Stem–loop formation drives RNA folding in mechanical unzipping experiments

Paolo Rissone[a], Cristiano V. Bizarro[b], and Felix Ritort[a,1]

[a]Small Biosystems Laboratory, Condensed Matter Physics Department, University of Barcelona, Barcelona 08028, Spain; and [b]Instituto Nacional de Ciência e Tecnologia em Tuberculose, Centro de Pesquisas em Biologia Molecular e Funcional, Pontifícia Universidade Católica do Rio Grande do Sul, 90616-900 Porto Alegre, Brazil

Edited by Zev Bryant, Bioengineering, Stanford University, Stanford, CA; received December 12, 2020; accepted November 29, 2021 by Editorial Board Member Yale E. Goldman

Accurate knowledge of RNA hybridization is essential for understanding RNA structure and function. Here we mechanically unzip and rezip a 2-kbp RNA hairpin and derive the 10 nearest-neighbor base pair (NNBP) RNA free energies in sodium and magnesium with 0.1 kcal/mol precision using optical tweezers. Notably, force–distance curves (FDCs) exhibit strong irreversible effects with hysteresis and several intermediates, precluding the extraction of the NNBP energies with currently available methods. The combination of a suitable RNA synthesis with a tailored pulling protocol allowed us to obtain the fully reversible FDCs necessary to derive the NNBP energies. We demonstrate the equivalence of sodium and magnesium free-energy salt corrections at the level of individual NNBP. To characterize the irreversibility of the unzipping–rezipping process, we introduce a barrier energy landscape of the stem–loop structures forming along the complementary strands, which compete against the formation of the native hairpin. This landscape correlates with the hysteresis observed along the FDCs. RNA sequence analysis shows that base stacking and base pairing stabilize the stem–loops that kinetically trap the long-lived intermediates observed in the FDC. Stem–loops formation appears as a general mechanism to explain a wide range of behaviors observed in RNA folding.

RNA biophysics | nucleic acids thermodynamics | statistical mechanics | fluctuation theorems

Unzipping experiments permit one to investigate the physicochemical properties of nucleic acids, from the thermodynamics of duplex formation to the folding of secondary and tertiary structures. In particular, DNA hybridization finds diverse applications in the field of DNA nanotechnology, the construction of DNA origami, molecular robots, DNA walkers, switches, and nanomotors (1–5). In an unzipping experiment, the two strands of a duplex DNA or RNA molecule are mechanically pulled apart by exerting opposite forces on the two strands on one end. In this way, it is possible to measure a force–distance curve (FDC) that exhibits a sequence-dependent sawtooth pattern. DNA unzipping has been used to test the validity of the nearest-neighbor (NN) model (6–9) and to extract the 10 NN base pairs (NNBP) free-energy parameters at different salt conditions (10, 11). A precise knowledge of the NNBP energies might be also useful to unravel hidden energy codes in molecular evolution (12).

Here we derive the 10 NNBP RNA energies from unzipping experiments carried out on a 2-kbp RNA hairpin in monovalent (sodium) and divalent (magnesium) salt conditions. The NN model has many parameters requiring a sufficiently long RNA hairpin to infer them from unzipping experiments reliably. Two are the main difficulties of these experiments: First, the molecular synthesis of a long (a few kilobases) RNA hairpin is challenging; second, the FDC along the RNA sequence alternates reversible unzipping regions with irreversible ones that exhibit hysteresis and multiple long-lived intermediates (13, 14). Compared to DNA, where unzipping is practically reversible, a similar derivation of the RNA energies from irreversible FDCs is not possible. Here we derive the full equilibrium FDC in RNA by the piecewise assembly of the reversible parts and the reconstructed equilibrium ones for the irreversible regions. These are obtained by repeatedly unzipping and rezipping the RNA hairpin in these irreversible regions and using statistical physics methods based on fluctuation theorems. This allows us to derive the NNBP energies for RNA in sodium and magnesium and compare them with the results reported by the literature (15–18). Moreover, we demonstrate the validity of an equivalence rule for the free-energy salt corrections between sodium and magnesium at the level of individual NNBP. We find that NNBP free-energy parameters for a given magnesium concentration are equal to those in $77(\pm 49)$-fold sodium. This result is compatible with the 100/1 rule of thumb by which the nonspecific RNA binding affinity of 10 mM Mg^{2+} approximately equals that of 1 M Na^{+} (19). We provide a solid verification of this phenomenological result by measuring the NNBP RNA energies in sodium and magnesium. We study the irreversibility and hysteresis in the FDCs and hypothesize that this is caused by the formation of stem–loop structures along the unpaired single strands. Remarkably, the hysteresis along the unzipping–rezipping pathway directly correlates with the barrier energy landscape defined by the stem–loops that are formed at the junction separating single strands and duplex. A sequence analysis of the irreversible

Significance

Understanding how complementary RNA single strands anneal to form the native structure requires accurate knowledge of the energetics of hybridization and the kinetics driving the hairpin fold. The annealing dynamics are highly complex for RNA, where multiple long-lived intermediate states occur along the folding pathway. In this framework, single-molecule experiments allow us to derive the thermodynamics and kinetics of RNA folding and the characterization of the intermediates along the folding pathway. We mechanically unzip a 2-kbp RNA hairpin with optical tweezers in sodium and magnesium, deriving the 10 RNA nearest-neighbor free energies and modeling the hybridization kinetics in a barrier energy landscape of stem–loop secondary structures. Our study contributes to a better understanding of the RNA folding problem.

Author contributions: C.V.B. and F.R. designed research; P.R. and C.V.B. performed research; P.R. analyzed data; and P.R., C.V.B., and F.R. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. Z.B. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: ritort@ub.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2025575119/-/DCSupplemental.

Published January 11, 2022.

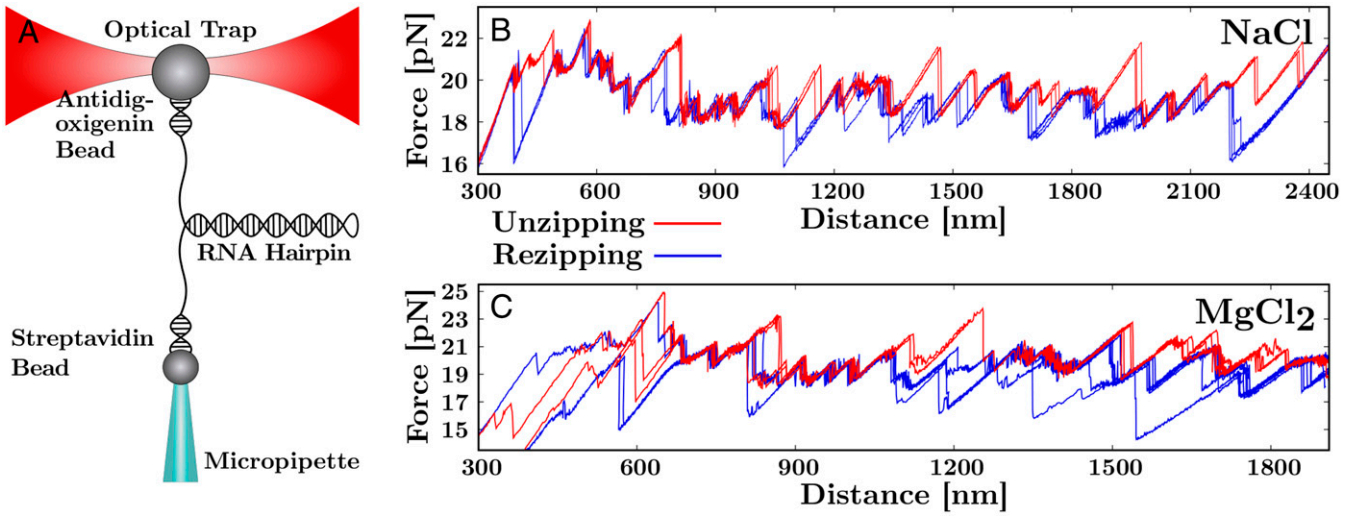

Fig. 1. Experimental setup and measured FDCs in sodium and magnesium. (*A*) Optical tweezers setup. The RNA hairpin is mechanically unzipped and rezipped by moving the optical trap. (*B*) Unzipping/rezipping FDCs (red/blue) in 500 mM NaCl. Hysteresis is apparent in some regions of the FDC. (*C*) Unzipping/rezipping FDCs in 10 mM $MgCl_2$. Magnesium enhances the amount of hysteresis compared to the sodium case. The irreversibility is so large that the initial and final regions of the FDCs remain inaccessible.

regions of the 2-kbp RNA and experiments on specifically designed short-RNA sequences demonstrates that base stacking and base pairing within the single-stranded RNA (ssRNA) promote the formation of stem–loop RNA structures transiently stabilized at forces as high as 20 pN. The stem–loops mechanism explains the slow kinetics and multiple trapping conformations observed in RNA folding, with implications for the RNA folding problem (13, 20–23).

Results

We used optical tweezers to pull a 2,027-bp RNA hairpin with short (29 bp) hybrid DNA/RNA handles. Details on the hairpin and the synthesis protocol are given in *Materials and Methods*. In our setup, the digoxigenin (DIG)-labeled and biotin-labeled handles of the hairpin are connected to anti-DIG (AD) and streptavidin-coated (SA) beads, respectively. The AD bead is optically trapped while the SA bead is immobilized by air suction at the tip of a micropipette (Fig. 1*A*). By moving the optical trap upward, the molecule gradually unzips from the completely folded double-stranded RNA (dsRNA) hairpin conformation (the native state *N*) to the completely unfolded and stretched ssRNA conformation (the unfolded state *U*), producing the characteristic sawtooth pattern of the FDC (red curves in Fig. 1 *B* and *C*). Once the hairpin is unfolded, the reverse process (rezipping) starts: The trap is moved in the opposite direction and the molecule gradually refolds into the native stem (blue curves in Fig. 1 *B* and *C*).

The experiments have been performed in buffers containing 100 mM tris(hydroxymethyl)aminomethane (Tris)-HCl (pH 8.1), 1 mM ethylenediaminetetraacetic, and 500 mM NaCl (monovalent salt) or 100 mM Tris·HCl (pH 8.1) and 10 mM $MgCl_2$ (divalent salt). Note that the ionic strength of the buffers has to be corrected by adding 100 mM Tris·HCl $\equiv 52$ mM $[Mon^+]$. Measured FDCs show that changing from $[Na^+]$ to $[Mg^{++}]$ strongly increases the irreversibility and hysteresis of the FDC. This makes the beginning (the first 200 bp, between 400 and 650 nm) and the end (the last 600 bp, between 1,800 and 2,200 nm) of the FDC experimentally inaccessible: The RNA hairpin does not hybridize in the experimental timescale (Fig. 1*C*). The observed hysteresis occurs in correspondence with specific regions along the FDC, each one limited by the equilibrated left (L) and right (R) states and exhibiting intermediate states I_p, with $p = 1, \ldots, P$. To efficiently sample the intermediates, we repeatedly unzipped and rezipped the RNA between the two limit positions (L, R), typically collecting 100 trajectories per region. We have identified eight irreversible regions in sodium (Fig. 2) and three in magnesium (Fig. 3). Regions in sodium are numbered from 1 to 8. In magnesium, regions are numbered from 2 to 4/5 to underline the matching of the RNA sequences in those regions in sodium and magnesium, as evidenced by the number of opened base pairs. The larger hysteresis observed in magnesium makes regions 4 and 5 in sodium merge into a single irreversible region (4/5). The missing regions (1 and 6 to 8) in magnesium result from their inaccessibility, as explained above. Although a few regions do not contain intermediates (e.g., region 5 in Fig. 2*A*), most of them exhibit more than one. The level of complexity of the unzipping–rezipping FDCs can be high; e.g., region 3 in magnesium shows seven states (five intermediates plus L and R, Fig. 3*B*).

To derive the NNBP energies we computed the equilibrium FDC by applying the extended fluctuation relation (*Materials and Methods*), which has been introduced to recover the free energy of thermodynamic branches (24), kinetic states (25), and ligand-binding energies (26, 27). This allowed us to reconstruct the equilibrium FDCs (black line in Figs. 2 and 3) for seven molecules in sodium and four molecules in magnesium.

Derivation of the NNBP Energies for RNA. In the NN model the free energy of formation ΔG_0 of a DNA and RNA duplex is defined as the sum over all adjacent NNBPs along the sequence, $\Delta G_0 = \sum_i \Delta g_{0,i}$ with $\Delta g_{0,i}$ the free energy of NNBP motif i. There are 16 different NNBPs whose energies are degenerated due to Watson–Crick complementarity, reducing the free energies set $(\Delta g_{0,i})$ to 10 parameters. The NNBP energies have been extracted from melting experiments of short RNA duplexes of varying sequence and length (15–18). These values are accessible in the Mfold server (28). Hereafter we refer to such energies as the RNA Mfold values. It is possible to further reduce this number from 10 to 8 independent parameters by considering the circular symmetry of the NN model (29, 30). This symmetry yields additional self-consistent relations for the dimer occupancies along the duplex: Of the 10 NNBP energies 2 can be expressed as linear combinations of the remaining 8 (11, 30, 31). The circular symmetry property has been tested and the eight parameters have been derived in DNA unzipping experiments (10, 11).

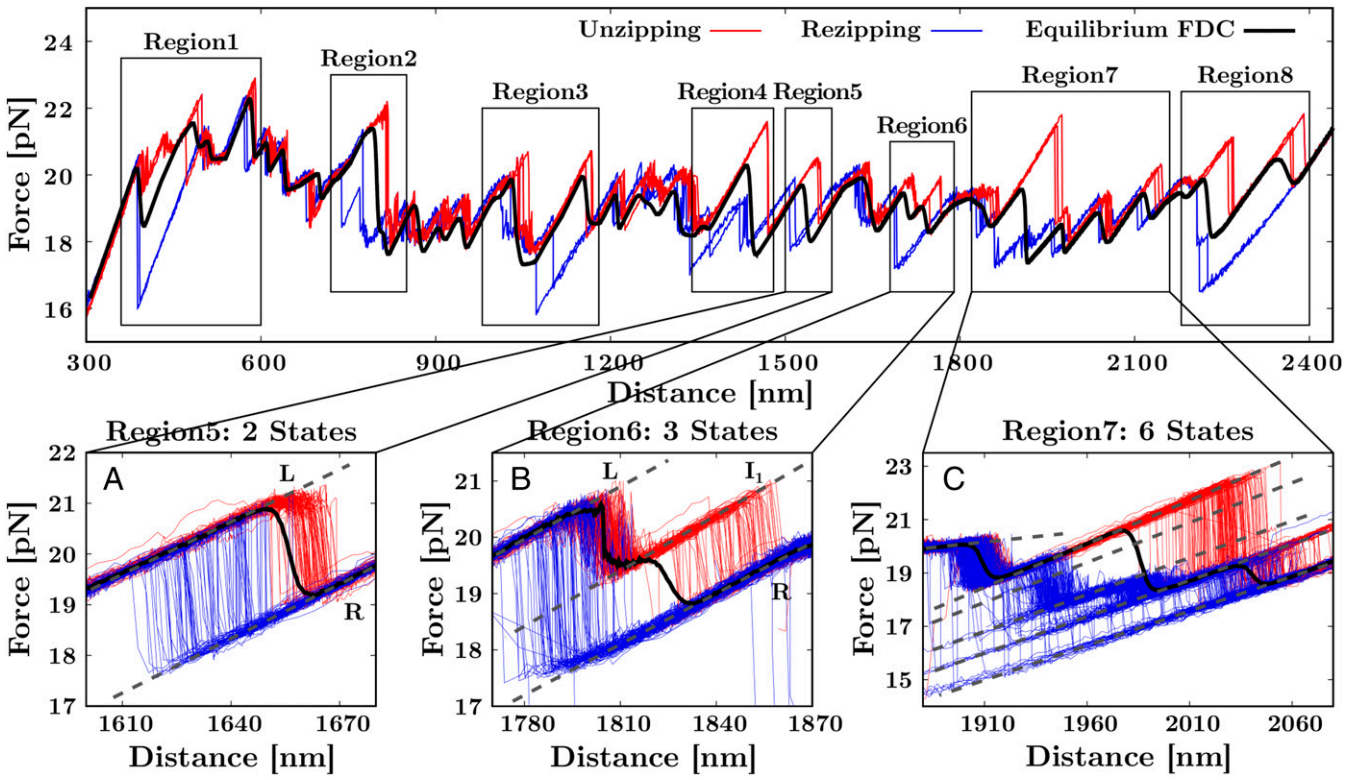

Fig. 2. Unzipping/rezipping FDCs (red/blue) in 500 mM NaCl. Black frames mark the irreversible regions. *Insets* show repeated pulling cycles in regions of increasing complexity. The intermediates (dashed gray lines) and the recovered equilibrium FDC (solid black line) are shown. *A* shows a two-states region (*L*/*R*) with no intermediates, whereas *B* and *C* report a three-states region (*L*, *R*, and the intermediate I_1) and a six-states region (*L*, *R*, and intermediates I_p with $p = 1, \ldots, 4$), respectively. The equilibrium FDC in the main box (black line) results by merging the reversible FDCs obtained for each region.

We derived the eight RNA NNBP and loop energies from the equilibrium FDCs in sodium and magnesium by using a Monte Carlo optimization algorithm, analogous to the DNA case (10, 11). The elastic parameters of the model include the persistence and contour lengths of the hybrid DNA/RNA handles ($P_{\mathrm{DNA/RNA}} = 10$ nm and $L_{\mathrm{DNA/RNA}} = 7.8$ nm) and those of the ssRNA ($P = 0.805$ nm and interphosphate distance $l_d = 0.68$ nm). The results, averaged over the different molecules, are summarized in Table 1 (columns 1 and 2) and shown in the main plot in Fig. 4*A*. The last two NNBP values (GC/CG and UA/AU) are obtained by applying the circular symmetry. These values support the validity of a salt equivalence rule between sodium and magnesium. To derive the rule we plotted the measured energies in $[Mg^{++}] = 0.01$ M as a function of the energies in $[Na^{+}] = 0.5$ M, fitting them to the relation

$$\Delta g_i^{\mathrm{Mg}}([\mathrm{Mg}^{++}]) = \Delta g_i^{\mathrm{Na}}([\mathrm{Na}^{+}]) - m \cdot \log\left(\frac{[\mathrm{Na}^{+}]_{\mathrm{eq}}}{[\mathrm{Na}^{+}]}\right), \quad \textbf{[1]}$$

where $[\mathrm{Na}^{+}]_{\mathrm{eq}} \equiv a \times [\mathrm{Mg}^{++}]$ is the magnesium concentration in sodium equivalents and a is the equivalence factor. $\Delta g_i^{\mathrm{Mg}}([\mathrm{Mg}^{++}])$ and $\Delta g_i^{\mathrm{Na}}([\mathrm{Na}^{+}])$ are the experimentally derived energies of motif i in (Mg^{++}) and (Na^{+}) at the respective salt concentrations in molar units. Finally, $m = 0.10 \pm 0.01$ kcal/mol is the NNBP-homogeneous monovalent salt correction parameter experimentally derived in ref. 19,

$$\Delta g_i^{\mathrm{Na}}([\mathrm{Na}^{+}]) = \Delta g_{0,i}^{\mathrm{Na}} - m \cdot \log([\mathrm{Na}^{+}]). \quad \textbf{[2]}$$

A least-squares fit to the data gives $a = 77 \pm 49$ (Fig. 4 *A*, *Inset*), which is compatible with the value $a \approx 100$ of previous studies (19). We expect that Eq. **1**, with a constant over a broad range of magnesium concentrations, holds if Mg^{++} correlations and competitive effects between sodium and magnesium are weak. This implies diluted magnesium solutions; i.e., $[Mg^{++}] < 0.05$ M (32, 33). With added sodium, Mg^{++} effects dominate when $R = \sqrt{[\mathrm{Mg}^{++}]}/[\mathrm{Na}^{+}] > 0.22\mathrm{M}^{-1/2}$ (34), which is the case in our experimental conditions ($R = 2\mathrm{M}^{1/2}$).

Given the measured energies (columns 1 and 2 in Table 1), we calculated the NNBP and loop values at the reference salt conditions of 1 M NaCl ($\Delta g_{0,i}^{\mathrm{Na}}$) and 1 M $MgCl_2$ ($\Delta g_{0,i}^{\mathrm{Mg}}$). By combining Eqs. **1** and **2**, we get

$$\Delta g_i^{\mathrm{Mg}}([\mathrm{Mg}^{++}]) = \Delta g_i^{\mathrm{Na}}([\mathrm{Na}^{+}]_{\mathrm{eq}}) = \Delta g_i^{\mathrm{Na}}(a \times [\mathrm{Mg}^{++}]). \quad \textbf{[3]}$$

The resulting energies in sodium and magnesium are given in columns 3 and 4 of Table 1, respectively.

For a direct comparison with the Mfold set, we use Eq. **3** to report the energies at $14\ \mathrm{mM}\ \mathrm{Mg}^{++} \equiv 1\ \mathrm{M}\ \mathrm{Na}^{+}$ (column 5 in Table 1), obtained by using Eq. **3**. Column 6 in Table 1 shows the 10 independent RNA Mfold energies plus the loop free energy. The last two NNBP values (indicated in brackets) are obtained from the circular symmetry relations applied on the other eight Mfold parameters. Note that the Mfold value for GC/CG (−3.82) is very different from our value in sodium (−3.01; Table 1, column 3). This discrepancy arises from the use of 8 parameters in our model while Mfold uses 10. Interestingly, by applying the circular symmetry property to the Mfold set we get for GC/CG the value −2.77, which is in better agreement with our value (−3.01). Note that the free energy of the loop in magnesium is not given in Table 1 as this value cannot be measured due to the inaccessibility of the last part of the unzipping curve. Results in Table 1 (columns 3, 5, and 6) are plotted in Fig. 4*B*, which shows the overall agreement between the unzipping free-energy values and those

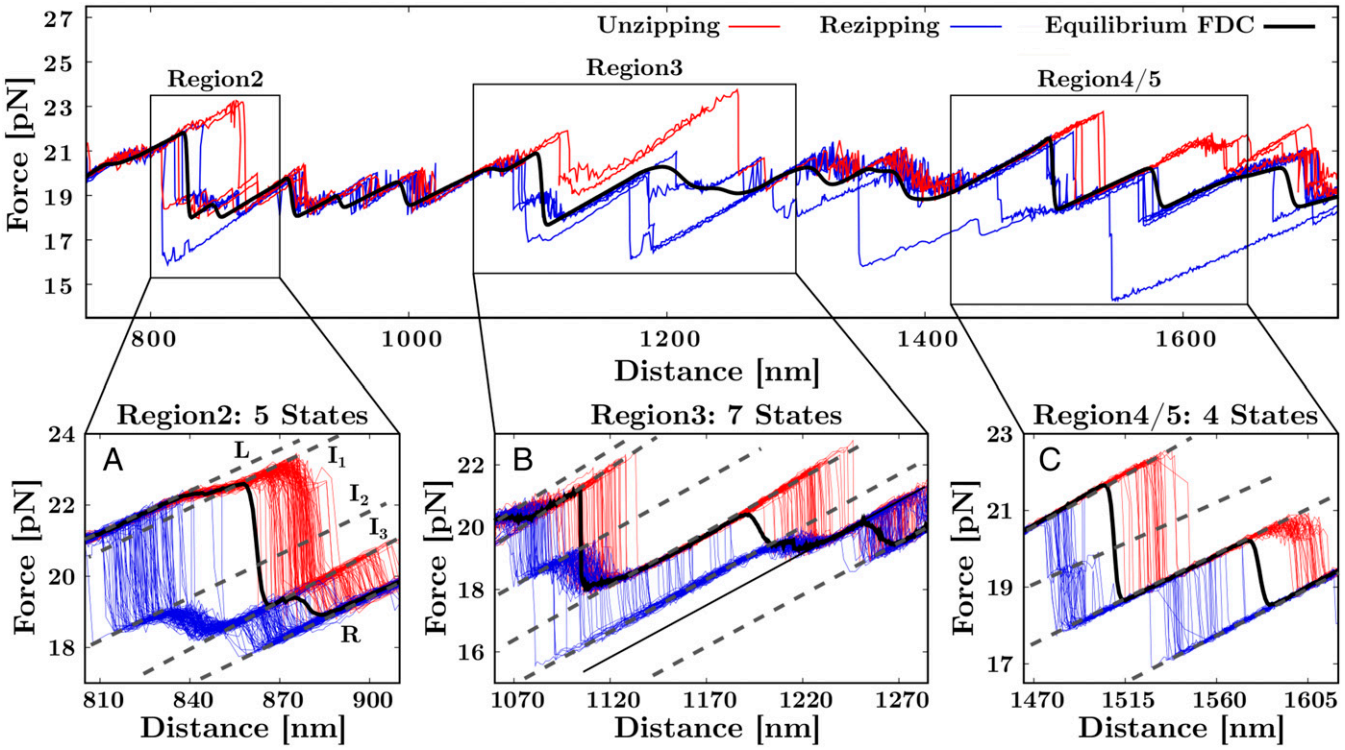

Fig. 3. Unzipping/rezipping FDCs (red/blue) in 10 mM $MgCl_2$. Black frames mark the irreversible regions along the sequence. *Insets* show repeated pulling cycles in regions of different complexity. Intermediates (dashed gray lines) and the recovered equilibrium FDC (solid black line) are shown. *A* shows a five-states region (*L*; *R*; and intermediates I_1, I_2, I_3), *B* shows a seven-states region (*L*, *R*, and intermediates I_p with $p = 1, \ldots, 5$), and *C* shows a four-states region (*L*; *R*; and intermediates I_1, I_2). The equilibrium FDC in the main box (black line) results by merging the reversible FDCs obtained for each region.

of Mfold. For the total hybridization free energy of the RNA hairpin the unzipping values predict $\Delta G_0^{\mathrm{Na}} = 4{,}031$ kcal/mol (1 M sodium) and $\Delta G_{14\ \mathrm{mM}}^{\mathrm{Mg}} = 4{,}082$ kcal/mol (14 mM of equivalent magnesium). These numbers compare well to the Mfold value $\Delta G^{\mathrm{Mfold}} = 4{,}086$ kcal/mol (1% relative error). The predicted FDCs computed with our free energies (columns 1 and 2 in Table 1) agree better with the experimental data than Mfold does, particularly for magnesium (green and orange lines versus the black line in Fig. 4 *C* and *D*). A comparison of the theoretical FDCs predicted by the Mfold set with those obtained with our energies at 1 M NaCl and 14 mM $MgCl_2$ is shown in Fig. 4*E*.

The salt rule for equivalent thermodynamics in sodium and magnesium does not necessarily imply an equivalent rule for kinetics. In ref. 19 a sequence known as CD4 hairpin was studied over three decades of monovalent and divalent salt concentrations in the diluted regime. Yet, the average unzipping force in magnesium was larger than in sodium at equivalent salt concentrations (*SI Appendix*, Fig. S10).

Stem–Loop Structures and Barrier Energy Landscape. Figs. 2 and 3 show that hysteresis is larger in magnesium than in sodium due to the longer lifetime of the intermediates in magnesium. We hypothesize a scenario where the formation of stem–loop structures

Table 1. Experimentally derived NNBP and loop RNA energies at T = 298 K

NNBP	$\Delta g^{\mathrm{Na}}_{500\ \mathrm{mM,i}}$ (1)	$\Delta g^{\mathrm{Mg}}_{10\ \mathrm{mM,i}}$ (2)	$\Delta g^{\mathrm{Na}}_{0,\mathrm{i}}$ (3)	$\Delta g^{\mathrm{Mg}}_{0,\mathrm{i}}$ (4)	$\Delta g^{\mathrm{Mg}}_{14\ \mathrm{mM,i}}$ (5)	Mfold (6)
AA/UU	−0.99(6)	−1.11(1)	−1.06(6)	−1.57(5)	−1.14(7)	−1.12
CA/GU	−1.81(6)	−2.12(1)	−1.88(6)	−2.58(5)	−2.15(7)	−2.14
GA/CU	−2.45(7)	−2.77(2)	−2.52(7)	−3.23(5)	−2.80(7)	−2.73
AU/UA	−1.20(4)	−1.06(4)	−1.27(4)	−1.52(6)	−1.09(8)	−1.09
GU/CA	−2.43(6)	−2.53(6)	−2.50(6)	−2.99(7)	−2.56(9)	−2.41
CC/GG	−3.33(4)	−3.21(4)	−3.40(4)	−3.67(6)	−3.25(8)	−3.26
CG/GC	−2.45(7)	−2.35(4)	−2.56(7)	−2.81(6)	−2.38(8)	−2.23
AG/UC	−2.16(5)	−1.96(5)	−2.23(5)	−2.42(7)	−2.00(9)	−1.93
GC/CG	−2.94(8)	−2.95(2)	−3.01(8)	−3.41(5)	−2.99(8)	−3.82[−2.77]
UA/AU	−1.03(10)	−1.26(7)	−1.10(10)	−1.72(8)	−1.29(10)	−1.36[−1.37]
Loop	0.16(3)	—	0.09(3)	—	—	0.14

Columns 1 and 2 show experimentally measured NNBP energies in 500 mM NaCl and 10 mM $MgCl_2$, respectively. The last two values (GC/CG, UA/AU) have been computed with the circular symmetry. Columns 3 and 4 show NNBP values reported at the standard conditions of 1 M NaCl and 1 M $MgCl_2$, respectively. Column 5 shows NNBP energies in magnesium reported at the concentration equivalent to 1 M $Na^+ \equiv 14$ mM Mg^{++}. Column 6 shows Mfold prediction for the 10 independent NNBP energies at 1 M NaCl. NNBP values computed with circular symmetry are also reported (square brackets). Note the loop free energy in magnesium is not given (main text). All energies are in kcal/mol and have been reported with the statistical error computed over the different molecules (in parentheses). NNBP follow the standard notation (for example, CA/GU stands for 5′-CA-3′ hybridized with 5′-UG-3′).

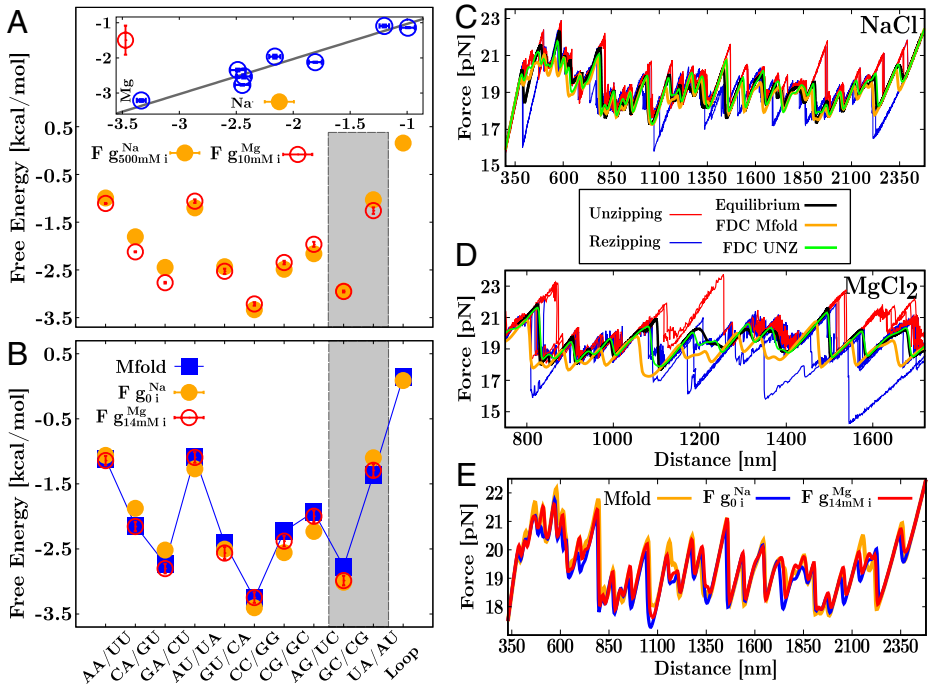

Fig. 4. NNBP free-energy parameters. (*A*) Main plot shows measured energies at 500 mM NaCl (orange) and 10 mM $MgCl_2$ (red). (*A*, *Inset*) Plot of the energies in Mg^{++} against those in Na^{+}. The fit according to Eq. **1** (gray line) gives the coefficient $a = 77 \pm 49$ (main text). (*B*) Comparison of the Mfold energies (blue) with the 1 M NaCl and (the equivalent) 14 mM $MgCl_2$ free-energy sets. The two parameters resulting from considering the circular symmetry have been highlighted (gray band). The loop free energy in magnesium has not been measured (main text). Note that in sodium the error is smaller than the size of the symbol. (*C* and *D*) Comparison of the unzipping, rezipping, and equilibrium FDCs (in red, blue, and black, respectively) measured in 500 mM NaCl and 10 mM $MgCl_2$ with the theoretical prediction obtained from Mfold (orange) and the energies reported in columns 1 and 2 of Table 1 (green). Mfold agrees better for sodium than for magnesium. (*E*) Comparison between the theoretical FDCs computed at the equivalent salt conditions $\Delta g^{\mathrm{Na}}_{0,i}$ (orange), $\Delta g^{\mathrm{Mg}}_{14\ \mathrm{mM},i}$ (red), and Mfold (blue) (columns 3, 5, and 6 of Table 1).

along the unpaired RNA strands kinetically traps the observed intermediates I_p for a given number n of formed base pairs along the FDC. The size of the force jumps observed in Figs. 2 and 3 indicates a number $\Delta n \in$ [50 to 150] of unzipped–rezipped base pairs between consecutive intermediates. The large stacking free energy of RNA loops facilitates the formation of stem–loop structures at forces as high as 20 pN where rezipping occurs (Fig. 5*A*). The stabilizing effect induced by loop formation has been demonstrated in experiments of blocking oligos in nucleic acids hairpins. By hybridizing to the complementary loop region these oligos prevent the formation of the native stem (35, 36). Stem–loops often contain hairpin-like folds with noncanonical base pairs (each colored structure in Fig. 5*A* corresponds to a different number of bases) stabilized by stacking and base-pairing interactions. To form the native stem the two single strands pulled under opposite forces must come close to each other. However, this process facilitates the formation of off-pathway (misfolded) stem–loop structures in the single strands. In Fig. 5 *B*, *Top* we depict the hairpin unzipping at position n (Fig. 5 *B*, *Top Center*) between two consecutive intermediates $I_p \to I_{p+1}$, which is slowed down by the transient formation of off-pathway (misfolded) structures (M_p; Fig. 5 *B*, *Top Left*) consisting of stem–loops in the single strands (blue segments) that kinetically trap the RNA. The intermediate I_{p+1} (Fig. 5 *B*, *Top Right*) is rescued upon releasing Δn bases forming the stem–loops (M_p). Note that kinetic trapping also occurs during rezipping for transitions $I_p \to I_{p-1}$. In the reversible regions intermediates have very short lifetimes and are not observed, meaning that kinetic trapping and hysteresis effects are negligible at the experimental pulling speeds.

The irreversibility of the unzipping–rezipping reaction can be understood by introducing a many-valley barrier energy landscape (BEL) that, for a given n, accounts for the off-pathway competing folds that can be formed in each single strand. We stress that the BEL is a nonstandard free-energy landscape describing the propensity of the hairpin to become kinetically trapped at a particular value of n by off-pathway conformations of high kinetic stability. The complexity of including all possible structures is enormous, and therefore we have restricted the analysis to the single stem–loops (loop BEL) stabilized by stacking and base pairing. Let us consider all consecutive segments of L bases along each of the two unpaired RNA strands (referred to as a and b). Let $\mathcal{S}_L^{(a,b)}$ be the set of all segments of length L contained in each strand of the RNA hairpin, $\mathcal{S}_L^{(a,b)} = \{[b_i, b_{i+L}]; 1 \le i \le N' = N - L\}$, where b_i and b_{i+L} stand for the initial and the final base of the segment on strands (a, b) (N being the total number of bases in the hairpin). For a given L-segment $[b_i, b_{i+L}]$ there are several competing folds, most of them stabilized by short complementary stems plus one or more loops of varying sizes (mostly 3 to 8 bases). We have searched for the optimal fold of lowest free energy, $\epsilon^0_{L,i}$, by using the DINAmelt web application (37, 38) based on Mfold. This yields the optimal set of energies $\{\epsilon^0_{L,i}\}^{(a,b)}$ for $\mathcal{S}_L^{(a,b)}$ at standard conditions. With the optimal set of stem–loop energies for a given L, we defined the loop BEL at force f and position n as

$$\Delta G_L(n,f) = \\ -k_B T \log \sum_{i,j=0}^{N-n} \exp\left(-\frac{\Delta g_L^{(a)}(i,f) + \Delta g_L^{(b)}(j,f)}{k_B T}\right), \quad [4]$$

where $\Delta g_L^{(a,b)}(i,f)$ is the total free-energy contribution per strand (a, b) of a stem–loop forming at distance i from the junction n at force f. Note that in Eq. **4** we assumed that all

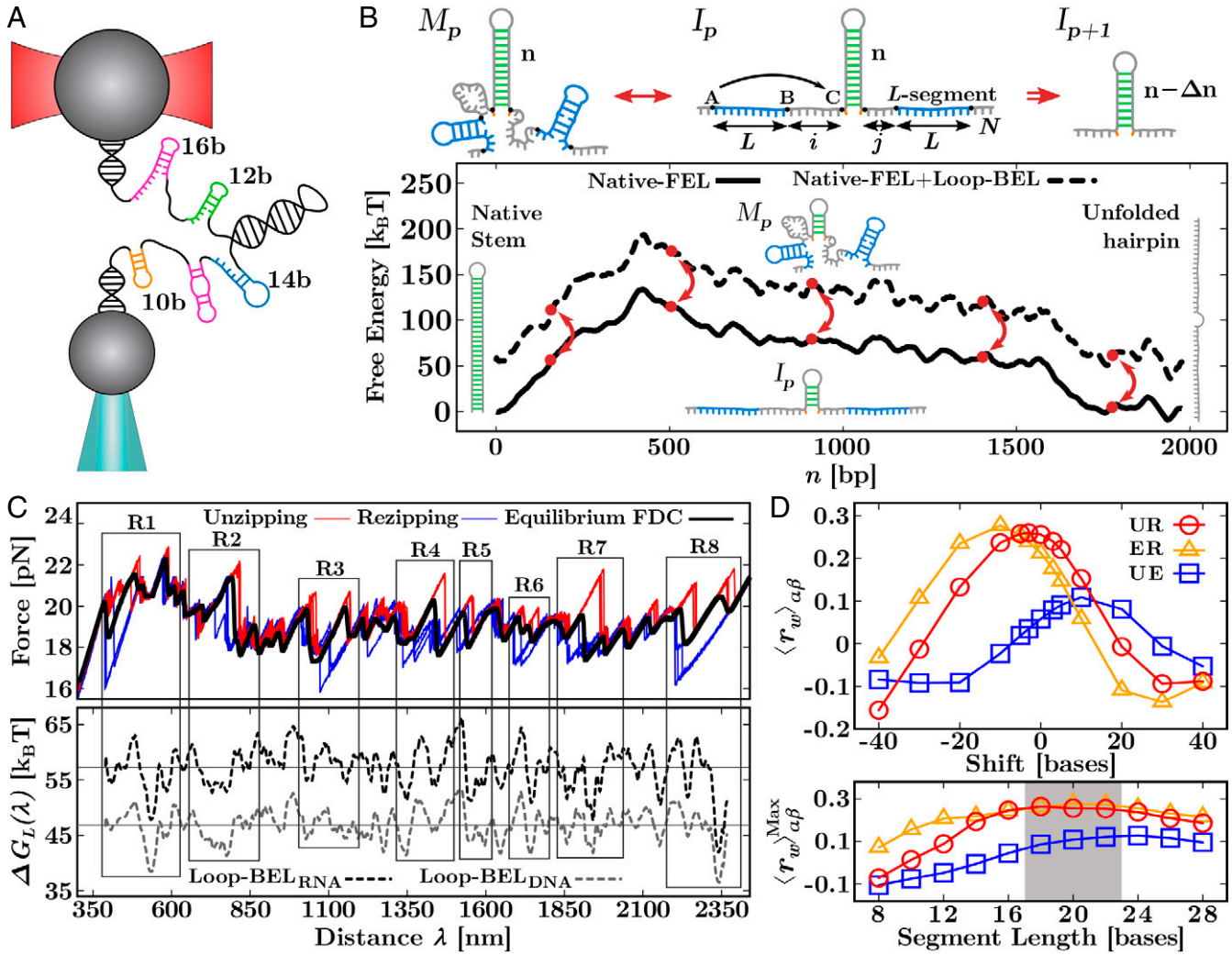

Fig. 5. Stem–loops barrier energy landscape and hysteresis. (*A*) Formation of the stem–loops during the unzipping (rezipping) process. Segments of different lengths (represented with different colors) along each RNA single strand form transient stem–loop structures. (*B, Top*) Transition between intermediates I_p and I_{p+1} (double red arrow). The formation of the off-pathway (misfolded) structures consisting of stem–loops (M_p) kinetically traps (left–right red arrow) the RNA at I_p, slowing down transitions $I_p \to I_{p+1}$ ($I_p \to I_{p-1}$) during unzipping (rezipping). (*B, Bottom*) Loop-BEL (dashed line) computed with Eq. **4** for $L = 20$ added to the native FEL of the hairpin (solid line). For a fixed n, the loop BEL is the free-energy difference between structures I_p and M_p (equal to the vertical distance between red points). Red arrows depict the kinetic trapping effect induced by the loop BEL. (*C*) Experimental FDCs in 500 mM NaCl (*Top*) and loop BELs at 19 pN (*Bottom*) computed for the RNA hairpin (dashed black line) and the equivalent DNA sequence (dashed gray line) for $L = 20$ bases. The mean values of the loop BEL (solid lines) are also shown. Loop-BEL minima correlate with the hysteresis regions R1 to R8. (*D, Top*) Average rolling correlation $\langle r_w \rangle_{\alpha\beta}$ as a function of the shift s between loop BEL and hysteresis profile for $\alpha\beta \equiv$ UR, ER, UE. (*D, Bottom*) Maximal average rolling correlation $\langle r_w \rangle^{\text{Max}}_{\alpha\beta}$ (corresponding to $s \approx 0, -10, +10$ for UR, ER, and UE, respectively) for each L. Hysteresis is found to be maximally correlated with stem–loops of length $L \sim [18, 22]$ bases.

L segments at the back of the junction are already hybridized into the native stem and do not contribute to the loop BEL (green base pairs in Fig. 5 *B*, *Top*). The term $\Delta g_L^{(a,b)}(i,f)$ is given by

$$\Delta g_L^{(a,b)}(i,f) = -\epsilon_{L,i}^{0(a,b)} + \int_0^f x_{L+i}(f')\,df', \qquad [5]$$

where $\epsilon_{L,i}^{0(a,b)}$ is the (positive) free energy of formation of the stem–loop at zero force along strand (a,b) and the integral stands for the energy cost to bring the $L+i$ bases from A to C at force f (Fig. 5 *B*, *Top*). The latter penalizes stem–loops that are formed far away from the junction because they cannot kinetically trap the stretched RNA. It has been computed with the worm-like chain (WLC) model (39)

$$f_{L+i}(x) = \frac{k_B T}{4P}\left[\left(1 - \frac{x}{(L+i)l_d}\right)^{-2} - 1 + \frac{4x}{(L+i)l_d}\right], \qquad [6]$$

with $l_d = 0.68$ nm the interphosphate distance (19, 40) and $P = 0.805$ nm the RNA persistence length (19, 41). To calculate the integral in Eq. **5** we inverted Eq. **6** (42). Note that Eq. **5** equals the free-energy difference between structures M_p and I_p in Fig. 5 *B*, *Top*.

We computed $\Delta G_L(n,f)$ at the average unzipping force $f \approx 19$ pN at 500 mM NaCl for L segments in the range $L = [8, 28]$, with $L = 8$ the minimum number of bases needed to form stem–loops. In Fig. 5 *B*, *Bottom* we show the native free energy landscape (FEL), $\Delta G_{\text{Native}}(n,f)$ (relative to the—fully unzipped—random coil state) as a black solid line. The contribution by the loop BEL for $L = 20$ has been added to the native FEL

(black dashed line) to stress the fact that it kinetically traps off-pathway stem–loop structures at fixed n (red arrows). The dashed line for the loop BEL emphasizes that this is a kinetic trapping landscape that does not describe transitions between contiguous n values. In Fig. 5 *C*, *Bottom* we show the loop BEL $\Delta G_L(\lambda)$ (dashed black line) for $L = 20$ together with the experimental FDC (Fig. 5 *C*, *Top*). The position n along the sequence in Eq. **4** has been converted to trap-pipette distance λ by using the elastic parameters, $\Delta G_L(\lambda) \equiv \Delta G_L(n, f)$. The position of the loop-BEL minima shows a correlation with the FDC regions of largest hysteresis (indicated by rectangles $R1$ to $R8$). To compare with the DNA case, we computed the loop BEL for the DNA analogous 2,027-bp sequence (obtained by replacing uracils by thymines) at the predicted average unzipping force (∼16.4 pN) at 500 mM NaCl (10). Despite that the profiles appear to be similar, the average barrier energy in DNA (∼47 k_BT, solid gray line) is lower than in RNA (∼57 k_BT, solid black line) because of the lower DNA unzipping force (which yields a lower elastic contribution in Eq. **5**). We stress that the loop BEL is overestimated as we have considered a restricted set (single stem–loops) among all possible competing structures. The lower the loop BEL is, the more stable the competing structures and the larger the irreversibility effects. The larger hysteresis in RNA apparently correlates with the higher kinetic stability of the stem–loops for RNA.

Correlation of Hysteresis with Stem–Loops Formation. To quantify the correlation between the loop BEL and the hysteresis, we introduced the hysteresis profile at position λ as a measure of the dissipated work over a given distance $\Delta\lambda$ (= 3 nm),

$$\Delta G^{\text{Hyst}}_{\alpha\beta}(\lambda) = -\int_{\lambda-\frac{\Delta\lambda}{2}}^{\lambda+\frac{\Delta\lambda}{2}} |f_\alpha(\lambda') - f_\beta(\lambda')|\, d\lambda', \qquad [7]$$

where α, β denote the unfolding (U), refolding (R), and equilibrium (E) FDCs, leading to three distinct profiles $\Delta G^{\text{Hyst}}_{\alpha\beta}(\lambda)$ with $\alpha\beta = $ UR,UE,ER. The minus sign in Eq. **7** has been introduced to positively correlate loop-BEL minima (maxima) with maximal (minimal) hysteresis. Eq. **7** has been averaged over several cycles and different molecules. Given the loop BEL, $\Delta G_L(\lambda) \equiv \Delta G_L(n, f = 19\ \text{pN})$ in Eq. **4**, and the hysteresis profile, $\Delta G^{\text{Hyst}}_{\alpha}(\lambda)$, we computed the Pearson correlation coefficient $r_w(\lambda) \in [-1, 1]$ over a given spatial window of size w as a function of λ. $r_w(\lambda) = 1$ $(r_w(\lambda) = -1)$ indicates fully correlated (anticorrelated) landscapes in that region. Correlation profiles $r_w(\lambda)$ have been calculated for $\Delta G^{\text{Hyst}}_{\alpha\beta}(\lambda)$ with $\alpha\beta \equiv \text{UR, UE, ER}$ (*SI Appendix*, Fig. S4). To assess the correlation between the loop BEL and the hysteresis profile $\alpha\beta$, we defined the average rolling correlation, $\langle r_w\rangle_{\alpha\beta}$, as the average taken over the entire landscape $r_w(\lambda)$. Another parameter for the correlation analysis is $\phi_{\alpha\beta}$, defined as the probability that $r_w(\lambda) \geq 0.5$ at a given λ averaged over the entire landscape. Although this parameter is a better estimator of positive correlations (*SI Appendix*, Fig. S5), here we show the standard average rolling correlation, $\langle r_w\rangle_{\alpha\beta}$. We used a sliding window of size $w \approx 100$ nm, the result being insensitive to w as far as it is comparable to the typical number of bases released in a force rip along the FDC (∼ 50 to 150 bases) (43) (*SI Appendix*, Fig. S6). In Fig. 5 *D*, *Top* we show $\langle r_w\rangle_{\alpha\beta}$ as a function of the shift s (in bases) of the loop BEL relative to the hysteresis profiles. $\langle r_w\rangle_{\alpha\beta}$ has been calculated for the L-segment length $L = 20$ at which correlation is maximal (see below). A positive shift $s > 0$ means that we are testing the correlation with the loop BEL in the rezipped region close to the junction (green base pair in Fig. 5 *B*, *Top*), whereas a negative shift $s < 0$ implies testing the correlation with the loop BEL ahead of the junction in the unzipped region (gray and blue base pairs in Fig. 5 *B*, *Top*). Remarkably, maximum correlation is found for $\alpha\beta \equiv$ UR and $s = 0$ (red circles in Fig. 5 *D*, *Top*) showing that stem–loops formation and hysteresis are highly correlated precisely at the junction. The position of the maximum in $\langle r_w\rangle_{\alpha\beta}$ shifts to $s > 0$ $(s < 0)$ for $\alpha\beta \equiv$ UE (ER) (blue squares and orange triangles, respectively; Fig. 5 *D*, *Top*). We note that for $\alpha\beta \equiv$ ER the maximum in $\langle r_w\rangle_{\text{ER}}$ is shifted leftward by $s \approx -10$ bases (orange triangles) and its value almost coincides with the $\alpha\beta \equiv$ UR case ($\langle r_w\rangle^{\text{Max}}_{\text{ER}} \sim \langle r_w\rangle^{\text{Max}}_{\text{UR}} \approx 0.25$, red circles). Therefore, the formation of stem–loops at a distance of ∼ 10 bases in the unzipped region slows down the refolding of the hairpin, leading to the hysteresis observed during the rezipping process. In contrast, the maximum of $\langle r_w\rangle_{\text{UE}}$ (blue squares) is shifted rightward ($s \approx +10$) with $\langle r_w\rangle^{\text{Max}}_{\text{UE}} \approx 0.1 < \langle r_w\rangle^{\text{Max}}_{\text{UR}} \approx 0.25$ (red circles). The asymmetry between UE and ER demonstrates that the largest source of irreversibility in the unzipping–rezipping experiment is the refolding process. Analogously, the rightward shift (∼ +10 bases) in $\langle r_w\rangle^{\text{Max}}_{\text{UE}}$ is related to breathing of stem–loops and the hysteresis effects observed in the unfolding FDCs. Finally, we analyzed the dependence of $\langle r_w\rangle^{\text{Max}}_{\alpha\beta}$ with the length L of the segments forming the stem–loops (Fig. 5 *D*, *Bottom*). All curves show a broad maximum for $L \approx 18$ to 22, meaning that this is the characteristic size of the stem–loops that kinetically trap the RNA intermediates during unzipping and rezipping.

These results are supported by various control analyses. In *SI Appendix*, Fig. S7 we report the average rolling correlation between loop BEL and hysteresis for different hairpins obtained by shuffling segments of the original sequence and with random sequences. This comparison shows a positive correlation for the original hairpin sequence, $\langle r_w\rangle^{\text{Max}}_{\text{UE}} \simeq 0.25$, whereas for the shuffled and random control sequences correlations are apparently lower (in the range $[-0.06{:}+0.07]$ and $[-0.04{:}-0.005]$ for the two controls in *SI Appendix*, Fig. S7 *A* and *B*, respectively). Finally, we computed the Pearson coefficient between loop BEL and hysteresis profile in the irreversible and reversible regions, separately. This analysis (*SI Appendix*, Fig. S8) shows a positive correlation only in the irreversible regions whereas in the reversible ones correlations are spurious due to thermal fluctuations and instrumental noise.

Discussion and Conclusions

Detailed knowledge of the energetics of hybridization of RNA is key to determine the thermodynamic stability of RNA structures, from dsRNA to tertiary RNAs, essential in many biophysical processes. We studied the kinetics of RNA hybridization by mechanically pulling a 2-kbp RNA hairpin with optical tweezers. By repeatedly unzipping and rezipping the RNA we measured the sequence-dependent FDCs in sodium and magnesium. The large hysteresis observed along the FDCs demanded nonequilibrium physics methods to derive the fully reversible FDC from the irreversible pulling data. In fact, quasi-static RNA unzipping experiments are not feasible as the lifetime of the intermediates requires pulling speeds that are exceedingly low. Estimates based on the Bell–Evans model range from 0.1 nm/s to 1 pm/s for irreversible hairpin segments of 30 to 40 bp.

By using an optimization algorithm, we derived the free energies of the 10 NNBPs in RNA (Fig. 4), finding good agreement with the Mfold values reported for sodium. To the best of our knowledge, NNBP energies are not currently available for RNA in magnesium. The highest difference between our energies and Mfold is found for CG/GC in sodium (Fig. 4*B*), a relevant motif prone to methylation that accumulates in many regulatory regions (44, 45). Moreover, the results for magnesium show the validity of a general salt equivalence rule 80/1 for which 10 mM Mg^{++} corresponds to 800 mM Na^+ (Fig. 4*A*). Although the scope of this result has been tested in a single-salt condition, its validity should span the dilute salt regime where cooperative salt effects are negligible ($[Mg^{++}] < 0.05$ M) and competition effects

with sodium are weak ($R = \sqrt{[Mg^{++}]}/[Na^{+}] > 0.22\,M^{-1/2}$). A salt equivalence rule has been disputed on the basis of experimental data obtained in bulk experiments using atomic emission spectroscopy in buffer-equilibrated samples (46). Although this technique is capable of determining the fraction of cations that are dissociated and bound to the RNA, it does not provide a direct measurement of free energies. Here we have demonstrated the validity of an 80/1 salt equivalence rule at the level of individual NNBP motifs. To date, this is the most direct confirmation of the validity of the 100/1 rule of thumb for the equivalence of the nonspecific binding energy of sodium and magnesium in RNA structures.

The strong hysteresis observed between the RNA unzipping and rezipping FDCs is driven by the collective effects of multiple stem–loop structures that kinetically trap the RNA. The effect is stronger in magnesium than in sodium, probably because the two charges of magnesium transiently stabilize nucleotide contacts to a higher extent. Irreversible regions are characterized by a high frequency of purine stacks and Watson–Crick bonds along the unpaired strands that lead to the multiple peaks observed in the experimental FDCs, even for forces as high as 20 pN. Note that stacking alone could not transiently stabilize stem–loops at such high forces; it is necessary for the concurrent formation of base pairs within each of the RNA strands. It is quite reasonable that such stem–loop structures also exhibit some degree of cooperativity; the more they proliferate the more they facilitate the formation of additional nearby stem–loops inhibiting native folding. Cooperative folding effects have been also found in DNA (47, 48), RNA (21, 49, 50), and proteins (51–53). The intermediates I_p in the unzipping–rezipping experiments are reminiscent of the cooperative foldons hypothesized to drive protein folding (23, 54). This cannot be otherwise, as the only way to form the native stem is to sequentially form the intermediates, one after the other, starting from the unfolded state. The remarkable effect of force is to increase the lifetime of the intermediates that would be difficult to detect in melting experiments.

We have shown that the hysteresis correlates with the transient stabilization of RNA stem–loop structures along each unpaired single strand. The formation of stem–loops in the proximity of the hybridization junction stabilizes the intermediates that enhance the hysteresis observed in the FDC. By defining a stem–loops barrier energy landscape (loop BEL, Fig. 5*B*), we found a correlation between the sequence regions where stem–loops are maximally stable (minima of the loop BEL) with those where hysteresis along the FDC is large (Fig. 5*C*). To support this interpretation we measured the correlation between the loop BEL and the hysteresis profiles (Eq. **7**) as a function of the relative shift between them (Fig. 5*D*). We have found that the hysteresis observed in the FDCs maximally correlates with the stem–loop formation at the hybridization junction. Additional test controls on shuffled and random sequences support the statistical significance of the measured correlation. Typical stem–loop sizes of about 20 bases are responsible for the observed hysteresis effects. Interestingly, this number is similar to that of foldon residues in protein folding (54). We stress that the loop BEL as a function of n is not a standard free-energy landscape as neither the trap-pipette distance λ nor n is a true reaction coordinate for the stem–loops. For a given n (λ) the loop BEL is a kinetic trapping landscape that quantifies off-pathway (misfolded) configurations M_p that compete with the folding intermediates I_p. Future work should lead to a better understanding of the stabilizing kinetics of these structures and the energy landscape describing transitions between them. We note that along the reversible regions the signal-to-noise ratio is very low due to instrumental drift and noise effects, which are detrimental in evaluating the correlation between sequence and hysteresis.

It is remarkable that hysteresis is observed in some specific regions of the FDC but not in others. To explain this, we have searched for specific sequence motifs that promote stacking, hybridization, and stem–loop formation within each single strand. We have searched for segments of length $N \geq 6$ bases containing consecutive purines for stacking (A,G) and complementary bases for hybridization (A,U and G,C) within each single strand for the irreversible and reversible regions (see *SI Appendix* for a detailed discussion). We find a higher frequency of purine stacks and hybridizing bases in the irreversible regions, showing that these regions enhance stem–loops formation and hysteresis.

Finally, we have designed a short RNA hairpin of 20 bp that ends in an A-rich dodecaloop to enhance stacking effects. The hairpin also contains many contiguous A,Us along the sequence promoting base pairing within the ssRNA. If pulled under equivalent salt conditions (100 nm/s, 1 M NaCl, and 10 mM $MgCl_2$, at 298 K) the native hairpin unzips around 21 pN. Interestingly, in magnesium the hairpin also forms an alternative misfolded structure (≈30% of the time) that is seldom observed in sodium (*SI Appendix*, Fig. S11). This result demonstrates that the presence of stacking and base pairing along the ssRNA facilitates misfolding. This effect is enhanced in magnesium, showing that kinetic effects between sodium and magnesium are nonequivalent. The same experiment but with a stem that does not contain contiguous bases capable of base pairing does not show the misfolded state either in sodium or in magnesium (*SI Appendix*, Fig. S12). Overall, these results demonstrate that concurrent stacking and hybridization among bases within the single strands lead to the observed irreversible effects.

Fluctuation relations have proved to be a fabulous playground to extract equilibrium information from irreversible pulling experiments in molecular structures from native RNAs (55, 56) to proteins (57) and ligand binding (58). Moreover, the well-defined reaction coordinate of the unzipping process shows that intermediates stabilization is induced by the formation of stem–loops along the RNA single strands. These results suggest that stem–loops formation is an essential step in RNA folding in in vitro and in vivo conditions. Indeed, numerical and theoretical studies of RNA-folding models have emphasized the importance of loop formation in the hybridization reaction (59–62). This might contribute to explaining a wide range of RNA behaviors, from misfolding (25) and multiplicity of native structures (63) to the RNA thermostatic and cold-denaturation phenomenon (64, 65). Ultimately, the promiscuity of transiently stable RNA structures might be related to the diversity of physiological responses observed when such RNAs interact with the human genome, as in the case of the RNA viruses.

Materials and Methods

Molecular Synthesis. We synthesized an RNA hairpin made of a stem of 2,027 equally represented canonical Watson–Crick base pairs (the occurrence of each NNBP motif is reported in *SI Appendix*, Table S2), ending in a tetraloop and inserted between short hybrid DNA/RNA handles (29 bp). Short handles ensure a sufficiently large signal-to-noise force and slower unzipping/rezipping kinetics (66) facilitating the detection of the intermediates occurring along the FDC. RNA constructs of a few kilobase pairs in length with a specific sequence are difficult to synthesize. In fact, the attempts to synthesize the hairpin as a single transcript from plasmids containing two copies of a DNA fragment coding for the hairpin stem failed probably due to hairpin nuclease SbcCD proteins recognizing long palindromic sequences and introducing double-stranded breaks (DSBs) on them (67). To circumvent this problem, we devised a synthesis protocol by which two RNA molecules, RNA1 and RNA2, are synthesized separately and then covalently joined using T4RNA ligase 2 (*SI Appendix*, Fig. S1). RNA1 contains a small 5′ sequence (region 1.1) that pairs with a digoxigenin-labeled DNA oligonucleotide to form a small DNA/RNA handle, a larger portion (region 1.2) that anneals with a reverse complementary strand from RNA2 molecule (region 2.2) to form the hairpin stem region, and a 3′ sequence (region 1.3) that contains the GAAA tetraloop. Apart from region 2.2, RNA2 molecule also contains a 3′ sequence used to form a second small DNA/RNA handle after annealing with a biotinylated DNA oligonucleotide. A detailed description of the synthesis is reported in *SI Appendix*.

Recovery of the Equilibrium FDC. The experimental FDCs show strong irreversibility localized in eight regions in sodium and four regions in magnesium (Figs. 2 and 3, respectively). Each region is limited by starting (left, L) and ending (right, R) trap positions where the RNA is in equilibrium and exhibits intermediates I_p being $p = (1, 2, \ldots, P)$. Let $\mathcal{S}$ be the set of $P + 2$ states containing L, R and the intermediates, $\mathcal{S} = (I_0 = L, I_1, I_2 \ldots, I_P, I_{P+1} = R)$. In our setting, during the forward process (F) the trap position λ is moved at a constant speed starting in I_0 at λ_0 and ending in I_p at λ. Similarly, in the time-reverse protocol (R) the trap position is moved back at the same velocity starting in I_p at λ and ending in I_0 at λ_0. Thus, the extended fluctuation relation reads (24, 25)

$$\frac{\phi_F^{I_0\to I_p}}{\phi_R^{I_p\to I_0}}\frac{P_F^{I_0\to I_p}(W)}{P_R^{I_p\to I_0}(-W)} = \exp\left[\frac{W-\Delta G_{I_0I_p}(\lambda)}{k_BT}\right], \qquad [8]$$

where $P_F^{I_0\to I_p}(W)$ ($P_R^{I_p\to I_0}(-W)$) is the partial distribution of the work W (defined as the work distribution conditioned to states I_0, I_p) measured along the F (R) protocol, $\Delta G_{I_0I_p}(\lambda) = G_{I_p}(\lambda) - G_{I_0}(\lambda_0)$ is the free-energy difference between states I_p at λ and I_0 at λ_0, and $\phi_F^{I_0\to I_p}$ ($\phi_R^{I_p\to I_0}$) is the fraction of paths along F (R) starting in I_0 (I_p) at λ_0 (λ) and ending in I_p (I_0) at λ (λ_0). k_B is the Boltzmann constant and T is the temperature. Note that all trajectories in the reverse process end up in I_0 whatever is the initial state I_p so that $\phi_R^{I_p\to I_0} = 1$. For a finite number of trajectories, a direct extrapolation of $\Delta G_{I_0I_p}(\lambda)$ from Eq. **8** leads to biased results. Therefore, we developed a method based on the combination of the extended Bennett acceptance ratio method (24) and the (extended) Jarzynski estimator (68) to extract the best estimate for $\Delta G_{I_0I_p}(\lambda)$ (see *SI Appendix* for the detailed description).

Given the energies of all the states occurring in a region, the equilibrium free energy is recovered as the potential of mean force taken over all the free-energy branches so that

$$\Delta G_{\rm eq}(\lambda) = -k_BT\log\left(\sum_{p=0}^{P+1}\exp-\left(\frac{\Delta G_{I_0\to I_p}(\lambda)}{k_BT}\right)\right). \qquad [9]$$

Eventually, the equilibrium FDC of the irreversible region is computed as $f_{\rm eq}(\lambda) = \partial\Delta G_{\rm eq}(\lambda)/\partial\lambda$. The recovery of the equilibrium FDC in all the irreversible regions in sodium and magnesium allowed us to reconstruct the whole equilibrium FDCs.

Data Availability. All study data are included in this article and/or *SI Appendix*.

ACKNOWLEDGMENTS. P.R. was supported by the Angelo della Riccia foundation. C.V.B. is a Research Career Awardee of the National Research Council of Brazil (CNPq) and was partly financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil, Finance Code 001. F.R. was supported by Spanish Research Council Grants FIS2016-80458-P and PID2019-111148GB-I00 and the Institució Catalana de Recerca i Estudis Avançats Academia Prizes 2013 and 2018.

1. C. E. Castro *et al.*, A primer to scaffolded DNA origami. *Nat. Methods* **8**, 221–229 (2011).
2. T. G. Cha *et al.*, A synthetic DNA motor that transports nanoparticles along carbon nanotubes. *Nat. Nanotechnol.* **9**, 39–43 (2014).
3. M. Hagiya, A. Konagaya, S. Kobayashi, H. Saito, S. Murata, Molecular robots with sensors and intelligence. *Acc. Chem. Res.* **47**, 1681–1690 (2014).
4. F. Wang, X. Liu, I. Willner, DNA switches: From principles to applications. *Angew. Chem. Int. Ed. Engl.* **54**, 1098–1129 (2015).
5. C. Jung, P. B. Allen, A. D. Ellington, A stochastic DNA walker that traverses a microparticle surface. *Nat. Nanotechnol.* **11**, 157–163 (2016).
6. H. Devoe, I. Tinoco Jr., The stability of helical polynucleotides: Base contributions. *J. Mol. Biol.* **4**, 500–517 (1962).
7. D. M. Crothers, B. H. Zimm, Theory of the melting transition of synthetic polynucleotides: Evaluation of the stacking free energy. *J. Mol. Biol.* **9**, 1–9 (1964).
8. K. J. Breslauer, R. Frank, H. Blöcker, L. A. Marky, Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3746–3750 (1986).
9. J. SantaLucia Jr., A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1460–1465 (1998).
10. J. M. Huguet *et al.*, Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15431–15436 (2010).
11. J. M. Huguet, M. Ribezzi-Crivellari, C. V. Bizarro, F. Ritort, Derivation of nearest-neighbor DNA parameters in magnesium from single molecule experiments. *Nucleic Acids Res.* **45**, 12921–12931 (2017).
12. H. H. Klump, J. Völker, K. J. Breslauer, Energy mapping of the genetic code and genomic domains: Implications for code evolution and molecular Darwinism. *Q. Rev. Biophys.* **53**, e11 (2020). Correction in: *Q. Rev. Biophys.* **53**, e14 (2020).
13. S. J. Chen, K. A. Dill, RNA folding energy landscapes. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 646–651 (2000).
14. X. Zhuang, M. Rief, Single-molecule folding. *Curr. Opin. Struct. Biol.* **13**, 88–97 (2003).
15. D. H. Mathews, J. Sabina, M. Zuker, D. H. Turner, Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940 (1999).
16. A. E. Walter *et al.*, Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9218–9222 (1994).
17. T. Xia *et al.*, Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **37**, 14719–14735 (1998).
18. S. M. Freier *et al.*, Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9373–9377 (1986).
19. C. V. Bizarro, A. Alemany, F. Ritort, Non-specific binding of Na^+ and Mg^{2+} to RNA determined by force spectroscopy methods. *Nucleic Acids Res.* **40**, 6922–6935 (2012).
20. J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins* **21**, 167–195 (1995).
21. S. A. Woodson, Compact intermediates in RNA folding. *Annu. Rev. Biophys.* **39**, 61–77 (2010).
22. D. U. Ferreiro, E. A. Komives, P. G. Wolynes, Frustration in biomolecules. *Q. Rev. Biophys.* **47**, 285–363 (2014).
23. S. W. Englander, L. Mayne, The case for defined protein folding pathways. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 8253–8258 (2017).
24. I. Junier, A. Mossa, M. Manosas, F. Ritort, Recovery of free energy branches in single molecule experiments. *Phys. Rev. Lett.* **102**, 070602 (2009).
25. A. Alemany, A. Mossa, I. Junier, F. Ritort, Experimental free-energy measurements of kinetic molecular states using fluctuation theorems. *Nat. Phys.* **8**, 688 (2012).
26. J. Camunas-Soler, A. Alemany, F. Ritort, Experimental measurement of binding energy, selectivity, and allostery using fluctuation theorems. *Science* **355**, 412–415 (2017).
27. P. Sonar *et al.*, Effects of ligand binding on the energy landscape of Acyl-CoA-binding protein. *Biophys. J.* **119**, 1821–1832 (2020).
28. M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
29. R. F. Goldstein, A. S. Benight, How many numbers are required to specify sequence-dependent properties of polynucleotides? *Biopolymers* **32**, 1679–1693 (1992).
30. P. Licinio, J. C. O. Guerra, Irreducible representation for nucleotide sequence physical properties and self-consistency of nearest-neighbor dimer sets. *Biophys. J.* **92**, 2000–2006 (2007).
31. D. M. Gray, I. Tinoco Jr., A new approach to the study of sequence-dependent properties of polynucleotides. *Biopolymers* **9**, 223–244 (1970).
32. Z. J. Tan, S. J. Chen, RNA helix stability in mixed Na^+/Mg^{2+} solution. *Biophys. J.* **92**, 3615–3632 (2007).
33. Z. J. Tan, S. J. Chen, Nucleic acid helix stability: Effects of salt concentration, cation valence and size, and chain length. *Biophys. J.* **90**, 1175–1190 (2006).
34. R. Owczarzy, B. G. Moreira, Y. You, M. A. Behlke, J. A. Walder, Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry* **47**, 5336–5353 (2008).
35. A. Bosco, J. Camunas-Soler, F. Ritort, Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions. *Nucleic Acids Res.* **42**, 2064–2074 (2014).
36. M. Manosas, X. G. Xi, D. Bensimon, V. Croquette, Active and passive mechanisms of helicases. *Nucleic Acids Res.* **38**, 5518–5526 (2010).
37. N. R. Markham, M. Zuker, DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* **33**, W577–W581 (2005).
38. N. R. Markham, M. Zuker, UNAFold: Software for nucleic acid folding and hybridization. *Methods Mol. Biol.* **453**, 3–31 (2008).
39. C. Bustamante, J. Marko, E. Siggia, S. Smith, Entropic elasticity of λ-phage DNA. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10009 (1991).
40. M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, T. Ha, Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys. J.* **86**, 2530–2537 (2004).
41. J. Camunas-Soler, M. Ribezzi-Crivellari, F. Ritort, Elastic properties of nucleic acids by single-molecule force spectroscopy. *Annu. Rev. Biophys.* **45**, 65–84 (2016).
42. A. Severino, A. M. Monge, P. Rissone, F. Ritort, Efficient methods for determining folding free energies in single-molecule pulling experiments. *J. Stat. Mech.* **2019**, 124001 (2019).
43. J. M. Huguet, N. Forns, F. Ritort, Statistical properties of metastable intermediates in DNA unzipping. *Phys. Rev. Lett.* **103**, 248106 (2009).
44. S. H. Cross, A. P. Bird, CpG islands and genes. *Curr. Opin. Genet. Dev.* **5**, 309–314 (1995).
45. M. Esteller, Epigenetics in cancer. *N. Engl. J. Med.* **358**, 1148–1159 (2008).
46. J. Lipfert, S. Doniach, R. Das, D. Herschlag, Understanding nucleic acid-ion interactions. *Annu. Rev. Biochem.* **83**, 813–841 (2014).
47. F. Schneider, N. Möritz, H. Dietz, The sequence of events during folding of a DNA origami. *Sci. Adv.* **5**, eaaw1412 (2019).
48. X. Viader-Godoy, C. Pulido, B. Ibarra, M. Manosas, F. Ritort, Cooperativity-dependent folding of single-stranded DNA. *Phys. Rev. X* **11**, 031037 (2021).
49. W. J. Greenleaf, K. L. Frieda, D. A. Foster, M. T. Woodside, S. M. Block, Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319**, 630–633 (2008).
50. M. T. J. Halma, D. B. Ritchie, T. R. Cappellano, K. Neupane, M. T. Woodside, Complex dynamics under tension in a high-efficiency frameshift stimulatory structure. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19500–19505 (2019).

51. D. B. Ritchie, M. T. Woodside, Probing the structural dynamics of proteins and nucleic acids with optical tweezers. *Curr. Opin. Struct. Biol.* **34**, 43–51 (2015).
52. E. A. Shank, C. Cecconi, J. W. Dill, S. Marqusee, C. Bustamante, The folding cooperativity of a protein is controlled by its chain topology. *Nature* **465**, 637–640 (2010).
53. A. Schug, J. N. Onuchic, From protein folding to protein function and biomolecular binding by energy landscape theory. *Curr. Opin. Pharmacol.* **10**, 709–714 (2010).
54. S. W. Englander, L. Mayne, The nature of protein folding pathways. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15873–15880 (2014).
55. D. Collin *et al.*, Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* **437**, 231–234 (2005).
56. G. Hummer, A. Szabo, Free energy profiles from single-molecule pulling experiments. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21441–21446 (2010).
57. P. T. Li, J. Vieregg, I. Tinoco Jr., How RNA unfolds and refolds. *Annu. Rev. Biochem.* **77**, 77–100 (2008).
58. X. Zhuang, Single-molecule RNA science. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 399–414 (2005).
59. C. Hyeon, D. Thirumalai, Mechanical unfolding of RNA: From hairpins to structures with internal multiloops. *Biophys. J.* **92**, 731–743 (2007).
60. T. R. Einert, P. Näger, H. Orland, R. R. Netz, Impact of loop statistics on the thermodynamics of RNA folding. *Phys. Rev. Lett.* **101**, 048103 (2008).
61. T. R. Einert, R. R. Netz, Theory for RNA folding, stretching, and melting including loops and salt. *Biophys. J.* **100**, 2745–2753 (2011).
62. T. R. Einert, H. Orland, R. R. Netz, Secondary structure formation of homopolymeric single-stranded nucleic acids including force and loop entropy: Implications for DNA hybridization. *Eur. Phys. J. E Soft Matter* **34**, 55 (2011).
63. J. Gralla, C. DeLisi, mRNA is expected to form stable secondary structures. *Nature* **248**, 330–332 (1974).
64. P. J. Mikulecky, A. L. Feig, Heat capacity changes in RNA folding: Application of perturbation theory to hammerhead ribozyme cold denaturation. *Nucleic Acids Res.* **32**, 3967–3976 (2004).
65. F. Iannelli, Y. Mamasakhlisov, R. R. Netz, Cold denaturation of RNA secondary structures with loop entropy and quenched disorder. *Phys. Rev. E* **101**, 012502 (2020).
66. N. Forns *et al.*, Improving signal/noise resolution in single-molecule experiments using molecular constructs with short handles. *Biophys. J.* **100**, 1765–1774 (2011).
67. J. K. Eykelenboom, J. K. Blackwood, E. Okely, D. R. Leach, SbcCD causes a double-strand break at a DNA palindrome in the Escherichia coli chromosome. *Mol. Cell* **29**, 644–651 (2008).
68. C. Jarzynski, Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78**, 2690 (1997).