoff, H., Kijms, J. & Berkhout, B. (2002). Regulated HIV-2 RNA dimerization by means of alternative RNA conformations. *Nucl. Acids Res.* **30**, 2647–2655.
22. Grundy, F. J., Rollins, S. M. & Henkin, T. M. (1994). Interaction between acceptor end of tRNA and the T Box stimulates antitermination in the *Bacillus subtilis* *tyrS* gene: a new role for the discriminator base. *J. Bacteriol.* **176**, 4518–4526.
23. Lodmell, J. S. & Dahlberg, A. E. (1997). A conformational switch in *Escherichia coli* 16S ribosomal RNA during decoding of messenger RNA. *Science*, **277**, 1262–1267.
24. Poot, R. A., Tsareva, N. V., Boni, I. V. & van Duin, J. (1997). RNA folding kinetics regulates translation of phage MS2 maturation gene. *Proc. Natl Acad. Sci. USA*, **94**, 10110–10115.

25. van Meerten, D., Girard, G. & van Duin, J. (2001). Translational control by delayed RNA folding: identification of the kinetic trap. *RNA*, **7**, 483–494.
26. Gulyaev, A. P., Franch, T. & Gerdes, K. (1997). Programmed cell death by *hok/sok* of plasmid R1: coupled nucleotide covariations reveal a phylogenetically conserved folding pathway in the *hok* family of mRNAs. *J. Mol. Biol.* **273**, 26–37.
27. Gerdes, K., Gulyaev, A. P., Franch, T., Pedersen, K. & Mikkelsen, N. D. (1997). Antisense RNA-regulated programmed cell death. *Annu. Rev. Genet.* **31**, 1–31.
28. Klostermeier, D. & Millar, D. P. (2002). Time-resolved fluorescence energy transfer: a versatile tool for the analysis of nucleic acids. *Biopolymers*, **61**, 159–179.
29. Klostermeier, D. & Millar, D. P. (2001). RNA conformation and folding studied with fluorescence energy transfer. *Methods*, **23**, 240–254.
30. Zuker, M. (2000). Calculating nucleic acid secondary structure. *Curr. Opin. Struct. Biol.* **10**, 303–310.
31. Morgan, S. R. & Higgs, P. G. (1996). Evidence for kinetic effects in folding of large RNA molecules. *J. Chem. Phys.* **105**, 7152–7157.
32. Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940.
33. Chen, S.-J. & Dill, K. A. (2000). RNA folding energy landscapes. *Proc. Natl Acad. Sci. USA*, **97**, 646–651.
34. Breton, N., Jacob, C. & Daegelen, P. (1997). Prediction of sequentially optimal RNA secondary structures. *J. Biomol. Struct. Dynam.* **14**, 727–740.
35. Gaspin, C. & Westhof, E. (1995). An interactive framework for RNA secondary structure prediction with a dynamic treatment of constraints. *J. Mol. Biol.* **254**, 163–174.
36. Wuchty, S., Fontana, W., Hofacker, I. L. & Schuster, P. (1999). Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers*, **49**, 145–165.
37. Flamm, C., Fontana, W., Hofacker, I. L. & Schuster, P. (2000). RNA folding at elementary step resolution. *RNA*, **6**, 325–338.
38. LeCuyer, K. A. & Crothers, D. M. (1994). Kinetics of an RNA conformational switch. *Proc. Natl Acad. Sci. USA*, **91**, 3373–3377.
39. Moine, H., Ehresmann, B., Ehresmann, C. & Romby, P. (1998). Probing RNA structure and function. In *RNA Structure and Function* (Simons, R. W. & Grunberg-Manago, M., eds), pp. 77–115, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
40. Hearst, J. (2000). Chemical and enzymatic methods. In *Nucleic Acids Structures, Properties, and Function* (Bloomfield, V. A., Crothers, D. A. & Tinoco, I. Jr, eds), pp. 45–78, University Science Books, Sausalito, CA.
41. Xu, J., Lapham, J. & Crothers, D. M. (1996). Determining RNA solution structure by segmental isotopic labeling and NMR: application to *Caenorhabditis elegans* spliced leader RNA 1. *Proc. Natl Acad. Sci. USA*, **93**, 44–48.
42. Leontis, N. B. & Moore, P. B. (1986). NMR evidence for dynamic secondary structure in helices II and III of the 5 S RNA of *Escherichia coli*. *Biochemistry*, **25**, 3916–3925.
43. Heus, H. A., Uhlenbeck, O. C. & Pardi, A. (1990). Sequence-dependent structural variations of hammerhead RNA enzymes. *Nucl. Acids Res.* **18**, 1103–1108.
44. Odai, O., Kodama, H., Hiroaki, H., Sakata, T., Tanaka, T. & Uesugi, S. (1990). Synthesis and NMR study of ribooligonucleotides forming a hammerhead-type RNA enzyme system. *Nucl. Acids Res.* **18**, 5955–5960.
45. Emsley, L. & Bodenhausen, G. (1990). Gaussian pulse cascades: new analytical functions for rectangular selective inversion and in-phase excitation in NMR. *Chem. Phys. Letters*, **165**, 469–476.
46. Geen, H. & Freeman, R. (1991). Band-selective radio-frequency pulses. *J. Magn. Reson.* **93**, 93–141.
47. Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, L. S., Tacker, M. & Schuster, P. (1994). Fast folding and comparison of RNA secondary structures. *Monatsh. Chem.* **125**, 167–188.
48. Turner, D. H. (2000). Conformational changes. In *Nucleic Acids* (Bloomfield, V. A., Crothers, D. M. & Tinoco, I. Jr, eds), pp. 259–334, University Science Books, Sausalito, CA.

Edited by J. Doudna

(Received 14 August 2002; received in revised form 25 October 2002; accepted 29 October 2002)

SCIENCE  DIRECT®
www.sciencedirect.com

Supplementary Material comprising seven Figures is available