Fast Folding of an RNA Tetraloop on a Rugged Energy Landscape Detected by a Stacking-Sensitive Probe

Krishnarjun Sarkar,[†] Konrad Meister,[†] Anurag Sethi,[†*] and Martin Gruebele[†‡§*]
[†]Department of Chemistry, [‡]Department of Physics, and [§]Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois

ABSTRACT We investigate the microsecond-timescale kinetics of the RNA hairpin g**a***cUUCGguc. The fluorescent nucleotide 2-aminopurine (**a***) reports mainly on base stacking. Ten kinetic traces and the temperature denaturation curve are globally fitted to four-state models of the free-energy surface. In the best-fitting sequential model, the hairpin unfolds over successively larger barriers in at least three stages: stem fraying and increased base-stacking fluctuations; concerted loss of hydrogen bonding and partial unstacking; and additional unstacking of single strands at the highest temperatures. Parallel and trap models also provide adequate fits: such pathways probably also play a role in the complete free-energy surface of the hairpin. To interpret the model states structurally, 200 ns of molecular dynamics, including six temperature-jump simulations, were run. Although the sampling is by no means comprehensive, five different states were identified using hydrogen bonding and base stacking as reaction coordinates. The four to five states required to explain the experiments or simulations set a lower limit on the complexity of this small RNA hairpin's energy landscape.

INTRODUCTION

Nucleic acids were the first biopolymers to have their folding and unfolding dynamics investigated systematically by fast temperature-jump (T-jump) methods (1). As experimental work has moved toward smaller structures to make direct connections with simulation, it has become clear that the energy landscapes of even the smallest RNA fragments, such as short tetraloop motifs, are highly rugged compared to small proteins (2–10). Multiphasic kinetics occur on timescales from picoseconds to milliseconds (11,12), and it can become difficult to disentangle loss of hydrogen bonding, base stacking, and evolving compactness as RNA unfolds.

Several experimental and computational studies have recently focused on the UNCG tetraloop motif with a short stem (one to four basepairs in the stem). Proctor et al. (6) showed that incorporation of bromine to make a UUC***G*** bromoguanosine loop sterically increases the unfolded-state free energy, favoring the folded state. Later Ma et al. (7) observed three unfolding relaxation phases (fast, medium, and slow) of the eight-nucleotide hairpin gcUUCGgc. Thus, at least four states were required to model its unfolding kinetics. A simple lattice model with an entropic correction factor accurately reproduced the experimental data. In accordance with the model, one of the four states, "**S**," was assigned as a kinetic trap, but based on its high fitted entropy, it was speculated that it could instead be a single-stranded state with highly variable base stacking. The question of base stacking was taken up quantitatively by Stancik et al. (13), who used structurally sensitive infrared fingerprints to detect gcUCCGgc unfolding induced by a T-jump. They concluded that the fastest phase monitors mainly base unstacking dynamics, whereas the medium (~5–10 μs) phase monitors mainly hydrogen-bond breakage. In silico, using replica exchange molecular dynamics simulations, Garcia and Paschek have studied the folding/unfolding equilibrium thermodynamics of the gcUUCGgc hairpin as a function of temperature and pressure, showing that the RNA exists in a diverse ensemble of conformations (10). Stock and colleagues analyzed replica exchange molecular dynamics trajectories of cUUCGg, concluding that it folds via intermediates and has a less cooperative melting transition than uCACGg (14).

To investigate further the relative role of hydrogen bond loss and base unstacking during RNA unfolding, we study the g**a***cUUCGguc tetraloop by fluorescence-detected T-jump kinetics and molecular dynamics. We insert an additional fluorescent 2-aminopurine residue **a*** into the sequence as a probe with stacking-sensitive fluorescence lifetime. As shown by Proctor et al., stems with two paired bases (8-mer) versus three paired bases (10-mer) do not show substantially different kinetics (15), although the longer-stem hairpin is slightly more stable.

We find that the 2-aminopurine probe yields three kinetic phases. The fast and medium phases have the same timescales as the corresponding phases observed by Ma et al. and Stancik et al. The slow phase has the same timescale as the slow phase observed by Ma et al. at high temperature. We globally fit all thermodynamic and kinetic fluorescence data to minimal four-state models. A sequential four-state model provides the most precise fit of the experimental data, although adequate parallel and trap model fits hint that such processes also play a role in the full free-energy landscape. This system is thus an ideal candidate for future

Submitted May 7, 2009, and accepted for publication June 30, 2009.

*Correspondence: sethi2@scs.uiuc.edu or gruebele@scs.uiuc.edu

Konrad Meister's present address is Lehrstuhl für Physikalische Chemie II, Ruhr-Universität Bochum, Bochum 44780, Germany.

Editor: Samuel Butcher.

0006-3495/09/09/1418/10 $2.00

doi: 10.1016/j.bpj.2009.06.035

multistate Markov chain modeling of the dynamics and thermodynamics (16,17).

At present, our 200 ns of molecular dynamics simulation reveal at least five structural ensembles. Taken together with the global experimental data analysis, the resulting picture is in general agreement with previous results, but does not neatly segregate hydrogen bonding and unstacking into separate phases. We distinguish two types of base stacking dynamics: increased stacking fluctuations, when bases rapidly move in and out of planar contact, and true base unstacking, where the distance between bases increases clearly (>3.5 Å) beyond the π-π interaction distance. Based on the simulations, we assign the fastest experimental phase to increased stacking fluctuations and fraying at the end of the stem. We assign the medium phase to concerted unfolding where the native hydrogen-bond network is entirely lost and simultaneously some unstacking occurs within single strands. We assign the slow phase, observed at the highest temperature, to a single-stranded state in which most stacking interactions are lost.

METHODS

RNA design and sample preparation

To study the folding of a small fluorescent RNA tetraloop hairpin similar to the UNCG motifs studied previously, we chose the sequence 5′-g**a***cUUCG-guc-3′ (stem nucleotides in lower case and loop nucleotides in upper case). "**a***" represents 2-aminopurine (2AP), an analog of adenine that we use as a fluorescence reporter in our folding and unfolding measurements (18–20). The 2AP fluorescence is known to be very sensitive to base stacking (19,20). It is also mildly sensitive to base pairing (20). The 2AP sensitivity to base pairing seems to come about indirectly through its effect on base stacking.

The 2AP-labeled oligonucleotide was purchased from Thermo Scientific (Waltham, MA). It was dissolved and stored in a low-salt $P_{10}E_{0.1}$ buffer (10 mM sodium phosphate and 0.1 mM Na_2EDTA, pH 7.1). This buffer is known to favor hairpin conformation (6). No changes in 2AP fluorescence were detected over a wide range of concentrations, indicating that the RNA mainly forms hairpins, not duplexes. The RNA concentration was measured by absorption, using an extinction coefficient of 83,400 M^{-1} cm^{-1} at 260 nm supplied by Thermo Scientific. The modeled molecular dynamics structure (see below) is analogous to the structure of the similar hairpin gcUUCGgc obtained from a ribosomal RNA by x-ray diffraction (21).

Thermal denaturation

Thermal unfolding was measured with a spectropolarimeter (JASCO, Easton, MD). The spectropolarimeter detected circular dichroism (CD), integrated fluorescence, and absorption. A thermoelectric cooler accompanied by a water bath was used to scan the temperature of the sample. Temperature melting curves were repeated several times with the same sample, as well as with different samples. The results were reproduced within measurement uncertainty. For direct comparison of the thermodynamic fluorescence denaturation experiments to the temperature-jump measurements, fluorescence was excited at 280 nm and the same Hoya B-370 (Hoya, Fremont, CA) filter was used in the fluorescence melt and T-jump measurements.

Temperature jump experiment

The relaxation kinetics of the RNA were obtained at final temperatures in the 313–353 K range by jumping the temperature of the RNA solutions by 10°. A fused silica sample cell of 200 μm path length was used. The temperature jump induced by a YAG laser heating pulse, Raman-shifted to 1.9 μm, had a dead time of 10 ns. Solvent-nucleic acid and solvent-solvent heat transfer is complete in <10 ps, well within our 10-ns dead time (11,22). After the jump, the temperature remains steady for >50 ms before diffusional cooling occurs; our total observation window is only 0.5 ms.

To probe hairpin relaxation kinetics, 2AP was excited every 14 ns at 281 ± 2 nm by a mode-locked tripled Ti:sapphire laser beam. In our experiments, the fluorescence lifetime of 2AP ranged from 3 to 3.5 ns (see Supporting Material), well within the 14-ns detection window. The fluorescence decays were digitized using a 500-ps resolution oscilloscope (Tektronix, Beaverton, OR), yielding 28 data points per 14-ns fluorescence detection window. The data were downloaded to a computer and analyzed by a program written in LabWindows (National Instruments, Austin, TX). A more detailed description of the setup is found in Ballew et al. (23). RNA concentrations of 160 μM and 500 μM were used. To obtain better signal/noise ratios, 30 temperature jumps were averaged for the 160-μM data and 11 for the 500 μM data. No difference was found between the two data sets within the signal/noise ratio, so duplex formation or aggregation is below the detection level. The size of the temperature jump was calibrated using the known tryptophan fluorescence lifetime change with temperature, obtained with a separate amino acid sample in the same cell configuration.

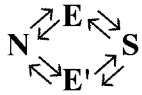
Data analysis and fitting

Both fluorescence lifetime and fluorescence intensity analyses were performed on the kinetics data, taken at five final temperatures. We used the Igor Pro software package for data fitting (23).

In the lifetime analysis, we fitted normalized fluorescence decays to a linear combination of the initial (just after the T-jump) and final (0.5 ms after the T-jump) fluorescence decays, allowing us to plot $\chi(t)$, the relative lifetime shift versus time (24). $\chi = 1$ means the fluorescence decay profile is identical to the one just after the T-jump, whereas $\chi = 0$ means the fluorescence decay profile has relaxed to the one at $t = 0.5$ ms. In the intensity analysis, we integrated the 28 data points collected per fluorescence decay. The time zero was determined from Raman scattering of the heating pulse in the sample cell, and decays near time zero (approximately five decays, or 70 ns of data) were not included in the analysis. The data was binned logarithmically, so that each phase and the baseline would contribute a comparable number of data points to the fit.

Two four-state models were used to fit all steady-state denaturation and kinetic fluorescence data simultaneously using nonlinear least squares:

$$\mathbf{N} \rightleftarrows \mathbf{E} \rightleftarrows \mathbf{U} \rightleftarrows \mathbf{S} \tag{1}$$



(2)

We label the states with single letters (in bold print) **N** for "native", **E** for "frayed", **U** for "unfolded", and **S** for "unstacked" to preserve the nomenclature of Ma et al. (7), although the interpretation of **S**, as suggested by its name, will be somewhat different. A simultaneous fit allowed us to use a small number of parameters to account for all of the data, and to better constrain the models. The free energies of the four states and of the three free-energy barriers connecting them were fitted to the linear model

$$\begin{aligned} \Delta G_i &= \Delta G_i^{(0)} + \Delta G_i^{(1)}(T - T_{i0}) \\ \Delta G_{i\leftrightarrow j}^{\dagger} &= \Delta G_{i\leftrightarrow j}^{(0)\dagger} + \Delta G_{i\leftrightarrow j}^{(1)\dagger}(T - T_{i0}) \end{aligned} . \tag{3}$$

The subscripts "i" and "j" run over states and T_0 is a reference temperature. Note that $\Delta G_i^{(1)}$ incorporates both the entropy and some enthalpy temperature dependence. A thermodynamically more consistent treatment of entropy and enthalpy could be achieved by fitting a constant heat capacity model (25). We use a linear model for consistency with the baseline analysis discussed in the next paragraph.

TABLE 1 Equilibrium thermodynamic parameters and activation energies of the four-state sequential model

State	T_0 (K)	$\Delta G^{(0)}$ (kJ mol^{-1})	$\Delta G^{(1)}$ (kJ mol^{-1} K^{-1})
N	326.67 (13)	0	0.4297 (42)
E	326.85 (16)	0	0.3678 (30)
U	327.0 (3.0)	0	0
S	345.21 (76)	0	−0.77 (18)
Transition state	$\Delta G^{(0)\dagger}$ (kJ mol^{-1})	$\Delta G^{(1)\dagger}$ (kJ mol^{-1} K^{-1})	k_m (μs^{-1})
N ↔ E	7.44 (44)	0.183 (26)	1
E ↔ U	11.42 (08)	−0.0513 (82)	
U ↔ S	19.1 (9)	0.274 (82)	

The four-state model (**N**, native; **E**, frayed, **U**, unfolded, and **S**, unstacked) is summarized in Eqs. 1 and 3 and Fig. 1. Two standard deviation uncertainties in the last digits are given in parentheses.

For each state, the same signal baseline, linear in temperature, was fitted for the fluorescence-intensity-detected kinetics and thermodynamics. This reduction in parameters was possible because the same detection filter was used in steady-state and kinetics experiments. A separate signal baseline, also linear in temperature, was necessary for the lifetime analysis.

$$S_i(\text{intensity or lifetime}) = S_i^{(0)} + S_i^{(1)}(T - T_{i0}) \quad (4)$$

The actual free energies and baselines are very likely somewhat nonlinear over the whole temperature range studied, and this should be taken into account when interpreting the parameters in Table 1. We were able to obtain a fit essentially within measurement uncertainty with the linearity assumption, so more complex free energy formulas and baselines are not yet warranted by the data.

To simulate kinetics, the fitting program first calculated the free energy of each state at the prejump temperature, T_{before}, using Eq. 3 above. The initial population, $P_i(t=0) = \exp[-\Delta G_i/RT_{\text{before}}]/\Sigma_j \exp[-\Delta G_j/RT_{\text{before}}]$, of each state was then assigned. The initial population was then placed on the free-energy surface recalculated at the higher temperature just after the temperature jump. The relaxation kinetics was simulated by matrix multiplication. In the sequential scheme of Eq. 1, the rate propagation becomes

$$\begin{bmatrix} P_N(t+\Delta t) \\ P_E(t+\Delta t) \\ P_U(t+\Delta t) \\ P_S(t+\Delta t) \end{bmatrix} = \begin{bmatrix} -k_{NE} & k_{EN} & 0 & 0 \\ k_{NE} & -k_{EN}-k_{EU} & k_{UE} & 0 \\ 0 & k_{EU} & -k_{UE}-k_{US} & k_{SU} \\ 0 & 0 & k_{US} & -k_{SU} \end{bmatrix} \otimes \begin{bmatrix} P_N(t) \\ P_E(t) \\ P_U(t) \\ P_S(t) \end{bmatrix} \times \Delta t + \begin{bmatrix} P_N(t) \\ P_E(t) \\ P_U(t) \\ P_S(t) \end{bmatrix}, \quad (5)$$

and in the parallel scheme of Eq. 2, the rate propagation becomes

$$\begin{bmatrix} P_N(t+\Delta t) \\ P_E(t+\Delta t) \\ P_{E'}(t+\Delta t) \\ P_S(t+\Delta t) \end{bmatrix} = \begin{bmatrix} -k_{NE}-k_{NE'} & k_{EN} & k_{E'N} & 0 \\ k_{NE} & -k_{EN}-k_{ES} & 0 & k_{SE} \\ k_{NE'} & 0 & -k_{E'N}-k_{E'S} & k_{SE'} \\ 0 & k_{ES} & k_{E'S} & -k_{SE}-k_{SE'} \end{bmatrix} \otimes \begin{bmatrix} P_N(t) \\ P_E(t) \\ P_{E'}(t) \\ P_S(t) \end{bmatrix} \times \Delta t + \begin{bmatrix} P_N(t) \\ P_E(t) \\ P_{E'}(t) \\ P_S(t) \end{bmatrix}. \quad (6)$$

To calculate rate coefficients, we used $k_{ij} = k_m(T)\exp[-\Delta G_{ij}^{\dagger}/RT]$. The preexponential factor $k_m(T) = k_m\eta(298\text{ K})/\eta(T)$ with a temperature-dependent viscosity, taken from Weast (26), assumes solvent friction in the normal Kramers regime (27). The actual viscosity dependence may be different due to internal friction or local hydrodynamic effects, but the simplicity of Eqs. 3 and 4 currently does not warrant a more detailed treatment of the prefactor. The average activation energies from Eq. 3 were used to calculate directional activation energies as $\Delta G_{ij}^{\dagger} = \Delta G_{i\leftrightarrow j}^{\dagger} + \Delta G_{ij}/2$ and $\Delta G_{ji}^{\dagger} = \Delta G_{i\leftrightarrow j}^{\dagger} - \Delta G_{ij}/2$ to satisfy detailed balance.

Molecular dynamics simulation setup

Molecular dynamics simulations of RNA have been carried out with AMBER, CHARMM, and OPLS force fields, each with its advantages and disadvantages (28). Our molecular dynamics simulations of the solvated complexes were performed using NAMD2 (29) with the CHARMM27 force field nucleic acid parameters (30,31). Parameters for 2AP were adapted from Sarzynska et al. (32) and were further checked using GAMESS (33). The structure of the 10-mer was obtained from the PDB structure 1K2G. All hydrogen atoms were added to the RNA molecule using the Psfgen plugin in VMD (34). The RNA molecules were explicitly solvated with TIP3 water molecules (35). To simulate experimental concentrations of buffer ions, 20 potassium ions and 10 chloride ions were initially placed at the minima of the Coulombic electrostatic interaction energy 6.5 Å away from the RNA molecules. The minima were calculated on a uniform grid using the program Ionize (36). The models were solvated in a two-step process. In the first step, Solvate 1.0 (37) was used with two Gaussians to add two layers of water molecules to the system. This placed water molecules in the solvation shell of the RNA and ions. Then, the Solvate 1.2 plugin to VMD was used to place the bulk water, resulting in a box size of 55 Å × 55 Å × 70 Å and a system size of ~18,000 atoms.

Simulation protocols

The native-state simulations were done with periodic boundary conditions in an NPT ensemble. A pressure of 1 atm was achieved by using a Langevin piston, and the temperature was set to 273 K using Langevin dynamics. Electrostatics were calculated with the particle-mesh Ewald method (38). The van der Waals interactions were calculated using a switching distance of 10 Å and a cutoff of 12 Å. Time steps for updates of bonded, van der Waals, and electrostatic calculations were 1, 2, and 4 fs, respectively. The system free energy was initially minimized using a four-step protocol in which the water molecules were allowed to associate with the macromolecule before the macromolecule was allowed to move. The minimization steps were 1), all heavy atoms fixed (2000 steps); 2), heavy atoms fixed except for water and ions (3000 steps); 3), macromolecule backbone atoms fixed (5000 steps); and 4), all atoms free to move (20,000 steps). During the initial equilibration, the system was gradually heated from 100 K to 273 K according to the protocol in Eargle et al. (39), during which different parts of the system were harmonically constrained. After this 4-ns equilibration, each system was run for 6 ns at 273 K. Under these conditions, the hairpin was stable.

Simulations at 298 K resulted in an unfolded RNA hairpin within 10 ns, indicating a much lower unfolding transition temperature of the molecular dynamics force field compared to experiment. To monitor the effect of salt concentration on the RNA hairpin's stability, we performed simulations at 298 K with three different salt compositions in the simulation box: 1), 10 potassium ions; 2), 20 potassium ions and 10 chloride ions; and 3), 3

magnesium ions, 14 potassium ions, and 10 chloride ions. None of these salt conditions shifted the melting point above 298 K. The disparity between the unfolding temperatures of the RNA hairpin in experiments and MD simulations indicates that the CHARMM parameters of RNA molecules have to be improved. Similar shifts in melting point have been observed in previous RNA molecular dynamics simulations (10).

Three different temperature jumps were simulated (from 273 K to 300 K, or to 333 K, or to 498 K at 216 atm). We started with two different conformations from the native-state simulation, for a total of six simulations. During each of these simulations, the temperature was increased uniformly from 273 K to the target temperature in 10 steps over a period of 250 ps. The pressure of the 498 K simulation was increased from 1 atm to 216 atm in a single step.

Analysis of the MD runs

Two variables were monitored in the analysis of the MD simulations: the number of native stem hydrogen bonds, and the integrity of native stem base stacking interactions. The general results obtained are independent of the exact definition of a broken hydrogen bond or lost stacking interaction. We describe the choices used in the figures in more detail.

The maximum number of hydrogen bonds in the RNA hairpin stem is eight. The number of native hydrogen bonds is calculated during each frame of the simulation. A hydrogen bond is formed under two conditions: the hydrogen-bond acceptor (A) atom and the hydrogen atom are within a cutoff distance of 3.5 Å from each other; and the measurement of the angle formed by the donor atom, the hydrogen atom, and the acceptor atom is 180 ± 30°. We also looked at hydrogen bonding in the loop. H-bonds from the crystal structure rapidly decayed to a low fluctuating level in the simulation and were not analyzed in detail. This could be real loop flexibility, or due to a force field that is too "floppy". However, Stock and co-workers compared the CACG and UUCG tetraloops with the same force field, and found the latter to be more flexible (14). Thus, at least some of the flexibility of the UUCG loop appears to be real.

The number of native base-stacking interactions between bases in the stem portion of the RNA hairpin (a maximum of four) is counted during each frame of the simulation. The maximum number of base-stacking interactions in the native structure is six in the RNA hairpin, including two loop-to-strand base-stacking interactions. Our RNA crystal structure taken from PDB 1K2G shows no base stacking in the loop, so a base stacking analysis in the loop was not performed. The three conditions under which a base-stacking interaction is formed are 1), two bases have nonhydrogen atoms that are within 3.5 Å of each other; 2), the normal vectors to the aromatic ring planes are at an angle of <30° of each other; and 3), the bases are not hydrogen-bonded to each other.

RESULTS

Thermal denaturation

The fluorescence melting curve of the RNA, obtained with a filter identical to that used in the T-jump experiments, is shown in Fig. 1 *A*. The native baseline persists up to 310 K, after which unfolding begins. The unfolding transition lasts to ~340 K, after which a strong unfolded baseline takes over. On its own, the data can be fitted by a two-state model with a midpoint at 330 K. When fitted by the same two-state model with linear baselines (Supporting Material), circular dichroism data show a midpoint that differs by a few degrees. Like the gcUUCGgc system studied earlier, the thermodynamic melting behavior of g**a***cUUCGguc alone is amenable to a deceptively simple two-state fit, if one is willing to neglect the small difference in CD and fluorescence melting points. The discrepancy between fluorescence- and CD-detected melting temperatures and the large slope differences between baselines hint that more states may be involved.

Temperature jump

The temperature-jump relaxation kinetics at three representative final temperatures are shown in Fig. 1, *B–G*. All temperature jumps were 10 ± 1°, and final temperatures in the range from 313 K to 353 K were monitored. The full set of data is included in the Supporting Material. Fluorescence lifetime and integrated fluorescence intensity were used as two different probes.

The lifetime-detected kinetics in Fig. 1, *B–D*, originated primarily from a fast component of the fluorescence decay ($\tau_{short} \approx 3$ ns) that accounted for ~70% of the fluorescence-decay amplitude. This component relaxed to a longer lifetime as the RNA unfolded after the T-jump. A longer fluorescence lifetime component of $\tau_{long} \approx 6$ ns that accounted for ~30% of the fluorescence amplitude switched within the T-jump dead time (<70 ns). We assign it to the intrinsic temperature dependence of unquenched 2AP fluorescence, although we cannot exclude the possibility of a fourth, very fast kinetic phase due to increased stacking fluctuations (vide infra).

Both the lifetime and intensity data show three resolvable kinetic components. At low temperature, the relaxation is biexponential, showing a fast phase of~2 μs along with a medium phase of 5–10 μs (depending on temperature). As the temperature increases, the fast phase amplitude decreases. Near the melting temperature, the kinetics appears nearly single-exponential. At even higher temperatures, the medium phase speeds up to a few microseconds as a new slow phase of ~50 μs appears. In the lifetime data, the slow phase has an amplitude with the same sign as the faster phases; in the intensity data, its amplitude has the opposite sign.

Four-state model fit

A four-state model is required minimally to explain the presence of the three resolved kinetic phases. Specifically, we tested Eq. 1 in the case where **U** is the most populated state at high temperature and **S** is a trap (trap model), Eq. 1 in the case where **S** is the most populated state at high temperature (sequential model), and Eq. 2 (parallel model). One barrier was held constant because of the barrier-prefactor correlation in the fit. The sequential model best fits the data. Its parameters are shown in Table 1. The fits are laid over the experimental data in Fig. 1 in red, and account for the data essentially within measurement uncertainty. The trap and parallel models adequately fit the data, but not within measurement uncertainty (for details, see Supporting Material). The blue trace in Fig. 1 *A* shows the best thermodynamic fit that could be achieved with the parallel model.

As indicated in Fig. 2 *B*, sequential and parallel (and trap) models are not mutually exclusive, but can be thought of as lower-dimensional projections that approximate a more

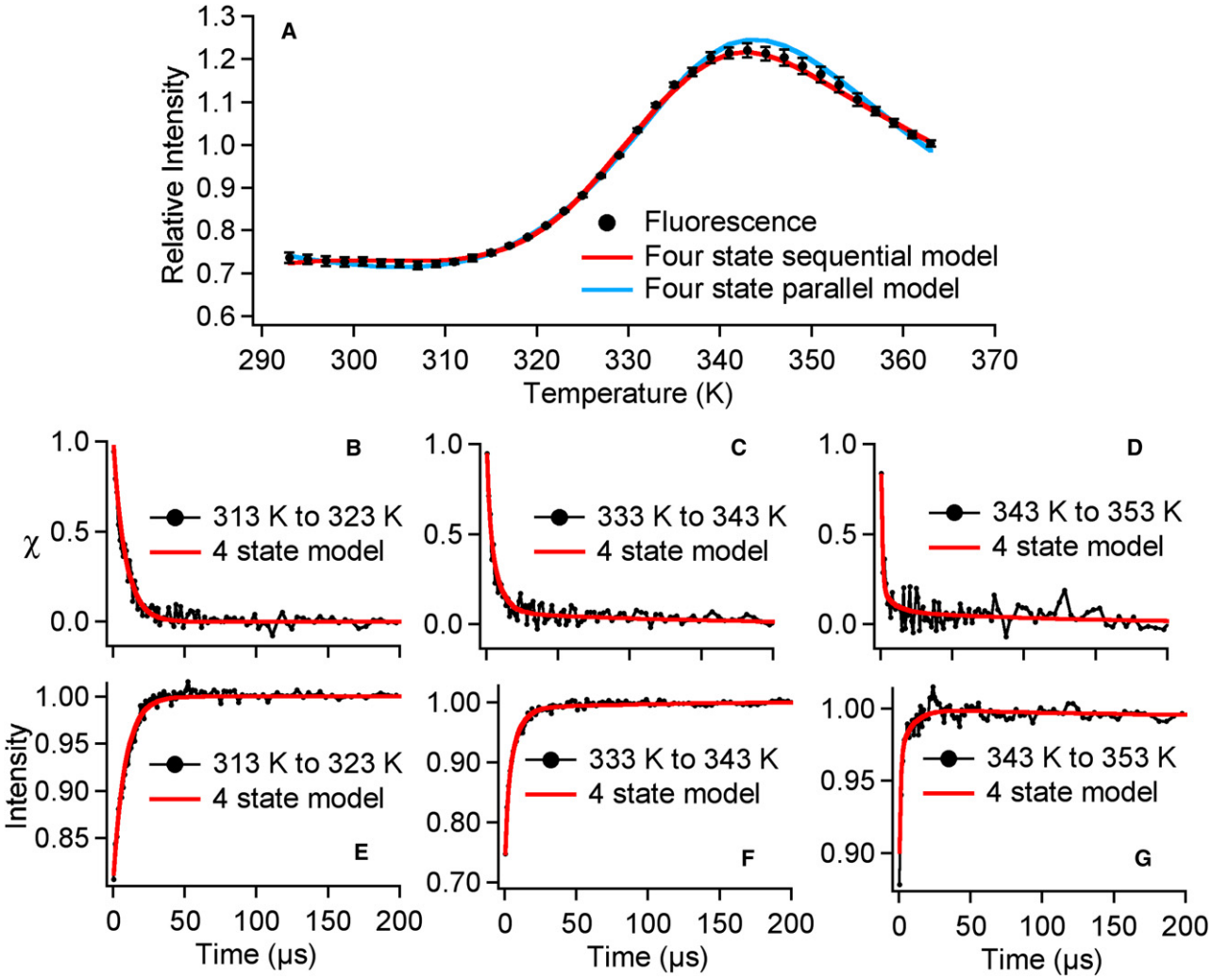


FIGURE 1 Experimental thermal denaturation and temperature-jump data fitted by a four-state sequential model. (*A*) Thermodynamic denaturation data and model fit (*red/light gray*). (*B–D*) Relative fluorescence lifetime change and model fit at three of the five temperatures measured (see also Supporting Material). (*E–G*) Relative fluorescence intensity change and model fit. A parallel model, which does not fit quite as well, is shown in blue for comparison (for details on this and a trap model, see Supporting Material).

comprehensive multidimensional free-energy surface. The requirement that thermodynamic and kinetic data be fitted simultaneously, with the same baselines for fluorescence intensity, was what distinguished the models and yielded the sequential model as the best "0^{th}-order" approximation. Rates alone would not be sufficient to distinguish the models.

The sequential model with populations shown in Fig. 2 *A* invokes the four states **N**, **E**, **U**, and **S** defined previously. These states are sequentially more unfolded, as evidenced by an increasingly negative slope of free energy with temperature, indicative of increased state entropy (Table 1, *upper*). With the sequential model, a linear free energy dependence on temperature was sufficient to account for the data. The orange arrows in Fig. 2 *B* indicate the dominant population transfer in this model.

The data could also be fitted to alternative trap or parallel four-state models. In the parallel model, population is transferred simultaneously to states **E** and **E′**, as indicated by the red dotted arrows in Fig. 2 *B*. However, this model resulted in an inferior fit of the data (Supporting Materials and Fig. 1 *A*, *blue trace*). The trap model also yielded an inferior fit, at the same time requiring additional quadratic free-energy parameters (Supporting Material). Thus, we favor the sequential model as the simplest approximation needed to describe the data. As will be seen in the discussion of the molecular dynamics simulations, the kinetic fitting models do not rule out the existence of additional states. Four states are simply the minimum number we needed to fit the resolved kinetic phases within measurement uncertainty.

Molecular dynamics simulations

To aid in the structural interpretation of the four states, we complemented the experiment with MD simulations (Figs. 3

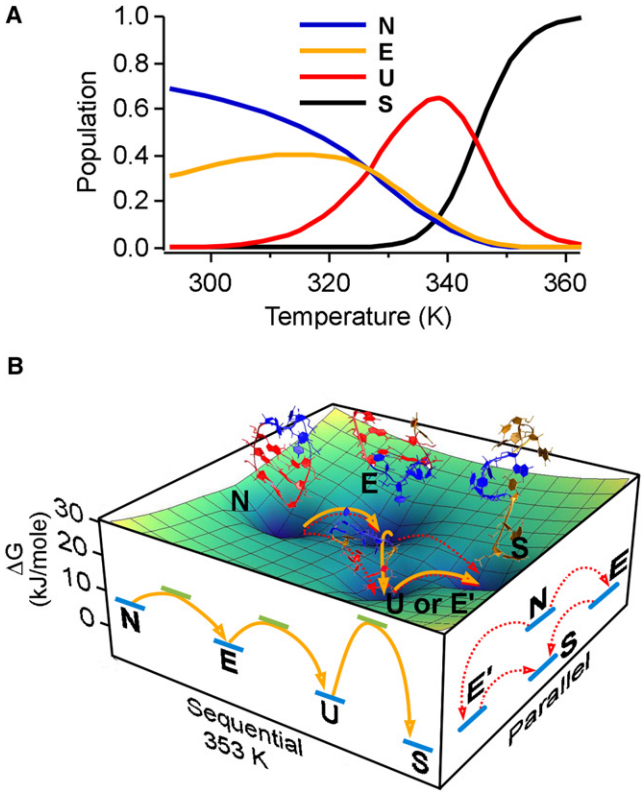


FIGURE 2 State populations and free energies: **N**, native state; **E,** frayed state; **U,** unfolded state without native hydrogen bonding but with residual stacking; **S,** single-stranded unstacked denatured state. (*A*) Populations from the four-state sequential model. (*B*) Free energies and structures of the four fitted states at 353 K on a two-dimensional free-energy landscape, projected onto the sequential and parallel models. Although the sequential model (*solid arrows*) provides the best one-dimensional fit, parallel processes (*dotted arrows*) and traps (see Supporting Material) can also make contributions to the complex dynamics of this RNA hairpin.

and 4). The RNA was simulated at 273 K, as well as two jumps each from 273 K to the final temperatures of 300, 333, and 498 K. The MD results could be grouped into five ensembles based on stem hydrogen bonding and stem base stacking (Fig. 3 *B*). Sequestration of populations into further subensembles cannot be ruled out with a small number of trajectories. Hydrogen bonding and stacking in the loop was found to be very labile even at low temperatures, and could not be used to classify states (Fig. 4, *upper left*). Before we interpret the experimental data with the aid of molecular dynamics trajectories, it should be cautioned that in silico, RNA is not as thermally stable as observed experimentally, and also unfolds much faster.

Representative microstate examples of the most prominent structural ensembles are shown in Fig. 3 *A*. The native state (**N**) has, on average, two-thirds of the native hydrogen bonds formed. On average, two-thirds of the stem bases are stacked. Yet even the equilibrium trajectory in Fig. 4, *upper left*, shows one major excursion to a frayed state with fewer hydrogen bonds and less base stacking. The frayed state (**E** or **E′**) is populated more extensively than the **U** and **S** states

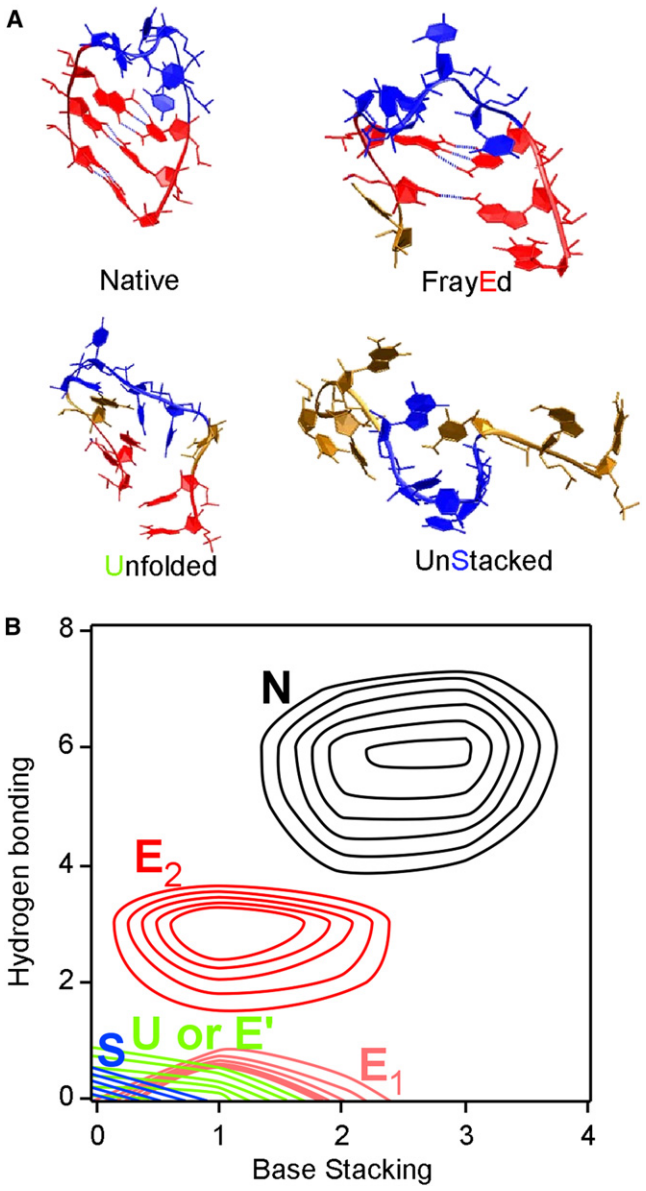


FIGURE 3 Molecular dynamics simulation structural results are shown. (*A*) Representative conformations of the native, frayed, unfolded, and fully unstacked states. Arrows in Fig. 4 indicate the times at which these configurations were sampled. *Blue*, loop; *red*, stacked stem bases; *ochre*, unstacked stem bases; *blue dotted lines*, native hydrogen bonds. (*B*) Conformational probability plot of stem hydrogen bonds and stem stacking interactions color-coded by temperature: *black*, 273 K equilibrium trajectory; *red*, after jumps to 300 K; *green*, after jumps to 333 K; *blue*, after jumps to 498 K. The outermost contour for each ensemble is at 50% of the highest-probability contour. At least five conformational ensembles or subensembles are detected by our limited MD sampling.

in the smallest simulated T-jumps from 273 to 300 K (Fig. 4, *upper right*). Immediately after the T-jump, the RNA shows increased stacking and hydrogen-bond fluctuations, even though these native interactions are not immediately broken. A concomitant loss of base stacking and hydrogen bonding then occurred in both 300 K T-jump simulations. Afterward, the frayed state sampled in the simulations has a variable

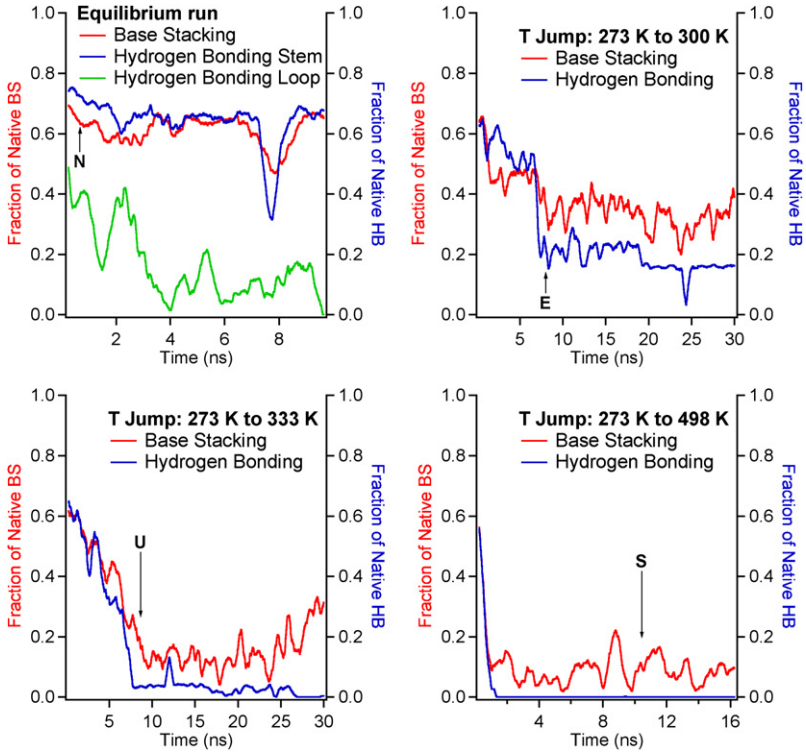


FIGURE 4 Molecular dynamics trajectory results are shown. Equilibrium molecular dynamics, at 273 K, is shown at the upper left; loop hydrogen bonds (*green*/*bottom trace*) and loop stacking (not shown) made relatively small contributions and were not analyzed further. The remaining three panels show simulations of three temperature jumps starting at time $t = 0$. Native stem base stacking (*red*/*light*) and native stem hydrogen bonding (*blue*/*dark*) are shown, with each time point averaged over a 0.5-ns window. The arrows indicate the point at which the conformations shown in Figs. 2 and 3 *A* were sampled. The T-jumps are each an average of two trajectories with different random initial conditions. The individual trajectories are shown in the Supporting Material.

number of native hydrogen bonds remaining in the stem (E_1 and E_2 substates in Fig. 3 *B*), whereas it still has 35% of base stacking (Fig. 4, *lower left*). A much more unfolded state, **U**, is reached in the T-jump to 333 K (Fig. 4, *lower left*). Most of the native hydrogen bonding is lost at this stage, although 20% of base stacking persists. Finally, in the temperature jump to 498 K (Fig. 4, *lower right*), even the base stacking is reduced to 10%, with large fluctuations as transient interactions are formed and broken during the simulation. Indeed, 10% may be an overestimate, based on our lax criterion for base stacking (see Methods). The hydrogen-bond and stacking patterns in Fig. 4 are averaged from two trajectories at each temperature. The individual simulations can be found in the Supporting Material.

On the whole, the loss of hydrogen bonding with increasing temperature is much more facile than the loss of single-stranded base stacking, but increased stacking fluctuations occur earlier than loss of hydrogen bonding. Based on our limited sampling, the MD simulations support the following sequence of events as temperature is increased:

$$\textbf{N}\text{ative} \rightarrow \text{Fray}\textbf{E}\text{d} \rightarrow \textbf{U}\text{nfolded} \rightarrow \text{Un}\textbf{S}\text{tacked}$$

At the same time, the MD simulations support the idea that multiple frayed states (**E** and **E**′) are populated in parallel from the native state.

To summarize, the native state has the majority of stem hydrogen bonds and stem stacking interactions formed at any given instant, but the loop is more flexible (Figs. 3 *B* and 4, *upper left*). The frayed states have increased stacking fluctuations and suffer disruption of hydrogen bonds, but retain considerable single-stranded stacking. The unfolded state has lost most hydrogen bonding, but retains some single-stranded stacking. The unstacked state maintains no significant hydrogen bonding, and little single-stranded stacking. Thus, stacking fluctuations and loss of stacking bracket loss of hydrogen bonding.

DISCUSSION

We use 2AP as an unfolding probe for RNA because it is mainly sensitive to base stacking (20,40,41). Stacking causes a reduction of the fluorescence intensity in two stages. When a stacking interaction first forms, static quenching occurs. Quenching is enhanced when 2AP is stacked with a purine compared to a pyrimidine. Once two bases are stacked, the fluorescence quenching still depends on fluctuations of the stacked geometry. More coplanar geometries with fewer fluctuations result in more quenching, referred to as dynamic quenching (19). The reverse of these processes would be increased stacking fluctuations followed by true unstacking of the bases beyond a critical distance. A loss of rigidity from breakage of hydrogen bonds also has been shown to cause a small change in 2AP fluorescence (20). On the other hand, it has been claimed, based on quantum calculations on 2AP, that hydrogen bonding has a minimal effect on 2AP fluorescence characteristics (42).

Our RNA stem is longer than the one used by Ma et al. (7) and Stancik et al. (13) by one basepair, and has a different loop base than the latter. Based on previous measurements, these differences are unlikely to change the general conclusions

we make here. Proctor et al. studied kinetics as a function of stem length, and found only small changes in rate (15). Ma et al. (7) studied a very different loop (UUUU), and still found four phases, though they occurred in slightly different proportions compared to the UUCG loop or UCCG loop (13).

In the study by Ma et al. (7), a four-state model like the one in Fig. 1 fitted three phases with a temperature dependence similar to that observed here. The favored interpretation of state **S** was a misfolded trap, based on its structural assignment in a lattice simulation. It was also speculated that state **S** could be a single-stranded state with little or no hydrogen bonding and variable single-strand stacking. Our best global model fit is in better agreement with the latter interpretation. The slope, $d\Delta G_i/dT$, becomes increasingly negative in the progression **N**-**E**-**U**-**S** (Table 1, *upper*), indicating again that **S** has the highest entropy of the four states. A state with low hydrogen bonding and stacking satisfies this requirement, but a misfolded compact state does not. Attempts to fit a trap that is depopulated in favor of **U** at the highest temperatures produces an inferior fit even though it requires more fitting parameters (Supporting Material). The molecular dynamics simulation shows that although stacking persists at moderate temperatures (Fig. 4, *upper right*), it is almost completely lost at higher temperatures.

In the study by Stancik et al. (13), base-stacking rearrangements are assigned to the earliest phase, and hydrogen-bond breaking follows thereafter based on structure-specific infrared transient absorption data. Our fast and medium phases (corresponding to unquenching of 2AP fluorescence) are in general agreement with the phases observed by Stancik et al., except that the two coordinates (stacking and hydrogen bonding) are not as neatly assignable to separate phases. Our data at 323 K have significant amplitudes for both the fastest (1–2 μs) and medium (6–8 μs) phases fitted by the four-state model. Since 2AP is stacking-sensitive, changes in stacking occur in both phases. The molecular dynamics simulation supports loss of stacking on three different timescales. Immediately after the temperature jumps in Fig. 4 ($t > 0$), larger fluctuations of the native stacking are observed. The stacking is then reduced in concert with loss of hydrogen bonding and continues to fluctuate more after the hydrogen bonding is broken. Only at the highest temperature is native stacking completely lost, even within single strands. By analogy, we assign our fast experimental phase to increased fluctuations of base stacking after the temperature jump and the medium phase to coordinated loss of hydrogen bonding and stacking, whereas the slow phase, observed only at high temperature, is due to further unstacking of the single RNA strand.

In another study focused specifically on loop dynamics, Porschke and co-workers conducted temperature jumps on GNRA type tetraloops (where R is purine), with a 2AP substitution in the loop (8). Both single- and double-exponential kinetics were observed. For the GA*AA mutant, a single 22-μs phase was seen. For the GAA*A mutant, a biexponential with lifetimes of 5 and 41 μs was observed, not unlike our medium and slow phases. It appears that 2AP samples somewhat less heterogeneity in the GNRA loop than in our stem context, but both stem and loop kinetics clearly contribute substantially to the overall complexity of the free-energy landscape.

The RNA free-energy landscape is thus much more rugged than the free-energy landscape of small peptides. With a prefactor of $k_m = 1\ \mu s^{-1}$ in Table 1, similar to that measured for β-sheet peptides (43), the barriers separating the four states range from 7 to 19 kJ/mol at the reference temperature, T_0, of each state. For small peptide hairpins with aromatic stacking interactions, simulations estimate a root-mean-square roughness of only ~2.5 kJ/mol (44) and measurements of relaxation rate fluctuations yield a root-mean-square value of 2 kJ/mol (45), findings that are corroborated by single-molecule experiments (46). Although four different probes (ultraviolet absorption, infrared absorption at two frequencies, and 2AP quenching) have so far resolved three phases, it is possible that more than four states exist. Both our fluorescence data and the infrared data (13) show unresolvable phases. Such phases may be intrinsic temperature dependences of the chromophores, or they may signal additional transitions involving smaller structural changes or fast relaxation down steep parts of the free-energy surface accessed by the RNA population right after the temperature jump. Picosecond temperature jumps have certainly resolved additional fast dynamics (11). It would be interesting to see whether future experiments with different or faster probes detect additional phases, or whether extensive simulations coupled with Markov analysis (16,17) allow classification of more states.

Our four-state model with 8–20 kJ/mol free-energy barriers should be viewed as a minimum model for roughness of the RNA free-energy landscape, even though it reproduces the data nearly within the signal/noise ratio. Together with the kinetic data, the molecular dynamics simulation indicates that stacking fluctuations contribute to unfolding on the shortest timescales, whereas concerted unstacking and hydrogen-bond loss contribute at intermediate times, and unstacking within single strands becomes important at higher temperatures. The currently available data can be fitted by a simple sequential model, but additional parallel states or traps cannot be ruled out. The fast timescale of the experiments makes RNA hairpins amenable to future simulation with improved force fields that better reproduce stability and absolute kinetic timescales than the simulations described here. The $<100\ \mu$s relaxation times observed for small RNAs unfortunately are not resolvable by currently available single-molecule experiments. Even for the simplest RNA secondary structures, the energy landscape may turn out to be more complex yet.

SUPPORTING MATERIAL

Six figures and two tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(09)01206-5.

K.M. was a visiting student from the Ruhr-Universität Bochum. A.S. thanks Professor Zan Schulten and John Eargle for helpful discussions.

This work was supported by National Science Foundation grant MCB 0613643. Computational work was carried out on a cluster supported by a National Science Foundation Chemistry Research Instrumentation and Facilities grant (CHE 0541659). K.S. thanks the Center for Physics of the Living Cell for funding.

REFERENCES

1. Williams, A. P., C. E. Longfellow, S. M. Freier, R. Kierzek, and D. H. Turner. 1989. Laser temperature-jump, spectroscopic, and thermodynamic study of salt effects on duplex formation by dGCATGC. *Biochemistry*. 28:4283–4291.
2. Kuznetsov, S. V., C. C. Ren, S. A. Woodson, and A. Ansari. 2008. Loop dependence of the stability and dynamics of nucleic acid hairpins. *Nucleic Acids Res*. 36:1098–1112.
3. Ansari, A., S. V. Kuznetsov, and Y. Q. Shen. 2001. Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc. Natl. Acad. Sci. USA*. 98:7771–7776.
4. Wang, X. J., and W. M. Nau. 2004. Kinetics of end-to-end collision in short single-stranded nucleic acids. *J. Am. Chem. Soc*. 126:808–813.
5. Ansari, A., and S. V. Kuznetsov. 2005. Is hairpin formation in single-stranded polynucleotide diffusion-controlled? *J. Phys. Chem. B*. 109: 12982–12989.
6. Proctor, D. J., E. Kierzek, R. Kierzek, and P. C. Bevilacqua. 2003. Restricting the conformational heterogeneity of RNA by specific incorporation of 8-bromoguanosine. *J. Am. Chem. Soc*. 125:2390–2391.
7. Ma, H. R., D. J. Proctor, E. Kierzek, R. Kierzek, P. C. Bevilacqua, et al. 2006. Exploring the energy landscape of a small RNA hairpin. *J. Am. Chem. Soc*. 128:1523–1530.
8. Menger, M., F. Eckstein, and D. Porschke. 2000. Dynamics of the RNA hairpin GNRA tetraloop. *Biochemistry*. 39:4500–4507.
9. Koplin, J., Y. G. Mu, C. Richter, H. Schwalbe, and G. Stock. 2005. Structure and dynamics of an RNA tetraloop: a joint molecular dynamics and NMR study. *Structure*. 13:1255–1267.
10. Garcia, A. E., and D. Paschek. 2008. Simulation of the pressure and temperature folding/unfolding equilibrium of a small RNA hairpin. *J. Am. Chem. Soc*. 130:815–816.
11. Ma, H. R., C. Z. Wan, A. G. Wu, and A. H. Zewail. 2007. DNA folding and melting observed in real time redefine the energy landscape. *Proc. Natl. Acad. Sci. USA*. 104:712–716.
12. Lin, M. M., L. Meinhold, D. Shorokhov, and A. H. Zewail. 2008. Unfolding and melting of DNA (RNA) hairpins: the concept of structure-specific 2D dynamic landscapes. *Phys. Chem. Chem. Phys*. 10: 4227–4239.
13. Stancik, A. L., and E. B. Brauns. 2008. Rearrangement of partially ordered stacked conformations contributes to the rugged energy landscape of a small RNA hairpin. *Biochemistry*. 47:10834–10840.
14. Villa, A., E. Widjajakusuma, and G. Stock. 2008. Molecular dynamics simulation of the structure, dynamics, and thermostability of the RNA hairpins uCACGg and cUUCGg. *J. Phys. Chem. B*. 112:134–142.
15. Proctor, D. J., H. R. Ma, E. Kierzek, R. Kierzek, M. Gruebele, et al. 2004. Folding thermodynamics and kinetics of YNMG RNA hairpins: specific incorporation of 8-bromoguanosine leads to stabilization by enhancement of the folding rate. *Biochemistry*. 43:14004–14014.
16. Chodera, J. D., N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope. 2007. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *J. Chem. Phys*. 126:155101.
17. Noe, F., and S. Fischer. 2008. Transition networks for modeling the kinetics of conformational change in macromolecules. *Curr. Opin. Struct. Biol*. 18:154–162.
18. Jean, J. M., and K. B. Hall. 2004. Stacking-unstacking dynamics of oligodeoxynucleotide trimers. *Biochemistry*. 43:10277–10284.
19. Jean, J. M., and K. B. Hall. 2001. 2-Aminopurine fluorescence quenching and lifetimes: role of base stacking. *Proc. Natl. Acad. Sci. USA*. 98:37–41.
20. Ballin, J. D., S. Bharill, E. J. Fialcowitz-White, I. Gryczynski, Z. Gryczynski, et al. 2007. Site-specific variations in RNA folding thermodynamics visualized by 2-aminopurine fluorescence. *Biochemistry*. 46:13948–13960.
21. Ennifar, E., A. Nikulin, S. Tishchenko, A. Serganov, N. Nevskaya, et al. 2000. The crystal structure of UUCG tetraloop. *J. Mol. Biol*. 304:35–42.
22. Wang, Z. H., Y. Pang, and D. D. Dlott. 2004. Vibrational energy dynamics of water studied with ultrafast Stokes and anti-Stokes Raman spectroscopy. *Chem. Phys. Lett*. 397:40–45.
23. Ballew, R. M., J. Sabelko, C. Reiner, and M. Gruebele. 1996. A single-sweep, nanosecond time resolution laser temperature-jump apparatus. *Rev. Sci. Instrum*. 67:3694–3699.
24. Ballew, R. M., J. Sabelko, and M. Gruebele. 1996. Observation of distinct nanosecond and microsecond protein folding events. *Nat. Struct. Mol. Biol*. 3:923–926.
25. Nishii, I., M. Kataoka, F. Tokunaga, and Y. Goto. 1994. Cold denaturation of the molten globule states of apomyoglobin and a profile for protein folding. *Biochemistry*. 33:4903–4909.
26. Weast, R. C. 1997. CRC Handbook of Chemistry and Physics. CRC Press, Cleveland, OH.
27. Kramers, H. A. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*. 7:284–304.
28. Hashem, Y., and P. Auffinger. 2009. A short guide for molecular dynamics simulations of RNA system. *Methods*. 47:187–197.
29. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, et al. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem*. 26:1781–1802.
30. MacKerell, A. D. J., B. Brooks, I. C. L. Brooks, L. Nilsson, B. Roux, et al. 1998. CHARMM: the energy function and its parameterization with an overview of the program. *In* The Encyclopedia of Computational Chemistry. P. von Ragué Schleyer, P. R. Schreiner, N. L. Allinger, T. Clark, and J. Gasteiger, et al., editors. John Wiley & Sons, New York. 271–277.
31. Foloppe, N., and A. D. MacKerell. 2000. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem*. 21:86–104.
32. Sarzynska, J., K. Kulinska, and T. Kulinski. 2003. Parametrization of 2-aminopurine and purine in CHARMM all-atom empirical force field. *Comput. Methods Sci. Technol*. 9:127–135.
33. Schmidt, M. W., K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, et al. 1993. General atomic and molecular electronic-structure system. *J. Comput. Chem*. 14:1347–1363.
34. Humphrey, W. F., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph*. 14:33–38.
35. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. 79:926–935.
36. Schulten, K.2009. Ionize. www.ks.uiuc.edu/Development/MDTools/ionize.
37. Grubmüller, H., and V. Groll. 2009. Solvate 1.0. *http://www.mpibpc.mpg.de/home/grubmueller/downloads/solvate/index.html*.
38. Darden, T., D. York, and L. Pedersen. 1993. Particle-mesh Ewald: an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys*. 98:10089–10092.
39. Eargle, J., A. A. Black, A. Sethi, L. G. Trabuco, and Z. Luthey-Schulten. 2008. Dynamics of recognition between tRNA and elongation factor Tu. *J. Mol. Biol*. 377:1382–1405.
40. Rachofsky, E. L., R. Osman, and J. B. A. Ross. 2001. Probing structure and dynamics of DNA with 2-aminopurine: effects of local environment on fluorescence. *Biochemistry*. 40:946–956.

41. Hall, K. B., and J. Williams. 2004. Dynamics of the IRE RNA hairpin loop probed by 2-aminopurine fluorescence and stochastic dynamics simulations. *RNA*. 10:34–47.

42. Hardman, S. J. O., and K. C. Thompson. 2006. Influence of base stacking and hydrogen bonding on the fluorescence of 2-aminopurine and pyrrolocytosine in nucleic acids. *Biochemistry*. 45:9145–9155.

43. Nguyen, H., M. Jaeger, J. Kelly, and M. Gruebele. 2005. Engineering a β-sheet protein towards the folding speed limit. *J. Phys. Chem. B*. 109:15182–15186.

44. Yang, W. Y., J. Pitera, W. Swope, and M. Gruebele. 2004. Heterogeneous folding of the trpzip hairpin: full atom simulation and experiment. *J. Mol. Biol*. 336:241–251.

45. Yang, W. Y., and M. Gruebele. 2004. Detection-dependent folding kinetics as a probe of folding landscape microstructure. *J. Am. Chem. Soc*. 126:7758–7759.

46. Schuler, B. 2005. Single-molecule fluorescence spectroscopy of protein folding. *Chem. Phys. Chem*. 6:1206–1220.