Nucleic Acids Research, 2025, **53**, gkaf681
https://doi.org/10.1093/nar/gkaf681
RNA and RNA-protein complexes

RNA adapts its flexibility to efficiently fold and resist unfolding

Sukjin S. Jang [1,2,†], Korak Kumar Ray [1,3,†], David G. Lynall [4,†], Kenneth L. Shepard [4], Colin Nuckolls [1,*], Ruben L. Gonzalez, Jr [1,5,*]

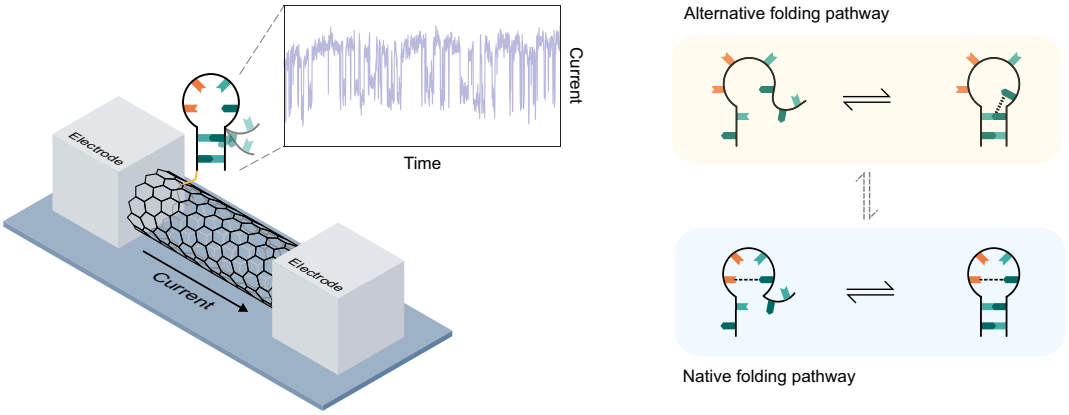
[1]Department of Chemistry, Columbia University, NY, NY 10027, United States
[2]Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, United States
[3]Present Address: MRC-Laboratory of Medical Sciences, London W12 0HS, United Kingdom
[4]Department of Electrical Engineering, Columbia University, NY, NY 10027, United States
[5]Department of Physiology and Cellular Biophysics, Columbia University Irving Medical Center, NY, NY 10032, United States

[*]To whom correspondence should be addressed. Email: rlg2118@columbia.edu
Correspondence may also be addressed to Colin Nuckolls. Email: cn37@columbia.edu
[†]The first three authors should be regarded as Joint First Authors.

Abstract

Recent studies have demonstrated that the mechanisms through which biopolymers like RNA interconvert between multiple folded structures are critical for their cellular functions. A major obstacle to elucidating these mechanisms is the lack of experimental approaches that can resolve these interconversions between functionally relevant biomolecular structures. Here, we dissect the complete set of structural rearrangements executed by an ultra-stable RNA, the UUCG stem-loop, at the single-molecule level using a nano-electronic device with microsecond time resolution. We show that the stem-loop samples at least four conformations along two folding pathways leading to two distinct folded structures, only one of which has been previously observed. By modulating its flexibility, the stem-loop can adaptively select between these pathways, enabling it to both fold rapidly and resist unfolding. This mechanism of stabilization through compensatory changes in flexibility broadens our understanding of stable RNA structures and we expect it to serve as a general strategy that can be employed by all biopolymers.

Graphical abstract



Introduction

The unique ability of biopolymers like nucleic acids and proteins to fold into and stably maintain complex three-dimensional structures is essential for their biological functions [1–5]. However, biopolymers rarely exist as a single static structure over the course of their functional lifetimes [6–8]. Instead, they form an ensemble of interconverting conformations with varying stabilities [3, 4, 9]. To understand the functions of biomolecules, it is essential to elucidate how these interconverting conformations contribute to the stability and dynamics of the ensemble as a whole.

In particular, the functional dynamics of nucleic acids, most prominently RNA, are governed by chemical properties distinct from proteins and other biomolecules. The four

Received: November 4, 2024. **Revised:** June 17, 2025. **Editorial Decision:** June 25, 2025. **Accepted:** July 8, 2025

ribonucleotide building blocks can form stable short-range interactions in a variety of combinations that collectively yield a wide range of possible RNA secondary structures [10]. Formation of a specific secondary structure further sets the stage for establishing the longer-range tertiary interactions typically responsible for the three-dimensional organization of the structured RNA [1, 3, 11]. Thus, elucidating the conformational dynamics of these foundational RNA secondary structural elements is essential for understanding the functional dynamics of larger-scale RNA structures.

One such foundational RNA secondary structural element is a stem-loop, consisting of a double-stranded stem capped by a single-stranded loop. At least two families of stem-loops, in which the loop of each family is composed of a specific subset of four-nucleotide 'tetraloop' sequences (Fig. 1), have been shown to be exceptionally stable and widely occurring [12–14]. Due to their prevalence and unanticipated stabilities, such ultra-stable stem-loops are thought to serve as nucleation sites for RNA folding, as elementary building blocks that anchor larger and more complex RNA structures, and as binding sites for RNA-binding proteins [15–18]. Despite decades of work investigating these ultra-stable RNA tetraloops [12, 13, 19–27]; however, the unique properties of these specific tetraloop sequences that allow them to confer exceptional stability to stem-loops over other loop sequences remain unknown.

Early structural work attributed their exceptional stability to the ability of these tetraloops to form unique loop structures that establish a network of non-canonical interactions. For example, the ultra-stable UUCG stem-loops (containing the UUCG tetraloop sequence) has been shown to adopt a loop structure that forms a trans-G-U wobble pair between G_{L4} and U_{L1} and a stacking interaction between U_{L1} and C_{L3} (Fig. 1A) [13, 27–29]. Consequently, this unique loop structure has suggested that UNCG tetraloops primarily serve as ultra-stable structural scaffolds and do not form interactions with other RNA motifs and proteins [15].

Recent nuclear magnetic resonance (NMR) and computational studies, however, show that stem-loops do not exclusively fold into the single static structures captured in early structural studies, but instead exist as a dynamic ensemble of varying loop conformations while retaining its stem structure [23, 24, 26, 30–32]. Moreover, ensemble fast kinetics studies directly characterizing the folding process of these tetraloops have shown that there are long-lived intermediates that significantly attribute to its stability [19–21, 33]. Finally, contrary to the popular view that UNCG tetraloops are mere structural scaffolds, a recent study analyzing existing structures of RNA stem-loops in the PDB database found that UNCG tetraloops can indeed form interactions with other RNA and protein partners. UNCG tetraloops in such contexts were found in various 'non-native' loop conformations that deviate from the loop structure identified in early structural studies [34]. Hence, the growing evidence that (i) UNCG loops remain highly dynamic even when it is stably folded and (ii) UNCG tetraloops involve significant intermediate states during its folding process suggests that the native loop structure cannot be the sole factor underlying its exceptional stability. Instead, the net stability of such an ensemble would arise collectively, from the number of populated intermediates, their individual thermodynamic stabilities, and their rates of interconversion. Unfortunately, despite the growing evidence that UNCG stem-loops exist in multiple conformations throughout its folding process, a single unifying model that explains its exceptional stability is lacking.

In this work, we have characterized the complete conformational ensemble of the UUCG stem-loop. Using a carbon nanotube-based nano-electronic device combined with a precise temperature-control setup, we have recorded the dynamics of a single UUCG stem-loop at microsecond resolution over a range of experimental temperatures (Fig. 1). We have identified the two distinct conformations the stem-loop folds into, and resolved the differences in the loop structures adopted by these conformations, that result in varying flexibilities of the folded stem-loop. Only one of these folded conformations, the more rigid one, has been previously identified as the 'native' folded conformation for this motif [13, 27, 29]. We demonstrate that the non-canonical interactions in the native loop structure stabilize this native folded conformation by increasing the rate of structure formation. On the other hand, we show that when faced with destabilizing perturbations, it is the ability of the stem-loop to sample the more flexible 'alternative' loop structure discovered here while maintaining a zipped stem that allows it to retain its overall folded structure and resist unfolding. Beyond underpinning the physical basis for the exceptional stability of the UUCG stem-loop, this mechanism of stabilization through compensatory changes in flexibility represents a fundamental shift in our understanding of RNA folding, structure, dynamics, and function. We predict this mechanism to be a general strategy employed in the folding of other structured biopolymers. Thus, this conceptual breakthrough will be crucial in the design of mutations and small-molecule ligands to both disrupt [35] and programmatically engineer [16] specific biomolecular functions.

Materials and methods

Design of the smFET experiments

The SWCNT-based smFET experimental platform, including details describing the theoretical basis for the function of an smFET device, practical aspects of how the current is measured, and how the temperature is controlled, has been previously described [36–38]. Briefly, an smFET device provides a measure of current through the SWCNT as a function of time. This current signal is extremely sensitive to the local charge environment near the surface of the SWCNT where the biomolecule is tethered. Because conformational changes, including folding and unfolding, of a biomolecule such as an RNA alters the electrostatic charge distribution near the surface of the SWCNT, the current versus time trajectory (hereafter referred to as the current trajectory) recorded by an smFET device very sensitively reports on these conformational changes.

The smFET experiments reported in this work utilized a chemically synthesized RNA stem-loop construct (Integrated DNA Technologies) that we refer to here as h-SL-bio (Supplementary Table S1). h-SL-bio consisted of a stem containing two G-C base-pairs and a loop composed of a four-nucleotide UUCG 'tetraloop' sequence. The 5′ end of h-SL-bio was modified with a hydrazide functional group that allowed us to tether the construct to the surface of the SWCNT of our smFET devices. The 3′ end of the h-SL-bio was further functionalized with a biotin for potential downstream validation of tethering with additional markers. Fortunately, tethering of h-SL-bio to the surface of SWCNT could be directly

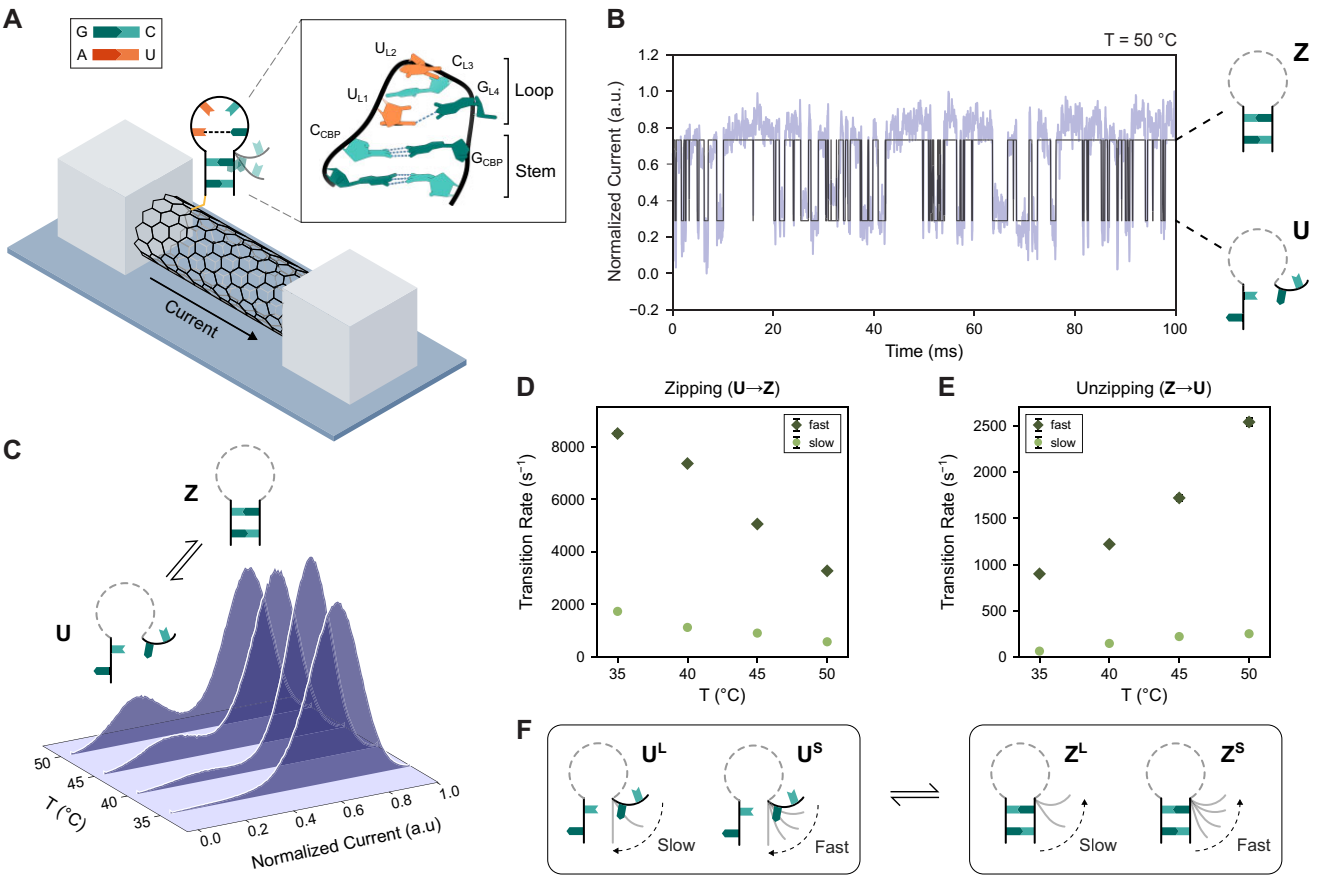


Figure 1. Following the zipping and unzipping kinetics of a single RNA stem-loop using single-molecule field effect transistor (smFET). (**A**) A schematic of the smFET experimental platform consisting of a single-walled carbon nanotube (SWCNT) connected between two electrodes, with the UUCG stem-loop tethered to the carbon nanotube surface. The inset details the three-dimensional structure of the stem-loop construct (PDB ID: 1F7Y) highlighting the stem, the loop, and the nucleotides forming the loop and the closing base-pair. (**B**) A 100 ms representative observation window of a normalized current versus time trajectory showing the real-time zipping dynamics of a single UUCG stem-loop as it transitions between **U** and **Z** conformations at 50°C. The idealized trajectory is overlaid over the raw signal. (**C**) Histograms of normalized current across an entire 60 s trajectory collected at 35°C, 40°C, 45°C, and 50°C. (**D**, **E**) The transition rates for (D) zipping and (E) unzipping of the stem-loop at varying temperatures. (**F**) A schematic of the observed slow and fast kinetic phases, and the corresponding conformations that they transition from.

validated using the current trajectories recorded by the smFET device (see below) and therefore downstream validation of tethering using the biotin functional group was unnecessary. To confirm that the 3′ biotin moiety does not alter the overall thermodynamic stability of the UUCG stem-loop, we conducted additional thermal melting experiments in which we compared the h-SL-bio construct to the same chemically synthesized RNA construct (Integrated DNA Technologies) that did not contain the biotin modification (a construct that we refer to as h-SL; Supplementary Table S1). The presence of the biotin did not perturb the thermodynamic properties of the stem-loop, demonstrating that it did not alter the folding dynamics of the stem-loop (Supplementary Fig. S1 and Supplementary Table S2).

A single molecule of the h-SL-bio construct was tethered to the SWCNT surface using an electrically coupled diazonium reaction [38, 39] (Supplementary Fig. S2A). Briefly, a source-drain current versus liquid-gate voltage (I_{sd}-V_{lg}) curve was obtained from the smFET device prior to running the diazonium reaction by applying a constant source-drain voltage (V_{sd}) and a sweeping V_{lg} from −0.5 V to +0.5 V through the platinum electrode of the smFET device in Recording Buffer [10 mM sodium phosphate (Na_2HPO_4)/sodium dihydrogen phosphate (NaH_2PO_4), pH = 7.0]. Before the introduction of 4-formylbenzene diazonium hexafluorophosphate (FBDP) to the devices, the device was placed at a V_{lg} of −0.5 V to prevent spurious diazonium reactions. The device was subsequently incubated in 300 μM FBDP dissolved in Recording Buffer to initiate the reaction. All devices were then slowly ramped towards a positive V_{lg} until a clear downward transition in current was observed, indicating the coupling of FBDP to the SWCNT surface (Supplementary Fig. S2B). Following the FBDP reaction, all devices were rinsed with 1 ml of Recording Buffer at 50°C to remove any residual FBDP salts, and a new I_{sd}-V_{lg} curve was obtained with the same conditions as before the diazonium reaction. The comparison between the I_{sd}-V_{lg} curves shows a clear reduction in conductance before and after FBDP exposure (Supplementary Fig. S2C). To tether each stem-loop construct, the hydrazide functional group at the 5′ end of the stem-loop construct was coupled to the aldehyde functional group of the SWCNT-tethered FDBP via a hydrazide-aldehyde coupling reaction. To achieve this, the device was incubated in 100 μM of the stem-loop construct in Reaction Buffer (100 mM sodium acetate (NaOAc), pH = 4.5) overnight (for ∼20 h) followed by a 2-h reduction in Reduction Buffer [10 mM $NaHPO_4$/$Na_2H_2PO_4$, pH = 7.0 and

50 mM sodium cyanoborohydride ($NaCNBH_3$)]. The device was then rinsed twice with 500 µl of Recording Buffer at 50°C to remove any non-specifically bound RNA. The smFET experimental data in this work derives from this single RNA molecule tethered to the smFET device.

The smFET experiments were performed in Recording Buffer. Current trajectories were recorded at varying temperatures (35, 40, 45, and 50°C) for 5 min at a V_{lg} of 0.1 V and a V_{sd} of 25 mV. To achieve the desired temperature, the microfluidic chamber was incubated in the heating chamber of the smFET instrument for 15 min at the specified temperature before data collection was initiated. The smFET device was sampled at a time resolution of 40 µs using a 5.3 kHz antialiasing filter.

Thermal melting experiments

To validate that our single SWCNT-tethered RNA molecule recapitulated the ensemble thermodynamic stability of the UUCG stem-loop, thermal melting experiments of the h-SL-bio construct were performed by monitoring the circular dichroism (CD) of the RNA at 280 nm as a function of temperature over a temperature window from 5°C to 95°C using a Chirascan V100 CD spectrophotometer. For this thermal melting experiment, a 10 µM sample of the h-SL-bio construct was prepared in Recording Buffer. Prior to the thermal melting experiment, the stem-loop construct was folded by pre-heating to 95°C and subsequent cooling to 5°C at a ramp rate of 2°C min^{-1}. CD measurements were recorded at every 0.2°C increment with a ramp rate of 1°C min^{-1} (Supplementary Fig. S1). Using a two-state model comprised of unimolecular transitions between single ***U*** and ***Z*** conformations [20], the equilibrium changes in enthalpy ($\Delta H^{(2)}_{eq}$) and entropy ($\Delta S^{(2)}_{eq}$) of folding for the stem-loop, where the superscript of (2) is used to denote the two-state model described above, was calculated using the following set of equations:

$$Y = (m_U T + b_U) + \left[(m_Z T + b_Z) - (m_U T + b_U)\right] \frac{K^{(2)}_{eq}}{1+K^{(2)}_{eq}},$$
$$K^{(2)}_{eq} = e^{\frac{\Delta H^{(2)}_{eq}}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right)}, \text{ and}$$
$$\Delta S^{(2)}_{eq} = \frac{\Delta H^{(2)}_{eq}}{T_m},$$

where Y is the observed CD signal, m_U and m_Z are the baseline slopes for the ***U*** and ***Z*** states, b_U and b_Z are the corresponding baseline intercepts, $K^{(2)}_{eq}$ is the two-state equilibrium constant for the zipping reaction, R is the universal gas constant (1.987 cal K^{-1} mol^{-1}), T is the absolute temperature of the measurement, and T_m is the absolute 'melting' temperature at which the corresponding free energy $\Delta G^{(2)}_{eq}$ (given by $\Delta G^{(2)}_{eq} = \Delta H^{(2)}_{eq} - T\Delta S^{(2)}_{eq}$) is zero. The above equations were fit to data collected for two independent replicates of the thermal melting experiment using a nonlinear curve fitting algorithm implemented with Matlab to yield the values of $\Delta H^{(2)}_{eq}$ and $\Delta S^{(2)}_{eq}$ for the stem-loop (Supplementary Table S2). These values were subsequently used to calculate the $\Delta G^{(2)}_{eq}$ for the UUCG stem-loop at the temperatures for the smFET experiments. The theoretical fractional occupancy of the ***U*** conformation (f_U) at each temperature was directly calculated from $K^{(2)}_{eq}$ as

$$f_U = \frac{1}{1 + K^{(2)}_{eq}}.$$

The $\Delta G^{(2)}_{eq}$ and corresponding f_U computed from the thermal melting experiments are compiled in Supplementary Table S3.

An additional set of thermal melting experiments were performed for the h-SL construct in a manner identical to the h-SL-bio construct described above (Supplementary Fig. S1). The data from two replicates of this thermal melting experiment was similarly used to estimate the $\Delta H^{(2)}_{eq}$ and $\Delta S^{(2)}_{eq}$ for the h-SL construct (Supplementary Table S2).

Identification of current states within current trajectories

The current trajectories for all of the experimental temperatures (35, 40, 45, and 50°C; Supplementary Fig. S3) were first individually background corrected using an iterative detection algorithm (as previously described in [38, 40, 41]) to correct for the slowly drifting current baseline inherent to smFET devices [37 38, 41]. The current trajectories at each temperature were subsequently divided into 200-ms time intervals, each containing 5 000 data points. Each interval was normalized such that the lowest current value for the interval corresponded to a normalized current signal of 0 a.u., and the highest current value corresponded to 1 a.u. This normalization further corrected for long-time-scale changes in the signal due to any remaining background current drift in the smFET device [41]. Intervals where excessive local background current drift and/or noise in the smFET device obscured transitions in the current trajectory resulting in an inability to clearly distinguish the two current states were discarded. In total, such intervals represented <3%–10% of the total intervals collected at each temperature point.

At all temperatures other than 35°C, the two current states present in the current trajectories could be identified and characterized using Gaussian mixture models (GMMs) inferred by employing a variational Bayesian expectation-maximization algorithm [42] implemented in the single-molecule data analysis platform, tMAVEN [43]. These GMMs yielded the means and standard deviations of the two observed states (Supplementary Fig. S3), which were subsequently used to generate idealized trajectories from the experimental data (based on the responsibilities [42] of the individual data points for each Gaussian distribution). The high-current state was centered around a normalized current value of ∼0.7 a.u. and the low-current state was centered around a normalized current value of ∼0.3 a.u. For the 35°C dataset, the relative proportion of the low-current state was so small that the signal for this state could not be reliably described as a Gaussian distribution using a GMM. In this case, a threshold of 0.4 a.u. (which corresponded to a normalized current value three standard deviations below the mean of the high-current state) was utilized to differentiate between the two states. Based on a previous smFET study of RNA stem-loops, the high-current state was assigned to zipped (***Z***) conformations of the stem-loop, and the low-current state was assigned to unzipped (***U***) conformations [41].

Calculation of two-state zipping model from smFET current trajectories

The fractional occupancies of the ***Z*** and ***U*** conformations (f_Z and f_U, respectively) were calculated from the smFET current trajectories for each experimental temperature using the GMM- and threshold-based kinetic models described above

(Supplementary Fig. S3). These fractional occupancies were subsequently used to calculate the two-state equilibrium constants, $K_{eq}^{(2)}s$, for the zipping reaction using

$$K_{\mathrm{eq}}^{(2)} = \frac{f_Z}{f_U},$$

and the corresponding two-state equilibrium free energies of zipping using

$$\Delta G_{\mathrm{eq}}^{(2)} = -RT \ln K_{\mathrm{eq}}^{(2)},$$

where R denotes the universal gas constant and T is the absolute temperature of the measurement. The $\Delta G_{eq}^{(2)}$ and corresponding f_U computed from the smFET current trajectories are compiled in Supplementary Table S3 and compared to the corresponding thermal melting experiments in Supplementary Fig. S4 (see Supporting Information for a detailed discussion of this comparison).

Inference of kinetic rates from current trajectories

Based on the idealized trajectories calculated as described above, dwell times in both $\boldsymbol{U}$ and $\boldsymbol{Z}$ conformations (defined as the contiguous time period the molecule spent in a $\boldsymbol{U}$ or $\boldsymbol{Z}$ conformation before transitioning to a $\boldsymbol{Z}$ or $\boldsymbol{U}$ conformation, respectively) were calculated. These dwell times were then used to generate the corresponding survival probability distributions, $F(t)$, defined as the fraction of dwells that were greater than a specific t, for each conformation (Supplementary Figs S5 and S6). Each survival probability distribution was fit to a single-exponential decay of the form

$$y(t) = Ae^{-kt},$$

where k represents the rate constant and A is the pre-exponential term and a double-exponential decay of the form

$$y(t) = A_1 e^{-k_1 t} + A_2 e^{-k_2 t},$$

where k_1 and k_2 represent the rate constants for the two kinetic phases, and A_1 and A_2 are the respective pre-exponential terms for the two phases that also give their relative proportions. Subsequently, residuals of the fits,

$$r(t) = F(t) - y(t),$$

were also calculated. These residuals showed that the single-exponential fits missed the initial fast decay phase of the survival probability distribution for both conformations at all temperatures (Supplementary Figs S5 and S6). Thus, the double-exponential decays better explained the survival distributions. The double-exponential decays each yielded a slow phase (k_{slow} and A_{slow}) that arises from a long-lived RNA conformation ($\boldsymbol{U}^L$ or $\boldsymbol{Z}^L$) and a fast phase (k_{fast} and A_{fast}) that arises from a short-lived RNA conformation ($\boldsymbol{U}^S$ or $\boldsymbol{Z}^S$). The kinetic parameter values obtained from the double-exponential fits at all temperatures for the zipping and unzipping transitions are compiled in Supplementary Tables S4 and S5, respectively.

Calculation of activation enthalpies and entropies using transition state theory

Each k was converted to the corresponding activation free energy ($\Delta G^{\ddagger}$) (Supplementary Fig. S7A and B) using the relation specified by transition state theory [44]

$$k = \kappa \frac{k_B T}{h} e^{-\frac{\Delta G^{\ddagger}}{RT}},$$

where T is the absolute temperature, k_B is the Boltzmann constant, h is the Planck constant, R is the universal gas constant, and κ is a correction factor that, in our calculations, is set to 1 [44]. For each kinetic phase in both the zipping and the unzipping transitions, the corresponding $\Delta G^{\ddagger}$s over the experimental range of temperatures were used to estimate the $\Delta H^{\ddagger}$s and $\Delta S^{\ddagger}$s using a linear fit to

$$\Delta G^{\ddagger}(T) = \Delta H^{\ddagger} - T\Delta S^{\ddagger}.$$

The $\Delta H^{\ddagger}$s and $\Delta S^{\ddagger}$s were assumed to be temperature-independent in the relatively small temperature range of the reported experiments, which agreed with the observed linear dependence of $\Delta G^{\ddagger}$ on temperature (Supplementary Fig. S7A and B). The calculated $\Delta G^{\ddagger}$s for both phases at all temperatures, and the fitted $\Delta H^{\ddagger}$s and $\Delta S^{\ddagger}$s, are compiled for the zipping and unzipping transitions in Supplementary Table S6.

Calculation of the equilibrium enthalpies and entropies between the kinetic phases

For both $\boldsymbol{U}$ and $\boldsymbol{Z}$ conformations, the relative proportions of the slow and fast kinetic phases were given by their respective pre-exponential terms (A_{slow} and A_{fast}) in the double-exponential fits (Supplementary Tables S4 and S5). Since these phases arose from the individual populations of the long-lived ($\boldsymbol{U}^L$ or $\boldsymbol{Z}^L$) and short-lived ($\boldsymbol{U}^S$ or $\boldsymbol{Z}^S$) conformations, the ratios between these two terms ($\frac{A_{\mathrm{slow}}}{A_{\mathrm{fast}}}$) were interpreted as the ratios between the proportions of these long-lived and short-lived RNA conformations. This allowed us to calculate the equilibrium constants (K_{eq}) between the two $\boldsymbol{U}$ conformations (and similarly, the two $\boldsymbol{Z}$ conformations) (Supplementary Fig. S7C and D) as

$$K_{\mathrm{eq}} = \frac{A_{\mathrm{slow}}}{A_{\mathrm{fast}}}.$$

Subsequently, these K_{eq}s, calculated over varying temperatures, were converted to the corresponding free energy differences, ΔG_{eq}, at these temperatures as

$$\Delta G_{\mathrm{eq}} = -\mathrm{RT} \ln K_{\mathrm{eq}},$$

where T is the absolute temperature and R is the universal gas constant. Similar to the activation parameters, these ΔG_{eq}s were fit to the linear equation

$$\Delta G_{\mathrm{eq}}(T) = \Delta H_{\mathrm{eq}} - T\Delta S_{\mathrm{eq}},$$

where ΔH_{eq} and ΔS_{eq} are the corresponding equilibrium enthalpic and entropic differences, that are also assumed to be temperature independent in the relatively small experimental range of temperatures studied here. The calculated ΔG_{eq}s at all temperatures, and the fitted ΔH_{eq}s and ΔS_{eq}s, are compiled in Supplementary Table S7.

Results

The UUCG stem-loop exists as an ensemble of two zipped and two unzipped conformations

We utilized a single-molecule field effect transistor (smFET) device to investigate the dynamics of a model stem-loop construct comprising a stem of two GC base pairs and a UUCG tetraloop [20, 22, 24, 33] (Fig. 1A and Supplementary Table S1; see the 'Material and methods' section) over a range of temperatures. Upon tethering the stem-loop to the smFET device [38, 39] (Fig. 1A and Supplementary Fig. S2; see the

'Material and methods' section), we observed a current signal that transitioned between a high-current state [normalized current of 0.7 arbitrary units (a.u.)] and a low-current state (normalized current of 0.3 a.u.) (Supplementary Fig. S3). Based on previous smFET studies [37, 41, 45, 46], we could assign the 0.7 and 0.3 a.u. states to zipped and unzipped stem-loop conformations, Z and U, respectively (Fig. 1B). Consistent with these assignments, the population of U increased as a function of temperature, reflecting the greater propensity of the stem-loop to be unzipped at higher temperatures (Fig. 1C). To confirm these assignments, we analyzed the equilibrium between these conformations using a two-state assumption (i.e. making the simplifying assumption that each state corresponded to a single conformation). This analysis accurately recapitulated the thermodynamic stability of this stem-loop as obtained from the standard two-state analysis of a conventional ensemble thermal melting experiment of the same construct (Supplementary Figs S1 and S4, and Supplementary Tables S2 and S3; see Supporting Information for a detailed description of this analysis). In addition, the populations of the two current states at each temperature were in striking agreement with the expected populations for a single stem-loop, demonstrating that the smFET device was indeed reporting on the behavior of a single, tethered stem-loop construct (Supplementary Fig. S4 and Supplementary Table S3; see Supporting Information for a detailed discussion of this comparison). Collectively, these results validated our assignment of the two current states and demonstrated that the observed zipping dynamics originated from a single stem-loop construct that had been tethered to the smFET device.

Although the two-state thermodynamic analysis described above confirmed that our smFET experiments elicited the same stem-loop stability that is obtained from ensemble experiments, the kinetics of both the zipping and unzipping transitions in the smFET experiments could not be explained by interconversions between single zipped and unzipped conformations. Specifically, the kinetics of both the zipping and unzipping transitions consisted of at least two distinct kinetic phases at all temperatures tested (Fig. 1D and E, Supplementary Figs S5 and S6, and Supplementary Tables S4 and S5). The slow and fast phases of zipping transitions from U to Z arise from long- and short-lived unzipped conformations, U^L and U^S, respectively (Fig. 1F). Similarly, the slow and fast unzipping transitions from Z to U arise from long- and short-lived zipped conformations, Z^L and Z^S, respectively (Fig. 1F). Whereas the zipping transition rates decreased with increasing temperature, the unzipping transition rates increased (Fig. 1D and E). This is a common characteristic of the folding kinetics of structured biomolecules [47–51].

The fact that these phases could be detected for both U and Z suggests that the rates of interconversion between U^L and U^S and between Z^L and Z^S must be slow compared to our 40 μs experimental acquisition time. The UUCG stem-loop thus exists as an ensemble of at least four conformations, two unzipped, U^L and U^S, and two zipped, Z^L and Z^S (Fig. 1F). However, because the normalized current values of U^L and U^S, as well as Z^L and Z^S, could not be distinguished, we were unable to identify transitions between the unzipped conformations or between the zipped conformations using our current analysis tools. This means we could not directly determine how these states were connected. For example, we could not ascertain whether U^L transitions specifically to one of the Z conformations or to both, and the same applies for U^S transitions.

The UUCG stem-loop folds through two distinct zipping pathways

We considered two potential kinetic schemes that could explain our observed dynamics, one where U^L predominately exchanges with Z^L, and U^S with Z^S (Fig. 2A), and one where U^S mainly exchanges with Z^L, and U^L with Z^S (Fig. 2B). These schemes together encompass all of the possible zipping and unzipping pathways the stem-loop may take (hereafter referred to simply as zipping pathways). To determine the kinetic scheme that was consistent with our experimental data, we used the thermodynamic parameters yielded by our temperature-dependent kinetics measurments. In each kinetic scheme, the net equilibrium change in thermodynamic parameters between two arbitrary states must be the same regardless of the path used to calculate them. For instance, the net equilibrium change in enthalpy (ΔH_{eq}) between two conformations (e.g. U^S and U^L), which represents the difference in stabilizing interactions between the two conformations, should be equal when calculated along a direct path (Path A, solid arrow in Fig. 2) or an indirect path (Path B, dashed arrows in Fig. 2). Similarly the net equilibrium change in entropy (ΔS_{eq}), which represents the difference in the flexibility of the conformations and/or the ordering of the surrounding solvent, should be equal when calculated along Paths A and B.

Our experiments allowed us to directly calculate the equilibrium change in enthalpy (ΔH_{eq}) and in entropy (ΔS_{eq}) between two of the four states (between U^L and U^S, and Z^L and Z^S) (Supplementary Fig. S7 and Supplementary Table S7; see the 'Material and methods' section for details). Our data, however, did not allow us to do the same between the U and Z states, so we needed to determine these indirectly. Using our temperature-dependent kinetics measurement, we determined the activation enthalpy ($\Delta H^{\ddagger}$) and activation entropy ($\Delta S^{\ddagger}$) that acted as barriers to the transition out of these four states (Supplementary Fig. S7 and Supplementary Table S6, see the 'Material and methods' section for details). We assumed that the zipping and unzipping transitions for our stem-loop construct do not proceed through any intermediate state. Thus, the transition state conformation ($TS^{\ddagger}$) that acts as a barrier for a zipping transition out of one of the U states corresponds to the same $TS^{\ddagger}$ that acts as a barrier for an unzipping transition out of one or the other of the Z states. By matching these $TS^{\ddagger}$s, and correspondingly, the related activation barriers, to each other (e.g. matching the transition state out of U^S to that out of Z^S in Fig. 2A or to that out of Z^L in Fig. 2B), we could calculate the net ΔH_{eq}s and ΔS_{eq}s between the U and Z states for both kinetic schemes (Fig. 2, a more detailed description of these calculations is provided in the Supporting Information).

While the direct Path A is the same in both schemes, we see that the different connectivity of the conformations in the two schemes lead to differences in the indirect Path B. Thus, if the first scheme is correct, equating the thermodynamic parameters for the two paths, we would have

$$\Delta H_{eq}(U^S \rightarrow U^L) = \Delta H_{eq}(U^S \rightarrow Z^S) + \Delta H_{eq}\left(Z^S \rightarrow Z^L\right) + \Delta H_{eq}\left(Z^L \rightarrow U^L\right), \text{ and}$$

$$\Delta S_{eq}\left(U^S \rightarrow U^L\right) = \Delta S_{eq}\left(U^S \rightarrow Z^S\right) + \Delta S_{eq}\left(Z^S \rightarrow Z^L\right) + \Delta S_{eq}\left(Z^L \rightarrow U^L\right),$$

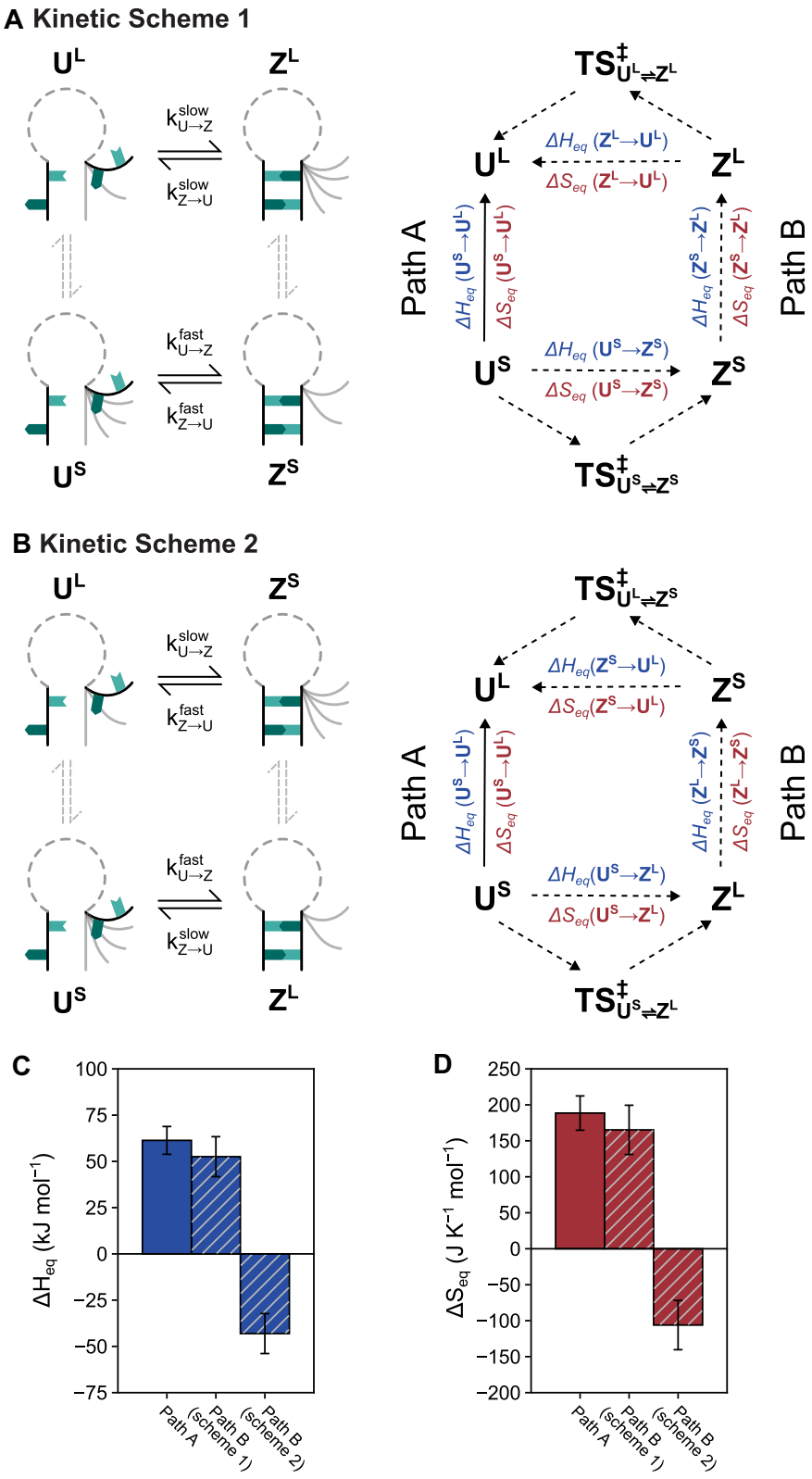


Figure 2. Two possible kinetic schemes for zipping and unzipping of the RNA stem-loop. The hypothetical kinetic schemes (left) showing predominant transitions (**A**) between $\mathbf{U^L}$ and $\mathbf{Z^L}$, and $\mathbf{U^S}$ and $\mathbf{Z^S}$, and (**B**) between $\mathbf{U^L}$ and $\mathbf{Z^S}$, and $\mathbf{U^S}$ and $\mathbf{Z^L}$, and the corresponding thermodynamic cycle depicting the direct Path A (solid arrow) and the indirect Path B (dashed arrows) between $\mathbf{U^S}$ and $\mathbf{U^L}$. The undetected transitions between the **U** conformations and **Z** conformations are shown with dashed equilibrium arrows. The bar graphs depict the calculated (**C**) $\mathbf{\Delta H_{eq}}$s and (**D**) $\mathbf{\Delta S_{eq}}$s for Path A and Path B for the two schemes in panels (A) and (B). The error bars represent the propagated errors of fit for the linear fits of free energies as a function of temperature (Supplementary Fig. S7).

where $\Delta H_{eq}(U^S \rightarrow Z^S)$, $\Delta H_{eq}(Z^L \rightarrow U^L)$, $\Delta S_{eq}(U^S \rightarrow Z^S)$, and $\Delta S_{eq}(Z^L \rightarrow U^L)$ are calculated from the activation parameters based on our placement of the corresponding $TS^{\ddagger}$s (Fig. 2A). On the other hand, if the second scheme is correct, we would similarly have

$$\Delta H_{eq}\left(U^S \rightarrow U^L\right) = \Delta H_{eq}\left(U^S \rightarrow Z^L\right) + \Delta H_{eq}\left(Z^L \rightarrow Z^S\right) + \Delta H_{eq}\left(Z^S \rightarrow U^L\right), \text{ and}$$

$$\Delta S_{eq}\left(U^S \rightarrow U^L\right) = \Delta S_{eq}\left(U^S \rightarrow Z^L\right) + \Delta S_{eq}\left(Z^L \rightarrow Z^S\right) + \Delta S_{eq}\left(Z^S \rightarrow U^L\right).$$

where $\Delta H_{eq}(U^S \rightarrow Z^L)$, $\Delta H_{eq}(Z^S \rightarrow U^L)$, $\Delta S_{eq}(U^S \rightarrow Z^L)$, and $\Delta S_{eq}(Z^S \rightarrow U^L)$ are calculated based on our placement of the corresponding $TS^{\ddagger}$s in the second scheme (Fig. 2B). Upon computing the net ΔH_{eq} and ΔS_{eq} for Path B based on the two schemes, we found that only those for the first scheme were equal to Path A (Fig. 2C and D).

UUCG stem-loop zipping is therefore characterized by two distinct kinetic pathways that exist in parallel to one another, a slow pathway consisting of transitions between U^L and Z^L, passing through $TS^{\ddagger}_{U^L \rightleftharpoons Z^L}$ (which we shall hereafter refer to as $TS^{\ddagger}_{L}$), and a fast pathway consisting of transitions between U^S and Z^S, passing through $TS^{\ddagger}_{U^S \rightleftharpoons Z^S}$ (which we shall hereafter refer to as $TS^{\ddagger}_{S}$), (Fig. 2A). In both pathways, zipping transitions occur at a faster rate than unzipping transitions. Thus, within the temperature range of our experiments, which is below the stem-loop melting temperature, T_m, of 57.0°C, the equilibria for both pathways favor the zipped conformations over the corresponding unzipped conformations. Notably, upon an increase in temperature, the slow pathway becomes progressively more populated over the fast pathway (Supplementary Tables S4 and S5), revealing that this zipping pathway of transitions between U^L and Z^L is more dominant at higher temperatures.

The two zipping pathways arise from differences in the structure of the loop

We next sought to identify what RNA structures correspond to U^L, U^S, Z^L, and Z^S. We expected our experimentally determined ΔH_{eq}s (Supplementary Fig. S7 and Supplementary Table S7) to predominantly report on the relative stabilizing interactions present in these conformations (with a negative ΔH_{eq} indicating formation of stabilizing interactions), and the ΔS_{eq}s (Supplementary Fig. S7 and Supplementary Table S7) to predominantly report on their conformational flexibilities (with a negative ΔS_{eq} indicating loss of conformational flexibility). U^L, in which the stem-loop possesses the least stabilizing interactions and is the most flexible, thus corresponds to the fully unfolded structure (Fig. 3A). Analogously, Z^S, in which the stem-loop forms the most stabilizing interactions while also being the most rigid, corresponds to the previously characterized fully folded, native stem-loop structure (Fig. 3A).

To determine the identity of Z^L, we compared the two zipped conformations. The stem-loop has less stabilizing interactions and is more flexible in Z^L than in Z^S (Fig. 3A). Since Z^L and Z^S both correspond to zipped conformations, the RNA must have similar stem structures in both; it is therefore the structure of the UUCG loop that differs between the two conformations. Having established that this loop adopts the previously characterized native structure in Z^S [27, 29], the loop in Z^L must adopt an alternatively folded structure that lacks some of the native Z^S stabilizing interactions (e.g. the *trans*-G_{L4}-U_{L1} wobble pair and/or the stacking interaction between U_{L1} and C_{L3}) and is therefore more flexible. The absence of both of these interactions easily accounts for the ∼40 kJ mol^{-1} difference in stabilizing interactions that we observe between the two Z conformations. While the UUCG loop has previously been predicted to be capable of adopting multiple non-native structures [24, 30, 32, 52], this alternatively folded loop structure has not previously been experimentally observed.

To determine the identity of U^S, we similarly compared the two unzipped conformations. U^S possesses more stabilizing interactions and is significantly more constrained than U^L (Fig. 3A). Since U^S and U^L both represent unzipped conformations, these differences must also arise from differences in the structures of the loop in U^S and U^L. Since U^S directly transitions to Z^S, we hypothesize that U^S corresponds to a partially unfolded, frayed structure in which the stem-loop is unzipped, but the loop retains its natively folded structure (Fig. 3A). Persistence of the native loop structure in U^S is in good agreement with the ∼61 kJ mol^{-1} difference in stabilizing interactions we observe between U^S and U^L. Our results therefore reveal a hierarchical mechanism of UUCG stem-loop folding in which the loop structure dictates whether the stem-loop folds along a 'native' zipping pathway from U^S to the natively folded Z^S or an 'alternative' zipping pathway from U^L to the alternatively folded Z^L.

The alternative UUCG loop structure forms base-stacking interactions with the zipped stem

We next sought to further characterize the alternative stem-loop structure we discovered by identifying stabilizing interactions in Z^L that are not present in Z^S. Unzipping along the alternative pathway resulted in a greater disruption of interactions than unzipping along the native pathway ($|\Delta H_{eq}(U^L \rightarrow Z^L)| > |\Delta H_{eq}(U^S \rightarrow Z^S)|$, Fig. 3A). This greater disruption in stabilizing interactions suggested that the stem itself is more stable and resistant to unzipping in the alternative pathway than in the native pathway. We stress that this increased stability refers to the stem alone; when considering the stabilizing interactions contributed by both the stem and the loop, the natively folded Z^S is more stable overall than the alternatively folded Z^L ($\Delta H_{eq}(Z^L \rightarrow Z^S)$ is negative, Fig. 3A). Nonetheless, this greater stability of the stem in Z^L relative to Z^S is counterintuitive based on the widely described role of the native UUCG loop structure in preferentially stabilizing the zipped stem-loop [13, 25, 27, 28].

Since stem unzipping is not expected to lead to any changes in the loop structure itself, the increased stability of the zipped stem in Z^L must arise from the alternative loop structure forming more interactions with the stem than the native loop structure. Previous computational studies guided our identification of these interactions [24]. In the native loop structure, G_{L4} in the UUCG loop (Fig. 1A) adopts the *syn*-configuration to form the *trans*-G-U wobble base pair. An anticonfiguration of G_{L4} would inhibit the formation of this *trans*-G-U wobble interaction, but instead allow G_{L4} to stack against the adjacent G_{CBP} (G in the closing base pair; Fig. 1A) [24]. We hypothesize that it is this stacking interaction in Z^L that stabilizes the zipped

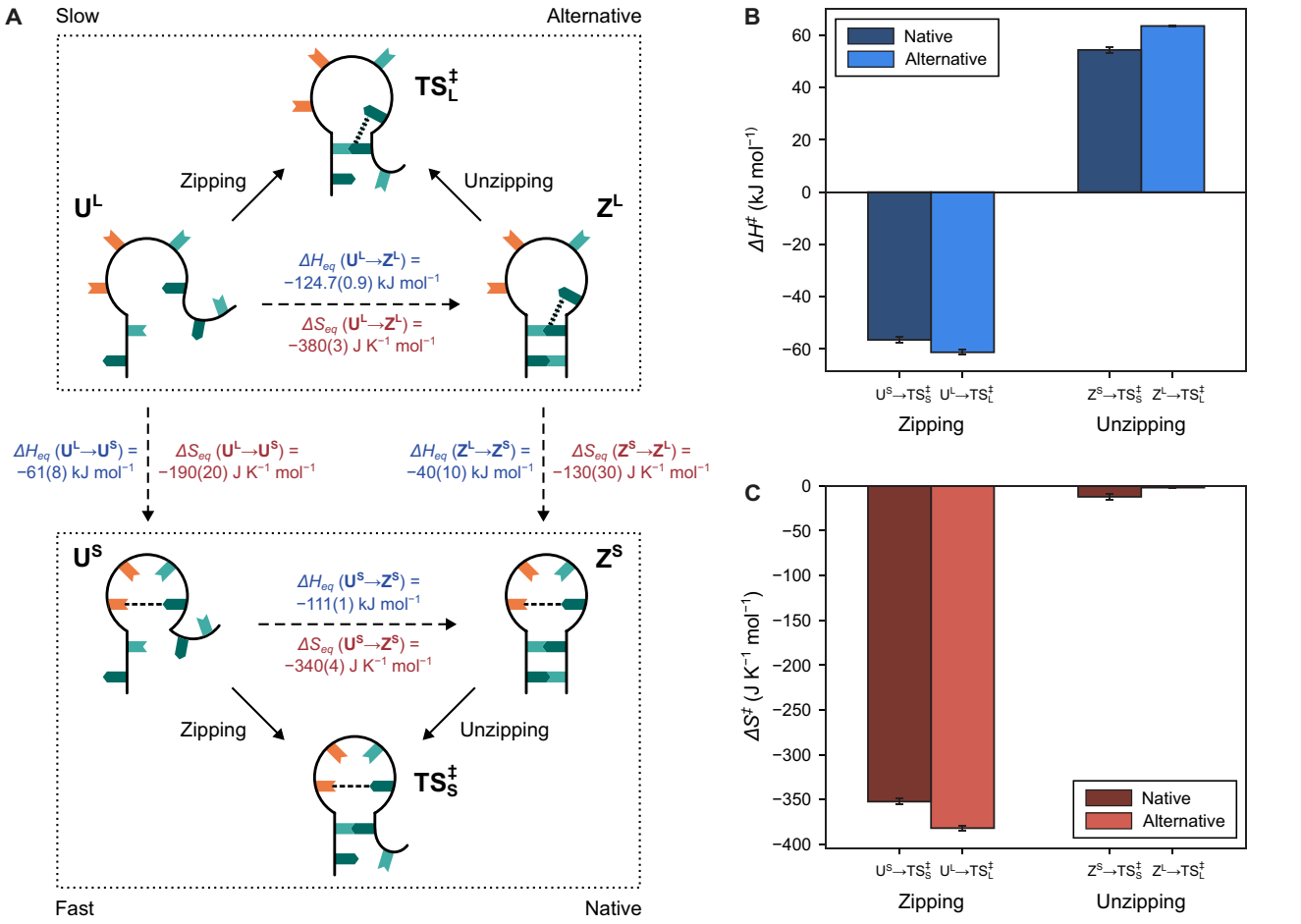


Figure 3. The structural identities of the **U** and **Z** conformations. (**A**) Schematic representations of the unzipped conformations (**U**), transition states (**TS**‡), and zipped conformations (**Z**) along the alternative (above) and native (below) zipping pathways. The calculated $\mathbf{\Delta H_{eq}}$s and $\mathbf{\Delta S_{eq}}$s between each of the observed **U** and **Z** conformations are shown. (**B**) Activation enthalpies ($\mathbf{\Delta H^{\ddagger}}$s) for the native and alternative pathways. (**C**) Activation entropies ($\mathbf{\Delta S^{\ddagger}}$s) for the native and alternative pathways. The errors (in parentheses) and error bars represent the propagated errors of fit for the linear fits of free energies as a function of temperature (Supplementary Fig. S7).

stem and needs to be overcome during unzipping in the alternative pathway (Fig. 3A). The stronger stabilizing interactions that the stem makes with the alternative loop structure versus the native loop structure, suggests that the role of the native loop structure in stem-loop folding goes beyond the formation of these interactions.

Formation of the closing base-pair determines the rate of stem zipping

To explore the role of the native loop structure in stem-loop folding, we sought to identify the conformations the stem-loop adopts in the $TS^{\ddagger}$s. Our measurements of $\Delta H^{\ddagger}$s and $\Delta S^{\ddagger}$s allowed us to directly access the stabilizing interactions and flexibilities of these transient RNA conformations [44]. For zipping transitions in both pathways, the negative $\Delta S^{\ddagger}(U^S \rightarrow TS^{\ddagger}_S)$ and $\Delta S^{\ddagger}(U^L \rightarrow TS^{\ddagger}_L)$ indicate that the $TS^{\ddagger}$s are significantly more constrained than the corresponding U conformations (Fig. 3B and C). Thus, these barriers are dominated by the challenge of constraining the stem in order to zip. In contrast, the $TS^{\ddagger}$s are almost equally as constrained as the corresponding zipped conformations [$\Delta S^{\ddagger}(Z^S \rightarrow TS^{\ddagger}_S)$ and $\Delta S^{\ddagger}(Z^L \rightarrow TS^{\ddagger}_L)$ are close to zero], but contain less stabilizing interactions [$\Delta H^{\ddagger}(Z^S \rightarrow TS^{\ddagger}_S)$ and $\Delta H^{\ddagger}(Z^L \rightarrow TS^{\ddagger}_L)$ are positive] (Fig. 3B and C). Specifically, $\Delta H^{\ddagger}(U^S \rightarrow TS^{\ddagger}_S)$ (-57 kJ mol^{-1}) and $\Delta H^{\ddagger}(U^L \rightarrow TS^{\ddagger}_L)$ (-61 kJ mol^{-1}) are almost exactly half of $\Delta H_{eq}(U^S \rightarrow Z^S)$ (-111 kJ mol^{-1}) and $\Delta H_{eq}(U^L \rightarrow Z^L)$ (-125 kJ mol^{-1}), respectively, suggesting that these $TS^{\ddagger}$s both possess half of the stabilizing interactions that are required for the stem to fully zip (Fig. 3A and B, and Supplementary Table S6). Given that zipping of the stem in our construct involves formation of only two base-pairs, this result suggests that one, and only one, of these base-pairs is already formed in the $TS^{\ddagger}$s, and it is the formation of this base-pair, associated with a significant loss in flexibility of the RNA, that determines the rate of stem zipping in both pathways.

Comparison of the two pathways revealed that formation of the first base-pairing interaction leads to greater stabilization in $TS^{\ddagger}_L$ than in $TS^{\ddagger}_S$ ($|\Delta H^{\ddagger}(U^L \rightarrow TS^{\ddagger}_L)| > |\Delta H^{\ddagger}(U^S \rightarrow TS^{\ddagger}_S)|$, Fig. 3B). We attribute this increased stabilization to the additional base-stacking interaction formed between G_{L4} and G_{CBP} in $TS^{\ddagger}_L$. Since this stacking interaction is likely to be formed only when G_{CBP} is base-paired, it is specifically the closing base pair that is formed in $TS^{\ddagger}_L$. Given the

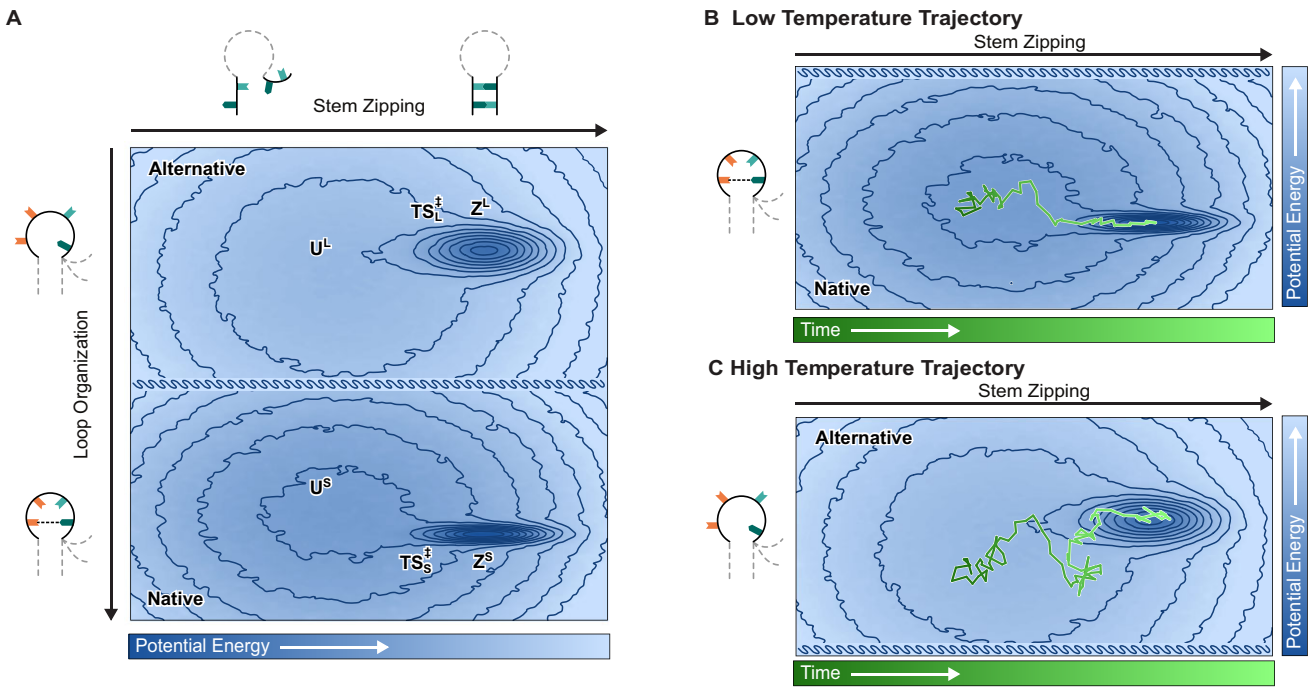


Figure 4. The physical basis of UUCG stem-loop stability. (**A**) The hypothetical hierarchical PEL for the folding of the stem-loop along the two reaction coordinates of loop organization and stem zipping. Each point in the landscape represents the potential energy of a specific stem-loop conformation projected onto these coordinates. The border between the native and alternative loop structure basin is shown as discontinuous to represent the lack of information about the barriers between the two in our analysis of the data. Two hypothetical tracks representing stem-loop folding in both basins are shown, (**B**) a short track capturing the fast guided search for Z^S in the narrow native loop structure basin at low temperatures and (**C**) a more diffusive track representing the slower search for Z^L in the wider alternative loop structure basin at high temperatures. For clarity, only the regions of the PEL containing the respective trajectories are shown in panel (B) and (C).

overall similarity between zipping in the alternative and native pathways, we expect $TS_S^‡$ to also contain just the closing base pair. The formation of the second base pair also leads to greater stabilization between $TS_L^‡$ and Z^L, than between $TS_S^‡$ and Z^S ($|\Delta H^‡(Z^L \rightarrow TS_L^‡)| > |\Delta H^‡(Z^S \rightarrow TS_S^‡)|$, Fig. 3B), as would be expected when the stacking network in $TS_L^‡$ subsequently expands to the outer base pairs in Z^L.

Our results thus reveal that it is the loss of flexibility associated with formation of the closing base-pair of the stem that serves as the barrier for zipping. On the other hand, the barrier for unzipping is almost exclusively due to the disruption of the outer base-pair (Fig. 3A), and therefore, is not associated with any significant increase in flexibility (Fig. 3C). The small differences in entropy that we observe between the Z conformations and the $TS^‡$s [$\Delta S^‡(Z^S \rightarrow TS_S^‡)$ and $\Delta S^‡(Z^L \rightarrow TS_L^‡)$] likely arise from changes in the ordering of the solvent when the outer base pair forms [53] rather than from changes in the flexibility of stem-loop itself.

UUCG stem-loop folding is a hierarchical two-step process

The results described above demonstrate that the process of UUCG stem-loop folding is two-step, using a set of hierarchical loop and stem dynamics. The UUCG loop, starting from the fully unfolded U^L, can either pre-organize into a native loop structure (U^S), in which case, stem-zipping proceeds rapidly to the Z^S conformation along the native pathway, or adopt an alternative loop structure, in which case, stem-zipping proceeds to the Z^L conformation. A powerful way to visualize the folding mechanism of the UUCG stem-loop is through a schematic of the potential energy landscape (PEL), where each point represents the potential energy of a single stem-loop microstate (Fig. 4A). Stable conformations appear as wells in PELs, where the height of the well represents the enthalpy or stability of the corresponding conformation, while its width represents the number of microstates that comprise the conformation, and thus, its entropy or conformational flexibility [54] (Supplementary Fig. S8A). In addition, for systems where the enthalpy and entropy are temperature-independent—an assumption we make for the relatively small temperature range probed in our experiments—a single PEL can capture the effect of perturbations such as changes in temperature on the system. Conformational dynamics, such as the folding of the UUCG stem-loop, can be captured as random walks on the landscape [55], with the probability of making an unfavorable (higher energy) step increasing with the temperature of the system (Fig. 4).

In our hierarchical PEL, we use two large basins to represent the ensembles of conformations in which the UUCG tetraloop adopts the native or alternative loop structure (Fig. 4A). These basins each contain one low-potential-energy well corresponding to the stable Z^S or Z^L conformations respectively. The natively folded Z^S, which has more stabilizing interactions and is more constrained than Z^L, corresponds to a deeper and narrower well. The wide and shallow regions of basins outside these wells represent the ensembles of flexible unzipped conformations the stem-loop can adopt in U^S and U^L respectively. Since the fully unfolded U^L has less stabilizing interactions and is more flexible than U^S, where the UUCG loop has already assumed a native loop structure, the corre-

sponding U^L basin is wider and shallower than the U^S one. Our existing data analysis tools cannot identify the transitions between the two basins in our experimental data (i.e. the transitions between the slow and fast phases of zipping and unzipping in our current trajectories). As a result, we could not extract any information regarding the transitions between the basins of our PEL and therefore, we represent the barrier between the two basins as a discontinuity.

Our results show that in both pathways, the barrier to zipping is purely entropic ($\Delta S^{\ddagger} < 0$) and is enthalpically favorable ($\Delta H^{\ddagger} < 0$). The corresponding $\mathbf{TS}^{\ddagger}$s are located at the entrance of the wells in each basins (Fig. 4A), demonstrating that even though the pathway of folding is energetically downhill (see Supplementary Fig. S8B for a one-dimensional representation of the PEL), the barrier to zipping comes from the significant loss of flexibility associated with the transition from the wide basins to the narrow wells. Thus, while our PEL is mostly illustrative, it is a powerful tool to capture and explain even subtle mechanistic details of the hierarchical two-step folding pathways of the UUCG stem-loop.

The native loop structure facilitates zipping by increasing the proximity between the closing bases

Using the above constructed PEL, we can now explore the mechanistic roles played by the RNA conformations we have observed in the folding of stem-loop. We see that the process of stem-loop folding from the fully unfolded U^L to natively folded Z^S can be represented on the PEL as the conformational search from the shallow and wide non-native loop basin into the deep and narrow Z^S well. This process involves a significant loss of flexibility and is thus entropically unfavorable. Our results show that the existence of the metastable intermediate, U^S, facilitates this folding process in two ways.

Firstly, a folding pathway comprising hierarchical loop and stem dynamics allows the entropic cost of stem-loop folding to be paid over two steps. The first part of the entropic cost of folding can be paid by nucleating the process using the pre-organization of the loop into the native structure in U^S, shifting the stem-loop to the native folding pathway. This constrains the conformational search for the native folded conformation within the deeper and, more importantly, narrower U^S basin, facilitating transitions to the narrow Z^S well (as shown by the track in Fig. 4B). By further coupling the stepwise loss of flexibility from U^L to Z^S with a gain in stabilizing interactions at each step (Fig. 4 and Supplementary Fig. S8B), this hierarchical separation of loop organization and stem zipping results in a guided conformational search for the natively folded stable but rigid stem-loop structure.

Secondly, while any structured intermediate would support the above guided conformational search to the native Z^S well, our results show how the native loop structure in U^S is particularly optimized to facilitate stem-zipping. We show above that the barrier to stem-zipping is the formation of the closing base-pair (Fig. 3A). Thus, RNA structures that increase the proximity between the two closing bases particularly facilitate zipping by reducing the change in flexibility that is needed to form the stem. This is the case for U^S, where formation of the first base-pair is associated with a smaller loss of flexibility than for the fully unfolded U^L [$|\Delta S^{\ddagger}(U^S \rightarrow TS_S^{\ddagger})| < |\Delta S^{\ddagger}(U^L \rightarrow TS_L^{\ddagger})|$, Fig. 3C]. The pre-organization of the native loop structure in U^S thus leads to a faster rate of zipping from U^S (Fig. 1D). Counterintuitively, in the alternative pathway, the stacking interaction between G_{L4} and G_{CBP} leads to a further loss of flexibility that needs to be overcome to form the stem, contributing to a more negative $\Delta S^{\ddagger}$, resulting in a slower rate of zipping despite the greater stabilization that results from forming this interaction. Taken together, this describes the mechanism through which pre-organization of the tetraloop significantly in U^S reduces the conformational space of the stem-loop and expedites the folding process.

The alternative zipped conformation resists unfolding through compensatory changes in flexibility

In addition to the guided-search mechanism of folding along the native pathway described above, we see that the alternative folding pathway also contributes significantly to the observed stability of the UUCG stem-loop. This is specifically observed under unfavorable conditions, such as at higher temperatures, where the alternative zipping pathway becomes progressively more dominant (Supplementary Tables S4 and S5). This effect can also be visualized using the PEL for UUCG stem-loop folding (Fig. 4).

When faced with destabilizing perturbations like higher temperatures, the guided conformational search of the native pathway cannot function as efficiently as in favorable conditions. The loss of flexibility required to access the narrow Z^S well is too unfavorable at higher temperatures, as seen by the greatly decreased zipping rate in the fast native pathway (Fig. 1D). Furthermore, at higher temperatures, given the rigidity of the Z^S native conformation, the increased thermal fluctuations of the stem-loop cannot be accommodated while remaining folded within the narrow Z^S well. Consequently, these thermal fluctuations result in a greater probability for transitions out of the Z^S well, overcoming the primarily enthalpic barrier to unzipping in the native zipped stem conformation (Fig. 3) and leading to an increased rate of unzipping of Z^S in the fast native pathway (Fig. 1D). Thus, increased thermal fluctuations both cause an unfolding of the native stem-loop structure and prevent its refolding into the narrow Z^S well.

Importantly, the Z^L well is wider than the Z^S well. Thus, the conformational search for this well from the fully unfolded U^L is easier than for the narrower Z^S well. This can be seen in the PEL by the more diffusive track in the alternative pathway basin—representing more thermal fluctuations of the RNA—still being able to find and zip into the wider Z^L well (Fig. 4C), albeit at a slower rate than the native pathway (Fig. 1D). Secondly, the wider Z^L well, reflecting the larger variation of more flexible loop structures that can be adopted in Z^L while maintaining a zipped stem, enables the stem-loop to accommodate increased thermal fluctuations remaining folded at higher temperatures. Furthermore, due to the additional base stacking interaction between the stem and the loop, Z^L possesses a higher barrier to exit than Z^S (Supplementary Fig. S8B). Thus, once in the Z^L well, the alternatively folded stem-loop can better resist unfolding in unfavorable conditions, as is seen by the slower rate of unzipping in the alternative pathway compared to the native pathway (Fig. 1D). Thus, shifting to the alternative pathway allows the stem-loop to both fold into the more flexible Z^L conformation and to resist unfolding at higher temperatures.

Z^L is clearly a kinetic trap in the folding process of the native UUCG stem-loop structure. However, the crucial role of this conformation in the stability of the folded stem-loop

structure as a whole leads us to suggest that Z^L does not represent merely a misfolded conformation, but is instead a previously unidentified, functionally relevant alternative structure of this RNA. While beyond the scope of this work, additional functional roles for this alternative UUCG loop structure need to be further explored. We note, for instance, that the alternative loop structure in Z^L (see above, Fig. 3A) may position the UUCG stem-loop to make tertiary contacts with other RNAs or proteins by freeing the bases in the loop. This conformation may thus have unexplored biological roles beyond stabilizing the stem-loop by resisting unfolding.

Discussion

We have elucidated the complete conformational ensemble of a fundamental RNA structural motif, the UUCG stem-loop, using a high-time-resolution, single-molecule experimental approach uniquely suited for studying the dynamics of such biopolymers. We characterize, in molecular detail, the competing pathways, metastable intermediates, and kinetic traps that collaborate to give rise to the exceptional thermodynamic stability of this RNA stem-loop. Our findings reveal that the UUCG stem-loop uses a hierarchical conformational search process to guide folding into the fully folded native structure. We further show how the specific ordering of this hierarchy into the two steps of loop structure organization and stem zipping facilitates the efficient collapse of the flexible RNA into the constrained native folded state (Fig. 4A) in the native zipping pathway—an entirely entropic effect. We further discover how this RNA stem-loop resists unfolding when challenged with destabilizing perturbations by shifting its equilibrium to a more flexible alternatively folded conformation (Fig. 4A), a conformation that may also have additional, hitherto uncharacterized biological roles. Our results elucidate how these dual mechanisms of constrained conformational search and compensatory changes to flexibility collaborate to give rise to the exceptional thermodynamic stability of the UUCG stem-loop.

The folding mechanism we have elucidated here has clear implications for our views of the biological roles of UNCG stem-loops (where N denotes any nucleotide). UNCG stem-loops, including UUCG stem-loops, have typically been considered to only serve as ultra-stable structural scaffolds [13, 15, 25, 27, 28]. This perspective has predominately been informed by static structures of the fully folded, native structure of the UUCG stem-loop which can only form a limited set of tertiary interactions [13, 25, 27, 28]. In contrast, our results show that when faced with perturbations that denature the stem-loop, it favors a more flexible alternative loop structure.

This alternative loop structure has less intramolecular interactions in the loop and we expect it to enable more tertiary interactions. Consistent with this view, recent study analyzing existing structures of RNA stem-loops in the PDB database found that UNCG stem-loops often exist in 'non-native' conformations when interacting with other RNAs and proteins [34]. Furthermore, studies combining NMR and molecular dynamics (MD) methods have shown that even in datasets in which the fully folded native structure of the UUCG was determined, the loop is highly dynamic and can populate minor alternative loop conformations [30, 32]. These alternative loop conformations mostly retain the syn-configuration and are likely to be distinct from the alternative loop we observe during the zipping and unzipping of the stem-loop. Nonetheless, the fact that UUCG stem-loops can adopt alternative loop conformations while it is stably folded highlights that alternative loop structures are prevalent in UUCG stem-loops and serve an important role in defining their stability. Altogether, these studies point to a unifying view that indeed UNCG loops exist in diverse alternative loop conformations that are more viable for tertiary interactions with other biomolecular partners.

The ability of the UUCG stem-loop to shift between the native and alternative loop structures in response to perturbations sets up an enticing model for the regulation of stem-loop interactions with other biomolecules. In this model, perturbations to the UUCG stem-loop, whether through changing conditions or the allosteric effect of the binding of ligands or other biomolecules, can shift the equilibrium between the two loop structures, tuning the ability of the stem-loop to interact with partners. This, of course, is a hypothesis which needs to be further investigated in light of our findings.

Finally, the complete folding mechanism of the UUCG stem-loop that we reveal here (Fig. 4) is in striking contrast to our current understanding, which attributes the exceptional thermodynamic stability of this RNA motif solely to the interactions within the unique native UUCG loop structure [13, 27–29]. While beyond the scope of this work, it will be worthwhile to determine how this folding mechanism maps onto other thermodynamically stable RNA motifs, including GNRA stem-loops (where N denotes any nucleotide and R any purine nucleotide), which are thought to have evolved a strategy for achieving their exceptional thermodynamic stability that is independent and distinct from that of the UNCG stem-loops [56]. It is worth noting, in this context, that the strategies for thermodynamic stability we describe here do not appear to be specific to stem-loops—or indeed, to RNA—and thus, we expect to see them in the folding mechanisms of a wide range of biomolecules. In particular, the strategy we have discovered here of resisting destabilizing perturbations through compensatory changes in the flexibility of the biomolecule represents an exciting addition to our understanding of biomolecular stability, the implications of which must be explored for broader classes of structured biopolymers.

Acknowledgements

We thank Dr Hashim al-Hashimi (Columbia University) and Dr Jeffrey Kieft (New York Structural Biology Center) for valuable comments on the manuscript. We also thank Dr Sarah Dubnik and other past and present members of the Gonzalez, Nuckolls, and Shephard labs for their assistance, comments, and advice. This work was carried out in part in the Clean Room, Electron Microscopy, and Shared Materials Characterization labs of Columbia Nano Initiative (CNI) Shared Lab Facilities and the Precision Biomolecular Characterization Facility (PBCF) at Columbia University.

Author contributions: Methodology: S.S.J., K.K.R., D.G.L., K.L.S., C.N., R.L.G.; Investigation: S.S.J., K.K.R., D.G.L., K.L.S., C.N., R.L.G.; Visualization: S.S.J., K.K.R., C.N., R.L.G.; Funding acquisition: C.N., R.L.G.; Project administration: K.L.S., C.N., R.L.G.; Supervision: K.L.S., C.N., R.L.G.; Writing – original draft: S.S.J., K.K.R., C.N.,

R.L.G.; Writing – review & editing: S.S.J., K.K.R., D.G.L., K.L.S., C.N., R.L.G.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

D.G.L. and K.L.S. have a financial interest in Quicksilver Biosciences, Inc., which is commercializing smFET technology for molecular diagnostic applications. The other authors declare no competing financial interest.

Funding

C.N. thanks Sheldon and Dorothea Buckler for their generous support. This research was supported by the National Science Foundation (CHE 2004016) and the National Institutes of Health (GM107417 and GM153724). Essential instrumentation in the PBCF was made possible by funding from the National Institutes of Health (S10OD025102).

Data availability

The open-source Python code for tMAVEN, which was used for the single-molecule data analysis in this work, is freely available *via* a Git repository (https://github.com/GonzalezBiophysicsLab/ tmaven). The current trajectories are also available on Zenodo at https://zenodo.org/records/14036502.

References

1. Tinoco I, Bustamante C. How RNA folds. *Science* 1999;**293**:191–4. https://doi.org/10.1006/jmbi.1999.3001
2. Dill KA, Ozkan SB, Shell MS *et al.* The protein folding problem. *Annu Rev Biophys* 2008;**37**:289–316. https://doi.org/10.1146/annurev.biophys.37.092707.153558
3. Mustoe AM, Brooks CL, Al-Hashimi HM. Hierarchy of RNA functional dynamics. *Annu Rev Biochem* 2014;**83**:441–66. https://doi.org/10.1146/annurev-biochem-060713-035524
4. Ganser LR, Kelly ML, Herschlag D *et al.* The roles of structural dynamics in the cellular functions of RNAs. *Nat Rev Mol Cell Biol* 2019;**20**:474–89. https://doi.org/10.1038/s41580-019-0136-0
5. Herschlag D, Bonilla S, Bisaria N. The story of RNA folding, as told in epochs. *Cold Spring Harb Perspect Biol* 2018;**10**:a032433.
6. Frauenfelder H, Sligar SG, Wolynes PG. The energy landscapes and motions of proteins. *Science* 1991;**254**:1598–603. https://doi.org/10.1126/science.1749933
7. Vicens Q, Kieft JS. Thoughts on how to think (and talk) about RNA structure. *Proc Natl Acad Sci USA* 2022;**119**:e2112677119. https://doi.org/10.1073/pnas.2112677119
8. Fraser JS, Murcko MA. Structure is beauty, but not always truth. *Cell* 2024;**187**:517–20. https://doi.org/10.1016/j.cell.2024.01.003
9. Wei G, Xi W, Nussinov R *et al.* Protein ensembles: how does nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell. *Chem Rev* 2016;**116**:6516–51. https://doi.org/10.1021/acs.chemrev.5b00562
10. Bloomfield VA, Crothers DM, Tinoco I. *Nucleic Acids: Structure, Properties, and Functions*. Sausalito: University Science Books, 2000.
11. Brion P, Westhof E. Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct* 1997;**26**:113–37. https://doi.org/10.1146/annurev.biophys.26.1.113
12. Antao VP, Lai SY, Tinoco I. A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res* 1991;**19**:5901–5. https://doi.org/10.1093/nar/19.21.5901
13. Varani G, Cheong C, Tinoco I Jr. Structure of an unusually stable RNA hairpin. *Biochemistry* 1991;**30**:3280–9. https://doi.org/10.1021/bi00227a016
14. Woese CR, Winker S, Gutell RR. Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". *Proc Natl Acad Sci USA* 1990;**87**:8467–71. https://doi.org/10.1073/pnas.87.21.8467
15. Hall KB. Mighty tiny. *RNA* 2015;**21**:630–1. https://doi.org/10.1261/rna.050567.115
16. Yesselman JD, Eiler D, Carlson ED *et al.* Computational design of three-dimensional RNA structure and function. *Nat Nanotechnol* 2019;**14**:866–73. https://doi.org/10.1038/s41565-019-0517-8
17. Bevilacqua PC, Blose JM. Structures, kinetics, thermodynamics, and biological functions of RNA hairpins. *Annu Rev Phys Chem* 2008;**59**:79–103. https://doi.org/10.1146/annurev.physchem.59.032607.093743
18. Yang X, Gerczei T, Glover LT *et al.* Crystal structures of restrictocin-inhibitor complexes with implications for RNA recognition and base flipping. *Nat Struct Biol* 2001;**8**:968–73. https://doi.org/10.1038/nsb1101-968
19. Sarkar K, Nguyen DA, Gruebele M. Loop and stem dynamics during RNA hairpin folding and unfolding. *RNA* 2010;**16**:2427–34. https://doi.org/10.1261/rna.2253310
20. Proctor DJ, Ma H, Kierzek E *et al.* Folding thermodynamics and kinetics of YNMG RNA hairpins: specific incorporation of 8-bromoguanosine leads to stabilization by enhancement of the folding rate. *Biochemistry* 2004;**43**:14004–14. https://doi.org/10.1021/bi048213e
21. Sarkar K, Meister K, Sethi A *et al.* Fast folding of an RNA tetraloop on a rugged energy landscape detected by a stacking-sensitive probe. *Biophys J* 2009;**97**:1418–27. https://doi.org/10.1016/j.bpj.2009.06.035
22. Mrazikova K, Mlynsky V, Kuhrova P *et al.* UUCG RNA tetraloop as a formidable force-field challenge for MD simulations. *J Chem Theory Comput* 2020;**16**:7601–17. https://doi.org/10.1021/acs.jctc.0c00801
23. Sorin EJ, Engelhardt MA, Herschlag D *et al.* RNA simulations: probing hairpin unfolding and the dynamics of a GNRA tetraloop. *J Mol Biol* 2002;**317**:493–506. https://doi.org/10.1006/jmbi.2002.5447
24. Chen AA, García AE. High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations. *Proc Natl Acad Sci USA* 2013;**110**:16820–5. https://doi.org/10.1073/pnas.1309392110
25. Nozinovic S, Fürtig B, Jonker HRA *et al.* High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Res* 2010;**38**:683–94. https://doi.org/10.1093/nar/gkp956
26. Chakraborty D, Collepardo-Guevara R, Wales DJ. Energy landscapes, folding mechanisms, and kinetics of RNA tetraloop hairpins. *J Am Chem Soc* 2014;**136**:18052–61. https://doi.org/10.1021/ja5100756
27. Ennifar E, Nikulin A, Tishchenko S *et al.* The crystal structure of UUCG tetraloop. *J Mol Biol* 2000;**304**:35–42. https://doi.org/10.1006/jmbi.2000.4204
28. Cheong C, Varani G, Tinoco I Jr. Solution structure of an unusually stable RNA hairpin, 5′GGAC(UUCG)GUCC. *Nature* 1990;**346**:680–2. https://doi.org/10.1038/346680a0
29. Nozinovic S, Fürtig B, Jonker HR *et al.* High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Res* 2010;**38**:683–94. https://doi.org/10.1093/nar/gkp956
30. Bottaro S, Nichols PJ, Vogeli B *et al.* Integrating NMR and simulations reveals motions in the UUCG tetraloop. *Nucleic Acids Res* 2020;**48**:5839–48. https://doi.org/10.1093/nar/gkaa399
31. Sorin EJ, Rhee YM, Nakatani BJ *et al.* Insights into nucleic acid conformational dynamics from massively parallel stochastic

simulations. *Biophys J* 2003;**85**:790–803. https://doi.org/10.1016/S0006-3495(03)74520-2
32. Borkar AN, Vallurupalli P, Camilloni C *et al.* Simultaneous NMR characterisation of multiple minima in the free energy landscape of an RNA UUCG tetraloop. *Phys Chem Chem Phys* 2017;**19**:2797–804. https://doi.org/10.1039/C6CP08313G
33. Ma H, Proctor DJ, Kierzek E *et al.* Exploring the energy landscape of a small RNA hairpin. *J Am Chem Soc* 2006;**128**:1523–30. https://doi.org/10.1021/ja0553856
34. Bottaro S, Lindorff-Larsen K. Mapping the universe of RNA tetraloop folds. *Biophys J* 2017;**113**:257–67. https://doi.org/10.1016/j.bpj.2017.06.011
35. Childs-Disney JL, Yang X, Gibaut QMR *et al.* Targeting RNA structures with small molecules. *Nat Rev Drug Discov* 2022;**21**:736–62. https://doi.org/10.1038/s41573-022-00521-4
36. Bouilly D, Hon J, Daly NS *et al.* Single-molecule reaction chemistry in patterned nanowells. *Nano Lett* 2016;**16**:4679–85. https://doi.org/10.1021/acs.nanolett.6b02149
37. Vernick S, Trocchia SM, Warren SB *et al.* Electrostatic melting in a single-molecule field-effect transistor with applications in genomic identification. *Nat Commun* 2017;**8**:15450. https://doi.org/10.1038/ncomms15450
38. Lee Y, Trocchia SM, Warren SB *et al.* Electrically controllable single-point covalent functionalization of spin-cast carbon-nanotube field-effect transistor arrays. *ACS Nano* 2018;**12**:9922–30. https://doi.org/10.1021/acsnano.8b03073
39. Wilson H, Ripp S, Prisbrey L *et al.* Electrical monitoring of sp3 defect formation in individual carbon nanotubes. *J Phys Chem C* 2016;**120**:1971–6. https://doi.org/10.1021/acs.jpcc.5b11272
40. Plesa C, Dekker C. Data analysis methods for solid-state nanopores. *Nanotechnology* 2015;**26**:084003. https://doi.org/10.1088/0957-4484/26/8/084003
41. Jang SS, Dubnik S, Hon J *et al.* Characterizing the conformational free-energy landscape of RNA stem-loops using single-molecule field-effect transistors. *J Am Chem Soc* 2023;**145**:402–12. https://doi.org/10.1021/jacs.2c10218
42. Bishop CM. *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
43. Verma AR, Ray KK, Bodick M *et al.* Increasing the accuracy of single-molecule data analysis using tMAVEN. *Biophys J* 2024;**123**:2765–80. https://doi.org/10.1016/j.bpj.2024.01.022
44. Fersht A. *Structure and Mechanism in Protein Science*. Singapore: World Scientific, 2017.
45. Sorgenfrei S, Chiu CY, Gonzalez RL Jr *et al.* . Label-free single-molecule detection of DNA-hybridization kinetics with a carbon nanotube field-effect transistor. *Nat Nanotechnol* 2011;**6**:126–32. https://doi.org/10.1038/nnano.2010.275
46. Choi Y, Olsen TJ, Sims PC *et al.* Dissecting single-molecule signal transduction in carbon nanotube circuits with protein engineering. *Nano Lett* 2013;**13**:625–31. https://doi.org/10.1021/nl304209p
47. Wallace MI, Ying L, Balasubramanian S *et al.* Non-Arrhenius kinetics for the loop closure of a DNA hairpin. *Proc Natl Acad Sci USA* 2001;**98**:5584–9. https://doi.org/10.1073/pnas.101523498
48. Ansari A, Kuznetsov SV, Shen Y. Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc Natl Acad Sci USA* 2001;**98**:7771–6. https://doi.org/10.1073/pnas.131477798
49. Munoz V, Thompson PA, Hofrichter J *et al.* Folding dynamics and mechanism of beta-hairpin formation. *Nature* 1997;**390**:196–9. https://doi.org/10.1038/36626
50. Howe CP, Greetham GM, Procacci B *et al.* Sequence-dependent melting and refolding dynamics of RNA UNCG tetraloops using temperature-jump/drop infrared spectroscopy. *J Phys Chem B* 2023;**127**:1586–97. https://doi.org/10.1021/acs.jpcb.2c08709
51. Howe CP, Greetham GM, Procacci B *et al.* Measuring RNA UNCG tetraloop refolding dynamics using temperature-jump/drop infrared spectroscopy. *J Phys Chem Lett* 2022;**13**:9171–6. https://doi.org/10.1021/acs.jpclett.2c02338
52. Borkar AN, Vallurupalli P, Camilloni C *et al.* Simultaneous NMR characterisation of multiple minima in the free energy landscape of an RNA UUCG tetraloop. *Phys Chem Chem Phys* 2017;**19**:2797–804. https://doi.org/10.1039/C6CP08313G
53. Dunitz JD. The entropic cost of bound water in crystals and biomolecules. *Science* 1994;**264**:670. https://doi.org/10.1126/science.264.5159.670
54. Karplus M. Behind the folding funnel diagram. *Nat Chem Biol* 2011;**7**:401–4. https://doi.org/10.1038/nchembio.565
55. Hänggi P, Talkner P, Borkovec M. Reaction-rate theory: fifty years after Kramers. *Rev Mod Phys* 1990;**62**:251–341. https://doi.org/10.1103/RevModPhys.62.251
56. Heus HA, Pardi A. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* 1991;**253**:191–4. https://doi.org/10.1126/science.1712983

Received: November 4, 2024. **Revised:** June 17, 2025. **Editorial Decision:** June 25, 2025. **Accepted:** July 8, 2025