334–342 *Nucleic Acids Research, 2006, Vol. 34, No. 1*
doi:10.1093/nar/gkj435

Stabilities of HIV-1 DIS type RNA loop–loop interactions *in vitro* and *in vivo*

Christina Lorenz, Nicolas Piganeau and Renée Schroeder*

Max F. Perutz Laboratories, Department of Biochemistry, University of Vienna, Dr Bohrgasse 9/5, A-1030 Vienna, Austria

Received September 7, 2005; Revised and Accepted December 21, 2005

ABSTRACT

RNA loop–loop interactions are a prevalent motif in the formation of tertiary structure and are well suited to trigger molecular recognition between RNA molecules. We determined the stabilities of several loop–loop interactions with a constant 6 bp core sequence and varying unpaired flanking nucleotides and found that the flanking bases have a strong influence on the stability and ion dependence of the kissing complex. In general, the stabilities determined in 1 M Na^+ are equivalent to those in the presence of near physiological Mg^{2+} concentrations. Therefore we further tested whether the stabilities determined *in vitro* and within yeast cells correlate, using a recently developed yeast RNA-hybrid system. For the majority of the loop types analyzed here, the melting temperatures determined *in vitro* are in good agreement with the relative β-galactosidase activity in yeast cells, showing that data derived from *in vitro* measurements reflect *in vivo* properties. The most stable interactions are the naturally occurring HIV-1 DIS MAL and LAI derived loops with the motif (5′ $A^A/_GN_6A$ 3′), emphasizing the crucial role of stable kissing complexes in HIV genome dimerization.

INTRODUCTION

RNA molecules are transcribed as linear polycationic chains and readily form secondary structures by Watson–Crick (WC) base pairing. This leads to the formation of A-form helices, which orient themselves in space to build the tertiary structure. RNA secondary structures can be very stable and the energies involved in their formation are most often significantly larger than those involved in tertiary interactions. RNA folding is clearly hierarchical and sequential, with the secondary structure forming fast and without the requirement of Mg^{2+} ions, while formation of the tertiary structure is complex and requires stabilization by divalent ions (1). Formation of the tertiary structure only minimally alters the secondary structure (2).

The rules governing secondary structure formation are now well defined through enormous experimental efforts undertaken to determine the energies that stabilize helices. As a consequence, prediction of RNA secondary structure has become very accurate (3–6). In contrast, tertiary structures are not yet possible to calculate or to predict due to lack of knowledge about the contribution of tertiary interactions to the overall stability of an RNA molecule and to the lack of sufficiently well-tested algorithms (7).

Among the interactions important for the formation of tertiary structures are WC base pairs, which can occur between two single-stranded regions, e.g. two hairpin loops forming so-called loop–loop interactions or kissing complexes. These complexes can form intermolecularly or intramolecularly and are in the latter case called pseudoknots. Intermolecular loop–loop interactions are a very prevalent motif and are found to participate in a wide range of biological functions (8). Because of the good accessibility of bases in hairpin loops, loop–loop interactions are well suited to trigger molecular recognition between two RNA molecules (9).

We started to analyze the stabilities of loop–loop interactions by using the HIV-1 type dimerization initiation signal (DIS) loop, asking the question whether WC base pairs obey similar rules in the tertiary structure and in the secondary structure context. The DIS hairpin loop consists of six WC base pairs flanked by two adenines or an adenine and a guanine on the 5′ side and one adenine on the 3′ side (5′ $A^A/_GN_6A$ 3′; Figure 1A). We found that indeed, rules governing WC base pair formation are followed, but loop–loop interactions consisting of a variety of six WC base pairs in the DIS context are more stable than the equivalent regular RNA duplexes by a constant factor of 4 kcal/mol in 1 M Na^+ (10). In this regard it is intriguing that the only autocomplementary sequences found in naturally occurring HIV isolates are either 5′-GUGCAC-3′ or 5′-GCGCGC-3′, suggesting that this high conservation of the central sequence is constrained for function. The key bases in these sequences are the central GC nucleotides,

*To whom correspondence should be addressed. Tel: +43 1 4277 54690; Fax: +43 1 4277 9528; Email: renee.schroeder@univie.ac.at

which are involved in a base triple interaction with the flanking nucleotides (11).

Only little is known about the intracellular folding properties of RNA molecules (12). First attempts to compare *in vitro* and *in vivo* properties of functional RNAs were done with the hairpin ribozyme (13,14) and with the T4 phage derived group I *td* intron (15,16). These reports demonstrated that reaction rates of wild-type and mutant ribozymes had similar relative activities *in vivo* and *in vitro* when using near physiological conditions, suggesting that conditions used to perform experiments *in vitro* support, with some limitations, the biological properties of the RNA molecules *in vivo*. Recently it was shown for the hairpin ribozyme in yeast that some feature, which is unique to the intracellular environment promotes folding of the RNA towards the thermodynamically most stable conformation bypassing kinetic traps and transcriptional polarity (17). We are therefore interested in comparing *in vitro* with *in vivo* data. All experiments done to determine thermodynamic parameters of RNA secondary structure for feeding algorithms were performed in 1 M Na^+ (3–5,18,19). For RNA tertiary structure interactions, which are most often dependent on divalent ions for their formation, the use of magnesium might be essential to determine thermodynamic parameters and obtain energy values that correlate with the intracellular conditions. Intracellular salt conditions lie in the range of 0.4 and 1 mM free Mg^{2+} (20).

In the work presented here, we measured the role of the flanking nucleotides in the stability of loop–loop interactions with a constant core sequence (5′-CCGACC-3′ = N_6) that does not contain the functionally constrained GC bases. This motif is derived from the HIV-1 DIS loop sequences LAI (5′-AAN_6A-3′) or MAL (5′-AGN_6A-3′). We analyzed the influence of the flanking nucleotides on the stability of the kissing complexes both *in vitro* and *in vivo*. The rationale for the *in vivo* experiments was the question whether parameters determined under *in vitro* conditions would also be valid within the intracellular environment. We recently developed a yeast system capable of detecting the intracellular interaction of RNA molecules and found that this system was suitable to at least partially quantify the interaction stability induced by the loop–loop contacts via β-galactosidase activity (21).

MATERIALS AND METHODS

RNAs

RNA oligomers for UV-melting experiments were synthesized by Dharmacon and 2′-deprotected as specified by the manufacturer. RNAs were stored at −20°C in water and RNA concentrations were determined after NaOH hydrolysis (pH was adjusted after hydrolysis) and absorption spectroscopy at 260 nm. The investigated RNA sequences were the following: each cognate hairpin pair XY contained a hairpin X with the sequence 5′-GGGCCGN$_n$CCGACCN$_n$CGGCCC-3′ and a hairpin Y with the sequence 5′-CCCGGCN$_n$GGUCGGN$_n$GCCGGG-3′. The nature and number of the flanking nucleotides N_n is depicted in Figure 1B.

X and Y RNAs for the *in vivo* RNA–RNA interaction assay were obtained by cloning of the corresponding insert either into pAN MS2-2, between XmaI and SphI, for X RNAs or into pIIIa m26, between SphI and SalI, for Y RNAs. Inserts were obtained by hybridization of the respective oligomers: X RNAs, 5′-GTCACCCGGG CGN$_n$CCGACCN$_n$ CGCCCGTGCA TGCACTG-3′ and 5′-CAGTGCATGC ACGGGCGN$_n$GG TCGGN$_n$CGCCC GGGTGAC-3′; Y RNAs, 5′-GTCAGCATGC CTGCAGGGCN$_n$ GGTCGGN$_n$GCC CTGCAGGTCG ACACTG-3′ and 5′-CAGTGTCGAC CTGCAGGGCN$_n$ CCGACCN$_n$GCC CTGCAGGCAT GCTGAC-3′, XΘ and YΘ RNA correspond to X_{1b} and Y_{1b} in Piganeau *et al*. (21). For cloning the hybridized oligomers were digested with XmaI/SphI and SphI/SalI, respectively.

UV spectra and calculation of thermodynamic parameters

UV-melting experiments were performed in 10 mM cacodylic acid, pH 6.8. The intrinsic salt concentration of this buffer is 10 mM Na^+ due to the addition of NaOH during pH adjustment. Thermodynamic parameters were determined in 1 M Na^+ final concentration without Mg^{2+} while T_m/log[ion] plots were measured at the indicated salt concentrations. Prior to the melting experiments the corresponding hairpins were mixed in the appropriate buffer at a 1:1 ratio, denatured for 5 min at 95°C and slowly cooled to room temperature. Absorbance versus temperature melting curves were measured at 260 nm with a heating (cooling) rate of 0.5°C/min and four to six consecutive runs on a Cary 100 Bio Varian Spectrophotometer.

Several methods can be used for extracting thermodynamic parameters from melting curves (22). We decided to determine melting transitions by calculating the first derivative of the absorbance versus temperature plot (curves were smoothed with a least square fit of second order polynomials). Although this method leads to systematic errors in the derived melting temperature (T_m) values (an underestimation of about 2°C) it hardly affects the accuracy of ΔG and ΔS values (10) and circumvents the necessity to fit lower and higher temperature base lines, which, in our study, would be error-prone due to the presence of two transitions.

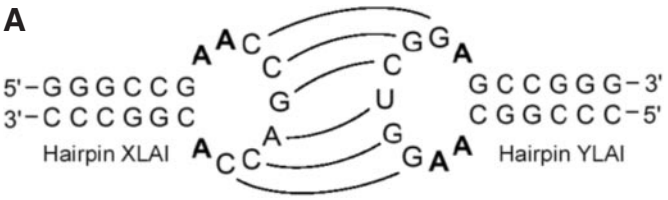
By measuring melting curves at various RNA concentrations and plotting the reciprocal T_ms versus ln(C_T/4), the thermodynamic parameters ΔS, ΔH and ΔG can be derived via a van't Hoff analysis (19). C_T is the total strand concentration, error bars indicate the SD of three independent experiments.

$$\frac{R}{\Delta H^\circ}\cdot \ln\left(\frac{C_T}{4}\right)+\frac{\Delta S^\circ}{\Delta H^\circ}=\frac{1}{T_m}$$

Theoretical free energies $-\Delta G^f$ for regular RNA duplexes with the central sequence 5′ CCGACC 3′ and the respective flanking nucleotides were calculated with the parameters determined by (6).

In vivo RNA-hybrid system and β-galactosidase activity assay

Yeast YBZ-1 cells were co-transformed with plasmids derived from pAN MS2-2 and pIIIa m26. pAN MS2-2 contains the LEU2 gene and expresses an RNA consisting of the MS2 binding domain and RNA X (RNAX-MS2), while pIIIa m26 carries the URA3 gene and expresses an RNA containing the m26 activation domain and RNA Y (m26-RNAY) (21). Transformed cells were selected on media lacking leucine and uracil (23). β-Galactosidase activity was quantitatively



B Loop sequences:

XLAI:	5'-**AA**CCGACC**A** -3'
YLAI:	3'- **A**GGCUGG**AA**-5'
XMAL:	5'-**AG**CCGACC**A** -3'
YMAL:	3'- **A**GGCUGG**GA**-5'
XAA:	5'- **A**CCGACC**A** -3'
YAA:	3'- **A**GGCUGG**A** -5'
XLAinv:	5'- **A**CCGACC**AA**-3'
YLAinv:	3'-**AA**GGCUGG**A** -5'
XMALinv:	5'- **G**CCGACC**AA**-3'
YMALinv:	3'-**AA**GGCUGG**G** -5'
X0:	5'- CCGACC -3'
Y0:	3'- GGCUGG -5'
XUUC:	5'-**UU**CCGACC**C** -3'
YUUC:	3'- **C**GGCUGG**UU**-5'
XUUU:	5'-**UU**CCGACC**U** -3'
YUUU:	3'- **U**GGCUGG**UU**-5'

Figure 1. (**A**) A schematic representation of the kissing interaction and the hairpins we used in this study, using XYLAI as example. The hairpins were derived from the HIV-1 DIS loop and consist each of a stem of six CG base pairs and a 9 nt loop. The six central bases of the loop directly participate in the kissing interaction by forming WC base pairs and were kept constant throughout the study. The three flanking nucleotides (displayed in bold) were varied and the influence of their nature and number upon complex formation was investigated. (**B**) A list of all hairpin pairs XY investigated. Only the loop sequences are shown, with the flanking nucleotides in bold.

determined as described in the Yeast Protocols Handbook from Clontech laboratories, using ONPG (*o*-Nitrophenyl-β-D-galactopyranoside) as substrate.

RESULTS

Design of RNA hairpin loops

We used the HIV-1 type DIS loop as a model system to study RNA loop–loop interactions. The DIS loop is a hairpin structure that contains a 9 nt loop consisting of a 6 nt self-complementary sequence preceded and followed by two and one purines, respectively. This original hairpin was redesigned in respect to the following criteria to suit the requirements of our study: (i) stems consisted of CG base pairs only to ensure a high thermal stability and thus a clear distinction between stem (secondary structure) and loop–loop (tertiary structure) melting transitions, (ii) no self-complementary sequence should be present in the loop to avoid homodimer formation and as a consequence extended duplex formation and (iii) inverted stem sequences were used to avoid extended duplex formation between heterodimers (Figure 1A). This design for RNA loop–loop interactions was previously used in our lab to investigate the properties of WC base pairings in a tertiary structure context (10). We now analyzed the contribution of the flanking nucleotides to the overall stability of the loop–loop interaction by varying the nature and number of the flanking adenines, whilst keeping the central WC base pairing sequence in the loop constant (5′-CCGACC-3′ for hairpins X and 5′-GGUCGG-3′ for hairpins Y) (Figure 1B).

UV-melting profiles of kissing complexes

UV-melting experiments for calculating thermodynamic parameters were performed in 1 M NaCl to allow comparison of our data with other published data. A representative plot of an absorbance versus temperature profile is shown in Figure 2A. Four consecutive heating and cooling runs for the hairpin pair XYMAL are superimposed. The first derivative of this melting curve (dA/dT) plotted versus temperature is shown in Figure 2B. In such a dA/dT plot, the maximum of a peak represents the inflection point of the corresponding melting transition; T_m values can thus be derived. In Figure 2B, the four consecutive heating and cooling profiles calculated from Figure 2A show one transition corresponding to a T_m of 62°C and the beginning of a second transition corresponding to a T_m at above 85°C. For the same hairpin pair, dA/dT versus temperature plots with different RNA concentrations were superimposed in Figure 2C. The first transition is RNA concentration-dependent and shifts from 60 to 66°C, while the second transition is RNA concentration-independent and remains at above 85°C. When the hairpins were melted separately in a monomolecular reaction, only the beginning of the second, concentration-independent transition was observed (Figure 2C). In control experiments with separate melting of the hairpins at lower salt concentrations (25 mM), this second transition became almost fully visible (data not shown). Therefore, we are confident that the concentration-independent transition at high temperature corresponds to the melting of the hairpin stems while the concentration-dependent transition at lower temperature represents the bimolecular melting of the kissing complex. Thermodynamic parameters for RNA kissing complexes were derived from van't Hoff analysis as described in Materials and Methods. As expected, the representative plot in Figure 2D shows that $1/T_m$ is linearly dependent on $\ln(C_T/4)$, where C_T represents the total strand concentration.

Table 1 summarizes the thermodynamic parameters for the investigated RNA loop–loop interactions. The measured T_ms and the calculated thermodynamic parameters $-\Delta H$, $-\Delta S$ and $-\Delta G_{37}$ as well as $-\Delta\Delta G^{k-d}$ are listed. ΔS and ΔH are directly derived from van't Hoff analyses, while ΔG is subsequently calculated by using the derived ΔS and ΔH values. $-\Delta\Delta G^{k-d}$ is the difference in free energy between the kissing interaction and a regular RNA duplex with the same central 6 nt sequence 5′-CCGACC-3′ [$-\Delta G_{37}$ for this duplex equals 9.21 kcal/mol, (10)]. At 1 M Na^+ all analyzed

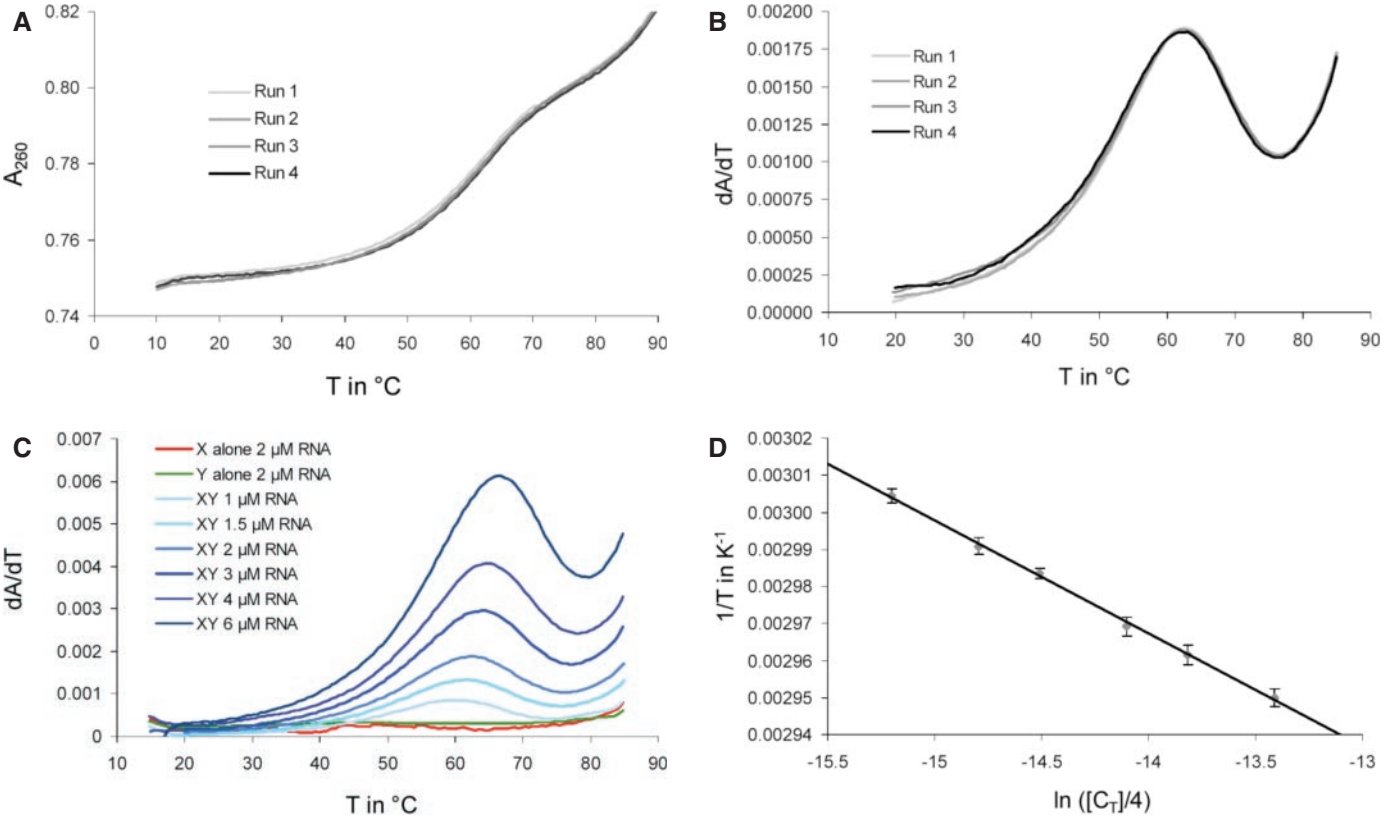


Figure 2. (**A**) An absorbance versus temperature plot. Four consecutive heating and cooling runs of the XYMAL interaction at 2 µM RNA are superimposed. (**B**) Superimposed first derivatives of the melting curves shown above. (**C**) Superimposed first derivatives of the melting curves of the XYMAL interaction for various RNA concentrations and of the hairpins XMAL and YMAL alone. (**D**) Van't Hoff analysis of the melting curves of XYMAL. Error bars indicate the SD of three independent experiments.

kissing complexes are more stable than the equivalent regular RNA duplex, displaying differences in free energy from 0.68 to 4.61 kcal/mol. Furthermore theoretical free energies for regular RNA duplexes possessing the central sequence and the respective flanking nucleotides were calculated ($-\Delta G^{f}$; f = flanking nucleotides). These values were also subtracted from the $-\Delta G_{37}$ values, resulting in $-\Delta\Delta G^{k-f}$, the difference in free energy between the kissing interaction and the respective regular RNA duplex with flanking nucleotides.

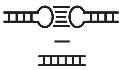
The naturally occurring HIV-1 LAI and HIV-1 MAL type loops (5′-AAN$_6$A-3′) and (5′AGN$_6$A-3′) show the highest $-\Delta G_{37}$ values, indicating that they are the most stable. The (5′-AN$_6$A-3′) loop, containing only one flanking A on either side, is less stable than the LAI and MAL loops but more stable than (5′-AN$_6$AA-3′) and (5′-GN$_6$AA-3′), in which the most 5′ adenine of the LAI and MAL loops is swapped to the 3′ side. The (5′-N$_6$-3′) loop possesses no flanking nucleotides but forms a loop–loop complex that is more stable than a regular RNA duplex by 2.8 kcal/mol, indicating that the tertiary structure context provides additional stability. Replacing the flanking purines with pyrimidines, (5′-UUN$_6$U-3′) and (5′-UUN$_6$C-3′), results in loop–loop complexes that display intermediate stabilities. They are less stable than (5′-AAN$_6$A-3′) and (5′AGN$_6$A-3′), suggesting that purines are preferred to pyrimidines, but more stable than (5′-AN$_6$AA-3′) and (5′-GN$_6$AA-3′).

Influence of ionic strength on melting temperature

The stability of RNA interactions is strongly dependent on the ionic strength of the surrounding environment, as negative charges of the backbone have to be compensated to enable folding into stable structures. Therefore we investigated the influence of monovalent and divalent ions on the thermal stability of RNA loop–loop interactions. UV-melting experiments within a range of Na^+ and Mg^{2+} concentrations were performed for several hairpin pairs and analyzed by plotting the T_m versus the logarithm of the ion concentration. The results are summarized in Table 2. As examples, plots for (5′-UUN$_6$U-3′) and (5′-AN$_6$A-3′) are shown in Figure 3A. As expected, increasing ion concentrations stabilize the loop–loop interactions. A similar dependence on sodium ions (ΔT_m/log[Na^+]) is observed for all interactions, as anticipated for unspecific electrostatic stabilization. In the case of magnesium, however, ion dependence (ΔT_m/log[Mg^{2+}]) varies from construct to construct. Two loop types show a stronger dependence on Mg^{2+} concentration, suggesting at least one additional specific binding site for Mg^{2+} in comparison with the other loop types. These findings indicate that for different loop types, different numbers of ions are bound upon complex formation.

It has been reported that the most 5′A in the hairpin loop is protonated with an approximate pK_a of 6.1–6.4 (24). We therefore tested the influence of pH on the kissing complex stability

Table 1. Thermodynamic parameters of the kissing interactions investigated

Loop–loop sequence	Name	T_m(°C) (2×10^{-6} M RNA)	-ΔH (kcal/mol)	$-\Delta S$ (cal/mol)	$-\Delta G_{37}$ (kcal/mol)	$-\Delta\Delta G^{k-d}$ (kcal/mol)	$-\Delta G^f$ (kcal/mol)	$-\Delta\Delta G^{k-f}$ (kcal/mol)
								
5′–**AA** CCGACC **A**–3′ 3′– **A** GGCUGG **AA**–5′	XYLAI*	59.4	68.0 ± 2.2	175.8 ± 0.9	13.5	4.3	13.0	0.5
5′–**AG** CCGACC **A**–3′ 3′– **A** GGCUGG **GA**–5′	XYMAL	62.0 ± 0.2	65.2 ± 1.0	165.6 ± 1.3	13.8 ± 0.1	4.6	12.5	1.3
5′– **A** CCGACC **A**–3′ 3′– **A** GGCUGG **A**–5′	XYAA	57.4 ± 0.4	56.5 ± 0.1	142.0 ± 1.3	12.4 ± 0.1	3.2	13.0	−0.6
5′– **A** CCGACC **AA**–3′ 3′–**AA** GGCUGG **A**–5′	XYLAlinv	45.1 ± 0.5	36.2 ± 1.9	84.8 ± 5.6	9.9 ± 0.1	0.7	13.0	−3.1
5′– **G** CCGACC **AA**–3′ 3′–**AA** GGCUGG **G**–5′	XYMALinv	46.0 ± 0.6	55.3 ± 0.7	144.6 ± 2.1	10.5 ± 0.1	1.3	12.5	−2.0
5′– CCGACC –3′ 3′– GGCUGG –5′	XYO	58.0 ± 0.3	47.8 ± 4.2	115.6 ± 12.8	12.0 ± 0.2	2.8	9.5	2.5
5′–**UU** CCGACC **C**–3′ 3′– **C** GGCUGG **UU**–5′	XYUUC	54.3 ± 0.2	57.8 ± 2.8	147.7 ± 8.4	12.0 ± 0.2	2.8	10.8	1.2
5′–**UU** CCGACC **U**–3′ 3′– **U** GGCUGG **UU**–5′	XYUUU	49.3 ± 0.1	56.9 ± 0.1	147.5 ± 4.2	11.2 ± 0.6	2.0	11.4	−0.2

T_m is shown for an RNA concentration of 2×10^{-6} M. The four columns to the right show ($-\Delta G_{37}$) ΔG values calculated for 37°C, ($-\Delta\Delta G^{k-d}$) the difference in free energy between the kissing interaction and a regular RNA duplex with the central 6 nt sequence (5′-CCGACC-3′), as measured by (10), ($-\Delta G^f$) calculated values for the stabilities of regular RNA duplexes with respective flanking nucleotides and ($-\Delta\Delta G^{k-f}$) the difference in free energy between the kissing interaction and the calculated stabilities of regular RNA duplexes with respective flanking nucleotides. (*) Data taken from (10).

Table 2. The influence of the ionic strength on the thermal stability of various loop–loop interactions is shown

Loop type	ΔT_m/lg[Na^+]	ΔT_m/lg[Mg^{2+}]
XYLAI	26.1	16.6
XYAA	23.1	20.1
XYUUU	22.9	12.2
XYO	23.0	10.6
XYLAlinv	23.3	9.1

Dependence on sodium or magnesium is indicated as ΔT_m/log[ion].

and found no role for the protonation in its stabilization (data not shown).

As previously observed, similar complex stabilities are obtained with Na^+ and Mg^{2+} concentrations that differ by three orders of magnitude (10). Measurements taken in 1 M Na^+ therefore result in similar thermal stabilities for loop–loop interactions as measurements taken in 1 mM Mg^{2+} (including 10 mM intrinsic Na^+), a concentration close to physiological conditions. This observation suggests that thermodynamic parameters determined *in vitro*, mostly at 1 M Na^+, might correlate with stabilities in a cellular environment where the concentration of Mg^{2+} ions is ~0.8 mM (20). To verify this assumption, we compared T_m values of all constructs at 0.8 mM Mg^{2+} and 1 M Na^+, shown in Figure 3B. In general, thermal stabilities at 0.8 mM Mg^{2+} and 1 M Na^+ are similar, except for two constructs (XYLAIinv and XYMALinv, where the asymmetric purines were swapped to the 3′ side). Their T_m values are significantly higher in 0.8 mM Mg^{2+} than in 1 M Na^+, by 6 and 9°C, respectively. To experimentally compare loop–loop interactions *in vitro* and *in vivo*, we analyzed the stabilities of all kissing complexes in a recently developed *in vivo* yeast RNA–RNA interaction assay.

Comparison of kissing interaction stabilities *in vitro* and *in vivo*

For the *in vivo* RNA–RNA interaction assay the yeast strain YBZ-1 was used (21). This strain is auxotrophic for leucine, uracil, adenine and histidine, carries *lacZ* and *HIS3* genes under the control of multiple lexA operators and expresses a LexA-MS2 coat protein fusion. This strain was co-transformed with the plasmids pAN MS2-2 and pIIIa m26, which express the individual hairpin RNAs X and Y as fusions to either the MS2 binding domain (RNAX-MS2) or the m26 activation domain (m26-RNAY). Upon interaction of the hairpins the m26 domain is tethered to the promoter of *lacZ*. This leads to activation of the reporter gene and expression of β-galactosidase according to the strength of the interaction (Figure 4). In order to determine the degree of activation and thus the stability of the loop–loop interaction, β-galactosidase activity was quantified as described in Materials and Methods and compared with the thermal stabilities obtained by UV melting at 1 M Na^+. In order to prove that the β-galactosidase activity we were observing was indeed due to the interaction of the corresponding hairpins XY, we co-transformed each RNAX-MS2 with a non-cognate m26-RNAYΘ and similarly each m26-RNAY with a non-cognate RNAXΘ-MS2. These transformants showed only background levels of β-galactosidase activity, as did transformants expressing only either XLAI-MS2 or m26-YLAI or no hairpin RNA at all (Figure 5A and C). In general, a rather good agreement in the relative stabilities of the different constructs was observed (Figure 5A). There is a linear correlation between the T_m determined *in vitro* at 1 M Na^+ and the β-galactosidase activity assayed *in vivo* for kissing interactions with an identical number of nucleotides in the loop (Figure 5B). Two previously analyzed complexes, LAIcore1 and LAIcore2 (20), containing different core

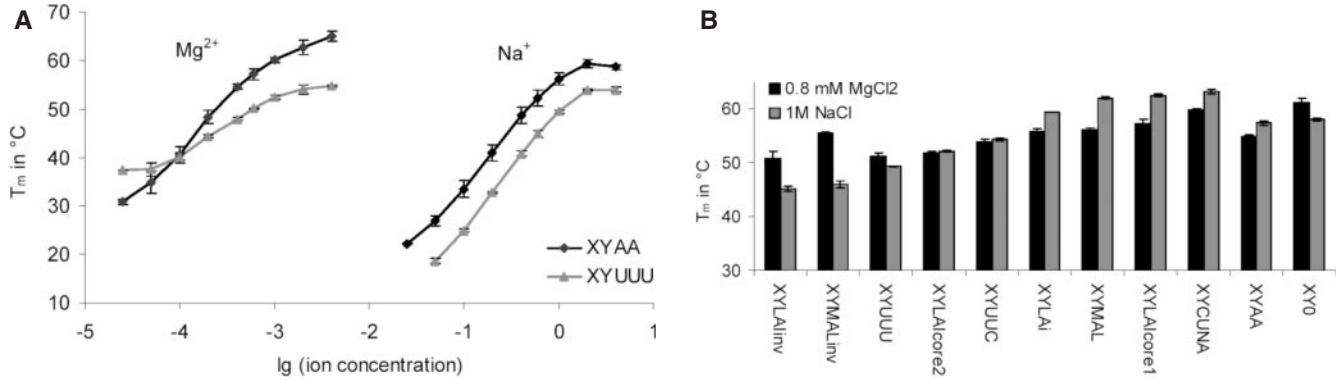


Figure 3. (**A**) Influence of the ionic strength of sodium and magnesium on the thermal stability of kissing complexes XYAA and XYUUU. The T_m was plotted versus the logarithm of the ion concentration. (**B**) T_m values for the indicated loop–loop interactions at 0.8 mM Mg^{2+} (black) or 1M Na^+ (gray). Error bars indicate the SD of three independent experiments. Three additional constructs were included, XYLAIcore1, YXLAIcore2 and XYCUN$_6$A, which are described in the text and in Figure 5.

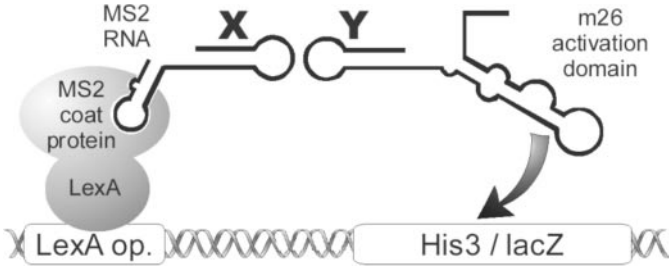


Figure 4. A schematic representation of the yeast RNA-hybrid system. The yeast strain YBZ-1 expresses a lexA-MS2 coat protein fusion and contains *lacZ* and *HIS3* genes under the control of several lexA operators. Upon interaction of the co-expressed hybrid RNAs RNAX-MS2 and m26-RNAY the m26 activation domain is tethered to the promoter of *lacZ* and activates gene expression in correlation to the strength of the interaction.

sequences (5′-AACCUGCCA-3′ and 5′-AAGAGAGGA-3′, respectively) as well as 5′-CUN$_6$A-3′, a loop–loop interaction containing eight complementary base pairs, were included in Figure 5 to show the reliability of the assay with additional sequences. The constructs (5′-AN$_6$A-3′ = XYAA) and (5′-N$_6$-3′ = XY0), which contain a different number of nucleotides in the loop, are exceptions. Although they have nearly the same T_ms *in vitro*, they show entirely different β-galactosidase activities. However, for kissing complexes with an identical number of nucleotides, thermal stabilities derived from *in vitro* UV-melting measurements at 1 M Na^+ correlate well with stabilities obtained from *in vivo* β-galactosidase activity assays.

DISCUSSION

Loop–loop interactions occurring in nature are the most stable

We used UV melting to determine the stabilities of RNA loop–loop interactions using a constant 6 bp core with varying flanking nucleotides. Several variations of these loops are found in isolates of HIV or related viruses (25). Our results show that the bases flanking the loop core sequence have a strong influence on the stability and on the ion dependence of the kissing interactions. The most stable complexes are those whose flanks resemble the naturally occurring HIV-1 LAI (5′-AAN$_6$A-3′) and MAL (5′-AGN$_6$A-3′) sequences. These complexes are more stable than regular RNA duplexes without flanking nucleotides by 4.3 and 4.6 kcal/mol, respectively. The asymmetric distribution of the flanking purines is not restricted to the DIS of HIV-1 MAL and LAI subtypes but is also present in related regions of other retroviruses, such as bovine leukemia viruses and avian viruses. These regions characteristically contain one more flanking purine at the 5′ side than at the 3′ side (26). The crystal structures of the HIV-1 MAL and LAI kissing complexes (27) as well as of the anticodon–anticodon interaction in $tRNA^{Asp}$ (28) and a loop–loop interaction between bases 418–423 and 2444–2449 in the 23S rRNA (29) show coaxial alignment of the two opposite stems due to unpaired bases at the 5′ side. The stabilization observed in the UV-melting data for this type of 5′ configuration is therefore likely to be caused by coaxial stacking.

Studies for the HIV-1 TAR hairpin loop using molecular dynamic simulations suggest that purine mismatches, mostly GA, as loop closing mismatches are water mediated. They are the best compromise between flexibility and stability through optimized stacking of the bases at the stem–loop junction and through formation of stable interbackbone hydrogen bonds (30,31). A GA mismatch as loop closing mismatch was also obtained in an *in vitro* selection for an RNA aptamer raised against the HIV TAR element (32). In a very comprehensive study, Lodmell *et al*. (33) isolated HIV-1 DIS loop sequences, which are best suited to undergo dimerization using SELEX. These studies resulted in a non-cannonical A-A base pair between positions 1 and 9 as the most frequently obtained loop closing mismatch. This correlates well with our findings, that (5′–AAN$_6$A-3′) and (5′-AGN$_6$A-3′) form the most stable kissing complexes.

Removal of one dangling adenine at the 5′ end leaving only one AA mismatch on each side of the loop–loop interaction (5′-AN$_6$A-3′) still results in a very stable complex. It is more stable by 3.2 kcal/mol than the regular helix, but already 1 kcal/mol less stable than the (5′-AAN$_6$A-3′) loop. This type of DIS loop was found in isolates HIV-1 OYI and SIVCPZ GAB and also exists in nature (25).

Interestingly, the loop–loop interaction with no flanking nucleotides (5′-N$_6$-3′) is more stable than a regular RNA duplex by 2.8 kcal/mol. This is reminiscent of the

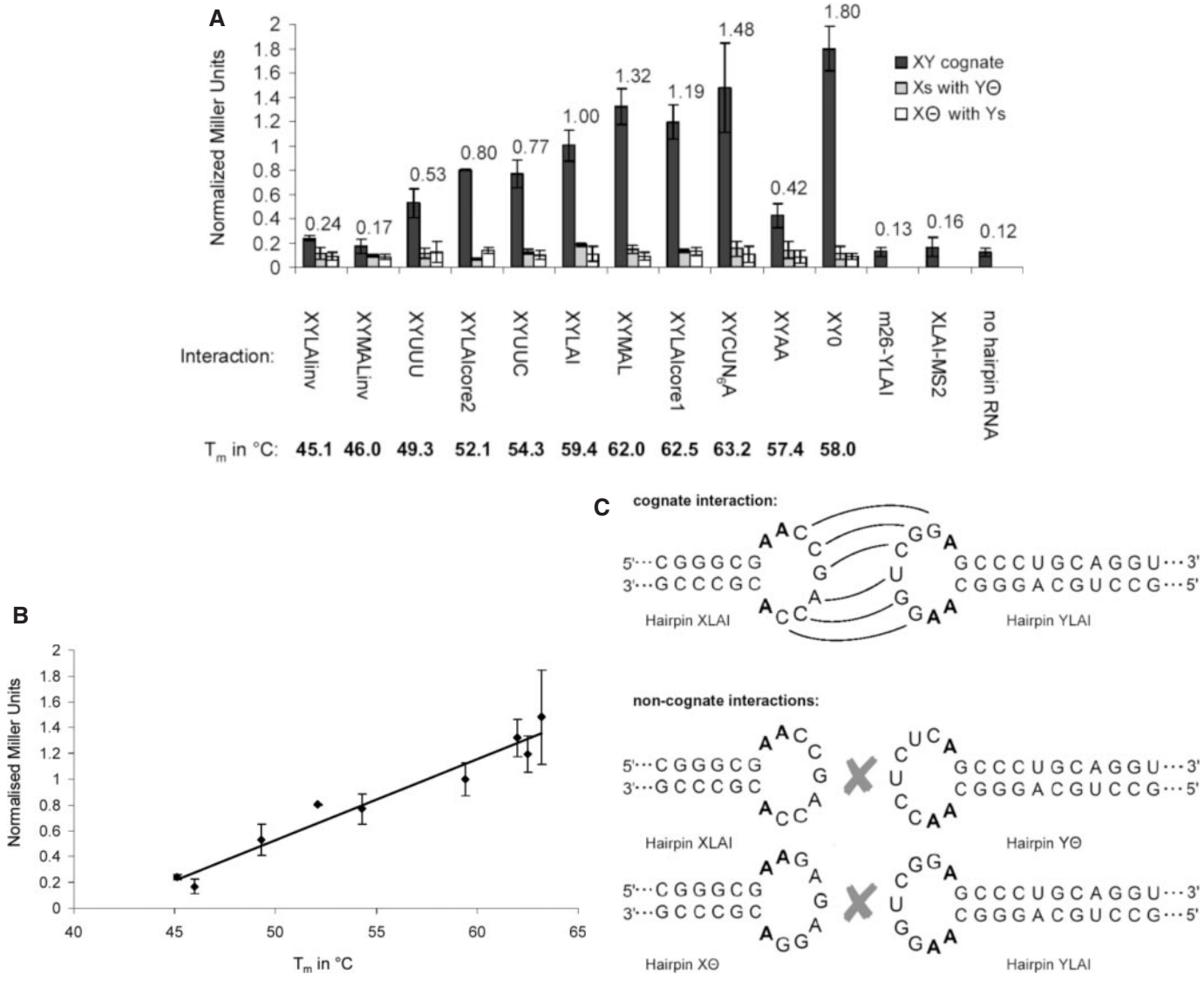


Figure 5. (**A**) β-Galactosidase assay of the different loop–loop interactions investigated. β-Galactosidase activity is expressed in Miller Units and was normalized relative to the activity of the XYLAI interaction. All hairpins were tested together with their cognate interaction partner (closed bars). As control all RNAs X and Y were assayed against non-cognate interaction partners, namely YΘ for RNAs X (light gray bars) and XΘ for RNAs Y (open bars). Furthermore transformants expressing only either XLAI-MS2 or m26-YLAI or no hairpin RNA at all are shown. T_ms for the cognate kissing complexes at a concentration of 2×10^{-6} M are indicated. Error bars show the SD of two independent triplicate experiments. (**B**) Correlation between the T_m *in vitro* and the β-galactosidase activity determined *in vivo* for kissing interactions containing 9 nt in the loop. (XYAA and YX0 contain 8 and 6 nt, respectively, and are not included.) (**C**) Schematic representation of cognate and non-cognate interactions considering XYLAI as example.

RNAI–RNAII complex, which consists of a loop–loop interaction of 7 bp without flanking sequences and continuous helical stacking from one stem through the loop–loop helix to the other stem. Further characteristics of this loop–loop motif are a strong bend in the loop–loop helix and bridging of the backbones across the major groove (34). This suggests that in the (5′-N_6-3′) interaction the two hairpin stems also stack on the loop–loop helix and thereby contribute to the stability of the kissing interaction.

Swapping the asymmetric purine to the 3′ side is detrimental to kissing complex formation

The most unexpected influence of the flanking sequence on the loop–loop stability was obtained by swapping the 5′ dangling adenine to the 3′ side. This destabilizes the interactions by 3.6 kcal/mol (5′-AN_6AA-3′) and 3.3 kcal/mol (5′-GN_6AA-3′), respectively, when compared with the MAL and LAI sequences. Hence, having less purines at the 5′ side than at the 3′ side results in an unfavorable configuration of the loop–loop interaction. The stems might perhaps no longer form a perfect coaxial alignment, resulting in a less stable interaction. The loop sequences (5′-AN_6AA-3′) and (5′-GN_6AA-3′) can be compared with the naturally occurring HIV-2 DIS loop, which also possesses one more purine at the 3′ side than at the 5′ side. In contrast to the HIV-1 DIS loops, which are clearly key players in genome dimerization, the role of the HIV-2 DIS loop in this process is unclear. The DIS loop itself could be sequestered intramolecularly by dimer-interfering elements and two additional dimerization elements exist in the HIV-2

leader RNA (35). Considering our findings, a possible explanation for the non-exclusive use of the HIV-2 DIS loop in the dimerization process might be due to its decreased kissing complex stability.

Switching from purines to pyrimidines induces destabilization of the kissing interaction

Changing the flanking purines to pyrimidines (5′-UUN$_6$U-3′ or 5′-UUN$_6$C-3′) still results in loop–loop interactions that are more stable than the equivalent regular RNA helixes without flanking nucleotides but less stable than kissing complexes flanked by purines. In the HIV-1 LAI and MAL kissing complexes, the most 5′ A and the 3′ A form a non-canonical intramolecular A-A pair by using the WC and Hoogsteen sides, respectively (36). This might explain the difference observed between flanking purines and pyrimidines. Furthermore, pyrimidine pairs as loop closing mismatches have been proposed to be too short to achieve a tight interaction, although the base mismatch is mediated by a water molecule (30). When comparing loop–loop interactions with flanking pyrimidines, we find that (5′-UUN$_6$U-3′) is less stable than (5′-UUN$_6$C-3′) by 0.8 kcal/mol. This is in good agreement with the observation that UU pairs as loop closing mismatches stabilize hairpins by 0.8 kcal/mol (37). Therefore, in the (5′-UUN$_6$U-3′) context, the single hairpins are stabilized by the UU pair and as a consequence kissing complex formation out of two such hairpins might be partially constrained making the (5′-UUN$_6$U-3′) interactions less stable than the (5′-UUN$_6$C-3′) interaction.

Good correlation of loop–loop stabilities determined *in vitro* with *in vivo* β-galactosidase activities

We studied kissing complex stabilities in a variety of Na^+ and Mg^{2+} concentrations and obtained similar thermal stabilities for the same interaction when we used Na^+ and Mg^{2+} concentrations that differed by approximately three orders of magnitude. Stabilities measured in 1 M Na^+ therefore coincide approximately with measurements in 1 mM Mg^{2+} (including 10 mM intrinsic Na^+), suggesting that stabilities obtained *in vitro* might correlate with stabilities inside a cellular environment. Therefore we analyzed the loop–loop complexes in a recently developed *in vivo* yeast RNA–RNA interaction assay.

The hairpin pairs containing the various HIV-1 derived sequences were cloned into appropriate constructs for expression in yeast and the strength of their interaction was monitored via their activation potential of the *lacZ* gene (21). We found a good linear correlation between the T_ms determined *in vitro* and the β-galactosidase activities in yeast (Figure 5). Two exceptions were observed, constructs AA (5′-AN$_6$A-3′) and 0 (5′-N$_6$-3′). Both constructs have a similar T_m but show very dissimilar β-galactosidase activities. In contrast to all other constructs they do not have 9 nt in their loops but 8 and 6, respectively. This could result in a different geometry of the loop–loop interaction and as a consequence the activation domain might be positioned differently relative to its interaction partner, which would lead to a different induction of the promoter.

In summary, the work presented demonstrates that character, number and position (5′ or 3′) of dangling nucleotides in loop–loop interactions have a significant influence on the stability of kissing complexes. Our results show that naturally occurring loop–loop interactions possess a composition of flanking nucleotides that provides maximal stability, indicating that nature selected for the most stable complexes. UV melting measurements with varying ion concentrations reveal that stabilities determined in 1 M Na^+ coincide with stabilities in 1 mM $MgCl_2$, a near physiological concentration, and experiments in yeast confirm that *in vitro* measurements appropriately describe *in vivo* properties of RNA kissing complexes.

ACKNOWLEDGEMENTS

This work was supported by the Austrian Science Fund FWF grants no. P16026 and Z-72. Funding to pay the Open Access publication charges for this article was provided by the FWF.

Conflict of interest statement. None declared.

REFERENCES

1. Brion,P. and Westhof,E. (1997) Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.*, **26**, 113–137.
2. Tinoco,I.,Jr and Bustamante,C. (1999) How RNA folds. *J. Mol. Biol.*, **293**, 271–281.
3. SantaLucia,J.,Jr and Turner,D.H. (1997) Measuring the thermodynamics of RNA secondary structure formation. *Biopolymers*, **44**, 309–319.
4. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.
5. Jaeger,J.A., SantaLucia,J.,Jr and Tinoco,I.,Jr (1993) Determination of RNA structure and thermodynamics. *Annu. Rev. Biochem.*, **62**, 255–287.
6. Xia,T., SantaLucia,J.,Jr, Burkard,M.E., Kierzek,R., Schroeder,S.J., Jiao,X., Cox,C. and Turner,D.H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson–Crick base pairs. *Biochemistry*, **37**, 14719–14735.
7. Onoa,B. and Tinoco,I.,Jr (2004) RNA folding and unfolding. *Curr. Opin. Struct. Biol.*, **14**, 374–379.
8. Wagner,E.G., Altuvia,S. and Romby,P. (2002) Antisense RNAs in bacteria and their genetic elements. *Adv. Genet.*, **46**, 361–398.
9. Brunel,C., Marquet,R., Romby,P. and Ehresmann,C. (2002) RNA loop–loop interactions as dynamic functional motifs. *Biochimie*, **84**, 925–944.
10. Weixlbaumer,A., Werner,A., Flamm,C., Westhof,E. and Schroeder,R. (2004) Determination of thermodynamic parameters for HIV DIS type loop–loop kissing complexes. *Nucleic Acids Res.*, **32**, 5126–5133.
11. Lodmell,J., Ehresmann,C., Ehresmann,B. and Marquet,R. (2001) Structure and dimerizytion of HIV-1 kissing loop aptamers. *J. Mol. Biol.*, **311**, 475–490.
12. Schroeder,R., Grossberger,R., Pichler,A. and Waldsich,C. (2002) RNA folding *in vivo*. *Curr. Opin. Struct. Biol.*, **12**, 296–300.
13. Donahue,C.P. and Fedor,M.J. (1997) Kinetics of hairpin ribozyme cleavage in yeast. *RNA*, **3**, 961–973.
14. Donahue,C.P., Yadava,R.S., Nesbitt,S.M. and Fedor,M.J. (2000) The kinetic mechanism of the hairpin ribozyme *in vivo*: influence of RNA helix stability on intracellular cleavage kinetics. *J. Mol. Biol.*, **295**, 693–707.
15. Brion,P., Schroeder,R., Michel,F. and Westhof,E. (1999) Influence of specific mutations on the thermal stability of the *td* group I intron *in vitro* and on its splicing efficiency *in vivo*: a comparative study. *RNA*, **5**, 947–958.
16. Waldsich,C., Masquida,B., Westhof,E. and Schroeder,R. (2002) Monitoring intermediate folding states of the td group I intron *in vivo*. *EMBO J.*, **19**, 5281–5291.
17. Mahen,E.M., Harger,J.W., Calderon,E.M. and Fedor,M.J. (2005) Kinetics and thermodynamics make different contributions to RNA folding *in vitro* and in yeast. *Mol. Cell*, **19**, 27–37.
18. Tinoco,I.,Jr, Borer,P.N., Dengler,B., Levin,M.D., Uhlenbeck,O.C., Crothers,D.M. and Bralla,J. (1973) Improved estimation of secondary structure in ribonucleic acids. *Nature New Biol.*, **246**, 40–41.

19. Serra,M.J. and Turner,D.H. (1995) Predicting thermodynamic properties of RNA. *Methods Enzymol*., **259**, 242–261.
20. Romani,A. and Scarpa,A. (1992) Regulation of cell magnesium. *Arch. Biochem. Biophys*., **298**, 1–12.
21. Piganeau,N., Schauer,U.E. and Schroeder,R. (2005) A yeast RNA-hybrid system for the detection of RNA–RNA interactions in vivo. *RNA*, **12**, 177–184.
22. Tinoco,I.,Jr and Schmitz,M. (2000) Thermodynamics of formation of secondary structure in nucleic acids. In Cera,E.D. (ed.), *Thermodynamics in Biology*. Oxford University Press, Oxford, pp. 131–176.
23. Bernstein,D.S., Buter,N., Stumpf,C. and Wickens,M. (2002) Analyzing mRNA-protein complexes using a yeast three-hybrid system. *Methods*, **26**, 123–141.
24. Mihailescu,M.R. and Marino,J.P. (2004) A proton-coupled dynamic conformational switch in the HIV-1 dimerization initiation site kissing complex. *Proc. Natl Acad. Sci. USA*, **101**, 1189–1194.
25. Human Retroviruses and AIDS (1995) A compilation and analysis of nucleic acid and amino acid sequences. In Myers,G., Korber,B., Wain-Hobson,S., Jeang,K.-T., Mellors,J.W., McCutchan,F.E., Henderson,L.E. and Pavlakis,G.N. (eds), Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.
26. Paillart,J.C., Marquet,R., Skripkin,E., Ehresmann,C. and Ehresmann,B. (1996) Dimerization of retroviral genomic RNAs: structural and functional implications. *Biochimie*, **78**, 639–653.
27. Ennifar,E., Walter,P., Ehresmann,B. and Ehresmann,C. (2001) Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNAdimerization initiation site. *Nature Struct. Biol*., **8**, 1064–1068.
28. Westhof,E., Dumas,P. and Moras,D. (1985) Crystallographic refinement of yeast aspartic acid transfer RNA. *J. Mol. Biol*., **184**, 119–145.
29. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, **289**, 905–920.
30. Golebiowski,J., Antonczak,S., Fernandez-Carmona,J., Condom,R. and Cabrol-Brass,D. (2004) Closing loop base pairs in RNA loop–loop complexes: structural behaviour, interaction energy and solvation analysis through molecular dynamic simulations. *J. Mol. Model*., **10**, 408–417.
31. Beaurain,F., Di Primo,C., Toulme,J.J. and Laguerre,M. (2003) Molecular Dynamic reveals the stabilizing role of loop closing residues in kissing interactions: comparison between TAR-TAR* and TAR aptamer. *Nucleic Acids Res*., **31**, 4275–4284.
32. Duconge,F. and Toulme,J.J. (1999) *In vitro* selection identifies key determinants for loop–loop interactions: RNA aptamers selective for the TAR RNA element of HIV-1. *RNA*, **5**, 1605–1614.
33. Lodmell,J.S., Ehresmann,C., Ehresmann,B. and Marquet,R. (2000) Convergence of natural and artificial evolution on an RNA loop–loop interaction: The HIV-1 dimerization initiation site. *RNA*, **6**, 1267–1276.
34. Lee,A.J. and Crothers,D.M. (1998) The solution structure of an RNA loop–loop complex: the ColE1 inverted loop sequence. *Structure*, **6**, 993–1005.
35. Lanchy,J.M., Ivanovitch,J.D. and Lodmell,J.S. (2003) A structural linkage between the dimerization and encapsidation signals in HIV-2 leader RNA. *RNA*, **9**, 1007–1018.
36. Paillart,J.C., Westhof,E., Ehresmann,C., Ehresmann,B. and Marquet,R. (1997) Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA. *J. Mol. Biol*., **270**, 36–49.
37. Vecenie,C.J. and Serra,M.J. (2004) Stability of RNA hairpin loops closed by AU base pairs. *Biochemistry*, **43**, 11813–11817.