

Folding of Human Telomerase RNA Pseudoknot Using Ion-Jump and Temperature-Quench Simulations

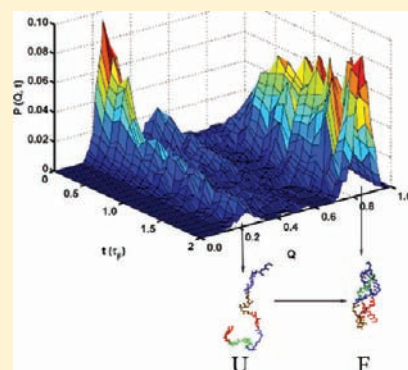
Shi Biyun,^{†,‡} Samuel S. Cho,^{†,#} and D. Thirumalai^{*,†,§}

[†]Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, United States

[‡]Bio-X Laboratory, Department of Physics and Soft Matter Research Center, Zhejiang University, Hangzhou 310027, China

[§]Department of Chemistry, University of Maryland, College Park, Maryland 20742, United States

ABSTRACT: Globally RNA folding occurs in multiple stages involving chain compaction and subsequent rearrangement by a number of parallel routes to the folded state. However, the sequence-dependent details of the folding pathways and the link between collapse and folding are poorly understood. To obtain a comprehensive picture of the thermodynamics and folding kinetics we used molecular simulations of coarse-grained model of a pseudoknot found in the conserved core domain of the human telomerase (hTR) by varying both temperature (T) and ion concentration (C). The phase diagram in the $[T, C]$ plane shows that the boundary separating the folded and unfolded state for the finite 47-nucleotide system is relatively sharp, implying that from a thermodynamic perspective hTR behaves as an apparent two-state system. However, the folding kinetics following single C -jump or T -quench is complicated, involving multiple channels to the native state. Although globally folding kinetics triggered by T -quench and C -jump are similar, the kinetics of chain compaction are vastly different, which reflects the role of initial conditions in directing folding and collapse. Remarkably, even after substantial reduction in the overall size of hTR, the ensemble of compact conformations are far from being nativelike, suggesting that the search for the folded state occurs among the ensemble of low-energy fluidlike globules. The rate of unfolding, which occurs in a single step, is faster upon C -decrease compared to a jump in temperature. To identify “hidden” states that are visited during the folding process we performed simulations by periodically interrupting the approach to the folded state by lowering C . These simulations show that hTR reaches the folded state through a small number of connected clusters that are repeatedly visited during the pulse sequence in which the folding or unfolding is interrupted. The results from interrupted folding simulations, which are in accord with non-equilibrium single-molecule folding of a large ribozyme, show that multiple probes are needed to reveal the invisible states that are sampled by RNA as it folds. Although we have illustrated the complexity of RNA folding using hTR as a case study, general arguments and qualitative comparisons to time-resolved scattering experiments on *Azoarcus* group I ribozyme and single-molecule non-equilibrium periodic ion-jump experiments establish the generality of our findings.



INTRODUCTION

In order to function, RNA molecules have to adopt nearly compact structures, thus making it necessary to understand their folding in quantitative detail. Folding landscapes of RNAs are rugged^{1–6} because they readily adopt alternate structures with free energies that are not significantly higher than the native fold. Consequently, the stability gap relative to $k_B T$ (k_B is the Boltzmann constant and T is the temperature), which is the free energy difference between the native state and other low free energy conformations,^{7,8} is only modest. As a result there is a substantial probability that RNA molecules are kinetically trapped in alternate misfolded structures. Trapping in low-lying free energy states gives rise to complex folding kinetics even in the formation of simple hairpins.^{9–12} Although the general principles governing the folding of RNA are in place,^{3,13–16} the interplay of multiple conflicting energy scales makes it difficult to map the major stages of RNA folding from an initial ensemble of unfolded conformations.

Folding (unfolding) of RNA can be triggered by increasing (decreasing) the concentration, C , of counterions. In typical ensemble experiments C (or T) is used to trigger folding or unfolding of RNA.^{1,17,18} Considerable insights into the folding mechanisms of RNA have come from such experiments,^{13,15,16,19,20} theoretical arguments,³ and simulations using coarse-grained (CG) models.^{21–24} These studies have shown that RNA folds by multiple parallel pathways and is succinctly summarized by the kinetic partitioning mechanism (KPM).^{16,25} According to KPM a fraction (Φ) of initially unfolded molecules reaches the native basin of attraction (NBA), whereas the remaining fraction is kinetically trapped in competing basins of attraction (CBAs). The long transition times from CBAs to NBA result in slow folding of ribozymes.^{5,26} Single-molecule fluorescence resonance energy transfer (smFRET) experiments

Received: October 2, 2011

Published: November 14, 2011

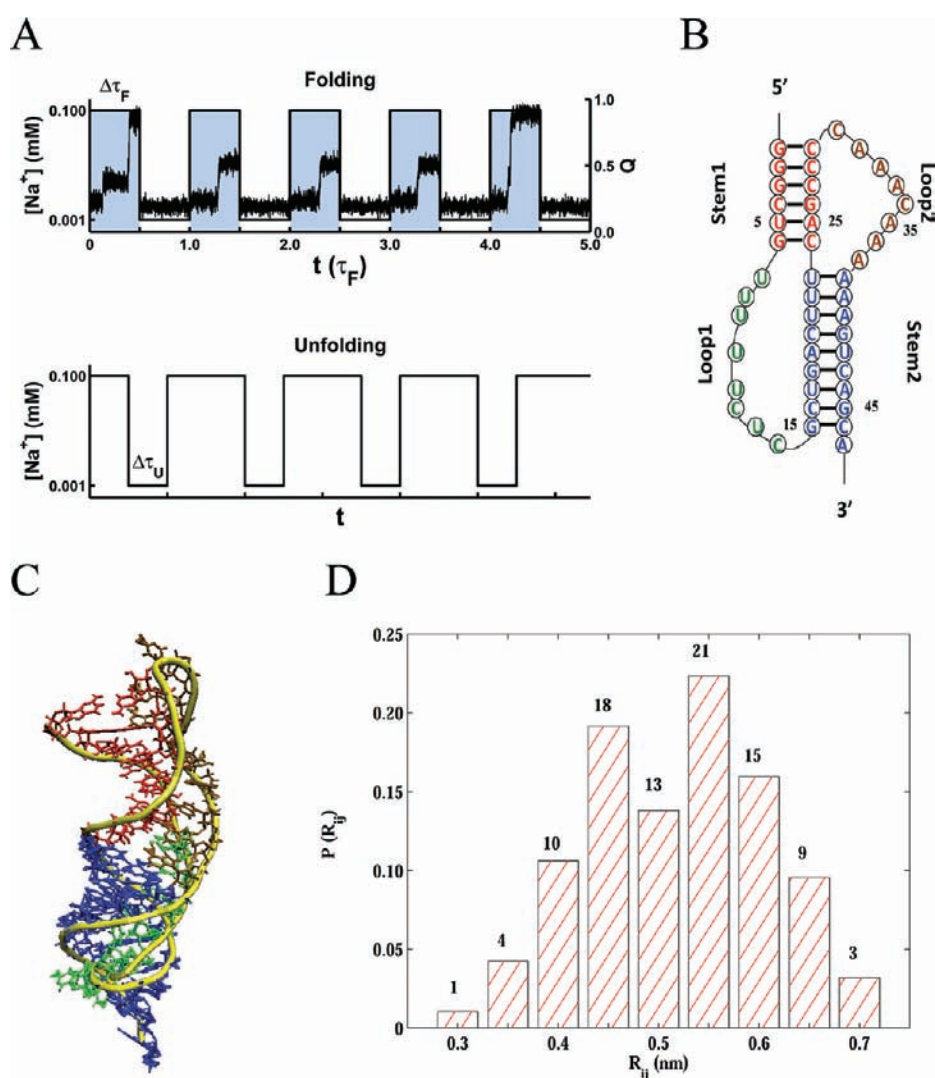


Figure 1. Ion pulse and structural features of human telomerase pseudoknot. (A) The top panel shows a folding pulse in which the ion concentration is increased from a low value to a high value for the duration $\Delta\tau_F$. Time-dependent changes in the fraction of native contacts, Q (scale for Q is on the right) in a sample trajectory is superimposed. The lower panel is an unfolding pulse in which C is lowered to interrupt folding. (B) Secondary structure of the 47-nucleotide hTR in which the stems and loops are labeled. (C) Tertiary structure of hTR. The backbone is colored in yellow. (D) Distribution of distances between the bases in the native state shown in (C). Two bases are in contact if the distance between them in an arbitrary conformation is within a ± 0.2 of that in the native state. Small R_{ij} values correspond to distances between bases in the stems.

performed by immobilizing *Tetrahymena* ribozyme on a surface²⁷ have given additional insights into RNA folding. Analysis of the smFRET trajectories following a single jump in ion concentration quantitatively confirmed the predictions of the KPM.²⁸ Cumulatively, these studies firmly established that RNA molecules fold by multiple pathways with an array of time scales, thus confirming the rugged nature of their folding landscapes.

Single ion-jump experiments^{29,30} are useful in describing the dominant ensemble of intermediate structures $\{I_i\}$ that accumulate under folding conditions. However, due to limitations in temporal resolution, they cannot be used to probe the ensemble of structures that are sampled prior to the formation of $\{I_i\}$ or the NBA. These states, if they are accessed during the folding process, would be “hidden” in conventional near-equilibrium smFRET experiments. The nature of hidden conformations can be revealed by interrupting folding after initiating an ion-jump.³¹ Consider an ion-pulse sequence (Figure 1A), which shows that

following an ion-jump the high C condition is maintained for a duration $\Delta\tau_F$ prior to a decrease in C . If $\Delta\tau_F$ is less than the C -dependent folding time, $\tau_F(C)$, or the time to reach $\{I_i\}$, then the conformations that are sampled in the pathways leading to the folded state or $\{I_i\}$ can be characterized. In smFRET experiments the states are described using the distribution of FRET efficiencies measured just prior to a decrease in T . By varying $\Delta\tau_F$ it is possible to probe structures that are hidden in typical experiments. The utility of pulse sequences, which was first demonstrated using theory and simulations³² and subsequently realized in experiments in the context of single-molecule force spectroscopy,³³ has been demonstrated in an ingenious experiment to reveal the underlying complexity of RNase P RNA folding.³¹ In the experimental study $\Delta\tau_F$, which for most part was kept constant, was apparently not long enough to complete the folding reaction even at high ion concentration.

Inspired by the interrupted multiple ion-jump single-molecule experiments³¹ we performed simulations using a coarse-grained model of the human telomerase (hTR) pseudoknot.³⁴ First, we calculated the thermodynamics of the hTR pseudoknot in terms of a phase diagram in the $[T, C]$ plane.³ The folded (F) and the unfolded (U) states are separated by a well-defined boundary, which shows that the melting temperature, T_m , increases nonlinearly as C increases. Folding kinetics using single ion-jump or T -quench shows that the flux to the folded state occurs through two dominant channels. In all trajectories folding is preceded by collapse of the hTR, as assessed by a decrease in the radius of gyration. The time-dependent changes in the radius of gyration show that even after substantial compaction the collapsed structures are not natively like. There are dramatic differences in the approach to the native state between folding initiated by C -jump or T -quench. By performing several interrupted ion-jumps, with varying waiting times $\Delta\tau_F$, we establish that there are multiple folding pathways reflecting the heterogeneity of the folding process. Conformations, which are difficult to characterize in a single ion-jump, become visible as $(\Delta\tau_F)/(\tau_F(C))$ is varied. These results, which are qualitatively similar to those found in the folding of RNase P RNA, show that the complete characterization of the states that are sampled during the folding process will require a combination of nonequilibrium experiments³¹ and simulations.

MODELS AND METHODS

Coarse-Grained TIS Model for RNA. We used a modified form of the three interaction site (TIS) model to represent RNA. The energy function and parameters are described in detail elsewhere.²¹ Briefly, in the TIS model each nucleotide is represented by three beads (or interaction centers) corresponding to the base, ribose sugar, and the charged phosphate group. The Hamiltonian for the TIS model is $H_T = H_C + H_{NB}$. The term, H_C , accounts for chain connectivity and rotational degrees of freedom. The term H_{NB} representing interactions between sites that are not directly connected to each other is $H_{NB} = H_{NC} + H_{stacking} + H_{elec}$. The term H_{NC} represents stabilizing native contact interactions, and $H_{stacking}$ is for base stacking interactions based on Turner's Rules.^{35–37} Finally, H_{elec} accounts for electrostatic repulsions between the charged phosphate groups. For interactions between phosphate groups P_i and P_j we use the Debye–Hückel potential,

$$V_{P_i P_j} = \frac{z_{P_i} z_{P_j} e^2}{4\pi\epsilon_0\epsilon_r r} e^{-r/l_D} \quad (1)$$

We vary the ion concentration in our simulations by changing the Debye length $l_D = ((8\pi l_B I)^{1/2})^{-1}$ for monovalent cations. Here, $l_B \approx 7 \text{ \AA}$ is the Bjerrum length at room temperature and I , the ionic strength, is proportional to the ion concentration. In our earlier study,³⁴ we showed that the melting temperatures of three pseudoknots obtained using the TIS model simulations were in excellent agreement with experimental measurements, which we consider to be a key validation of the coarse-graining procedure. We should note that there are other theoretical approaches that have also reproduced the thermodynamics of pseudoknot folding.³⁸

Simulation Details. We used Langevin dynamics in the low friction limit, which has been shown to enhance conformational sampling of proteins and RNA, to calculate various thermodynamic properties.³⁹ In order to produce realistic estimates of folding and unfolding times we performed Brownian dynamics simulations using a friction coefficient that is appropriate for water.⁴⁰ For each simulation condition (specified by ion concentration or temperature) we generated 100 trajectories, which seem sufficient to obtain converged results.

Simulations at friction coefficient corresponding to water viscosity allow us to map the simulation time scales to real times, as detailed in the Supporting Information of ref 34. All simulations were performed for the human telomerase pseudoknot (PDB code: 1YMO), referred to as hTR through out the paper.

Analysis of the Trajectories. To monitor the collapse transition we calculated the radius of gyration using

$$R_g = \sqrt{\frac{\sum_{ij} r_{ij}^2}{2N^2}} \quad (2)$$

where r_{ij} is the distance between sites i and j , and N is the total number of interaction centers in RNA. The folding and unfolding transitions are determined using the fraction of native contacts, Q . Two interaction sites are in tertiary contact if the distance between them is less than the cutoff, r_c . We only consider contacts between bases, which suffice to uniquely specify the conformations of hTR. With this definition there are 94 tertiary contacts in the TIS representation of the native structure of hTR (1YMO). The native contacts are computed by considering bases including those that are part of the two stems (see Figure 1B). With such a definition we are assured that in the folded state both secondary and tertiary interactions are fully developed. In the TIS representation of the native state of hTR there is a distribution of distances involving contacts between bases in the native state (Figure 1D). The values of r_c vary, depending on the contact. For each contact we allow a tolerance of about 20% to take thermal fluctuations into account. Thus, if a contact is at a distance R_N in the native state then it is assumed to be a native contact in any arbitrary conformation if the distance between them is at $R_N + 0.2R_N$.

We characterize the ensemble of unfolded conformations using a number of order parameters including q -score⁴¹ given by

$$q = \frac{1}{(N-1)(N-2)} \sum_{i < j-1}^N \exp\left[-\frac{(r_{ij} - r_{ij}^0)^2}{\sigma_{ij}^2}\right] \quad (3)$$

where ij labels the non-neighboring pair, r_{ij} is the distance between the two beads, and r_{ij}^0 is the native distance between two beads. The value of q varies from 0 to 1 with $q = 1$ corresponding to the native state. We also calculated the structural overlap function,⁴² $\chi = 1 - (N_K)/(N_T)$ where

$$N_K = \sum_{i < j-1}^N H(\delta - |r_{ij} - r_{ij}^0|) \quad (4)$$

with $N_T = [(n-1)n]/2$, $H(x)$ is the Heavyside function, and $\delta = 0.1 \text{ nm}$. Note that if $\chi = 0$ the conformation is in the naive state. At finite temperatures both χ and q -score have well-defined values. In the TIS representation of the 47-nucleotide hTR the value of $n = 140$ because the first nucleotide has only two interaction centers.

RESULTS

Phase Diagram in C and T Plane. We determined the phase diagram as a function of C and T of the experimentally well-characterized hTR pseudoknot.^{43–45} The hTR pseudoknot has 47 nucleotides with a secondary structure (Figure 1B) that corresponds to the commonly observed H-type RNA pseudoknot, which is defined as two double-stranded and base-paired helices that are each connected by single-stranded loops (Figure 1B and C). We designate the two helices as Stem1 (S_1) and Stem2 (S_2), and the loops as Loop1 (L_1) and Loop2 (L_2). Both the stems are predicted to be stable based on Mfold,⁴⁶ and could form RNA hairpins (S_1 and S_2) in isolation.³⁴ There are extensive Hoogsteen base triples between S_2 and L_1 , which

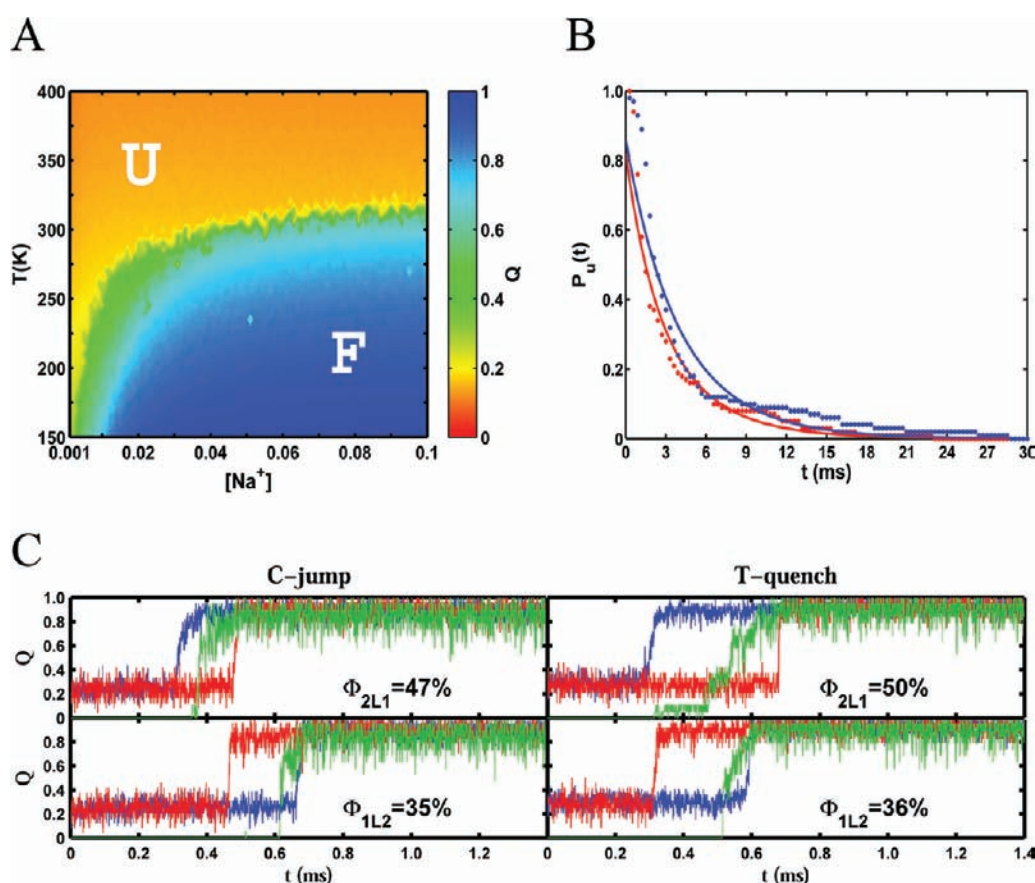


Figure 2. Phase diagram and folding kinetics. (A) phases of hTR pseudoknot as a function of temperature (T) and ion concentration (C). The concentration of monovalent ions is in mM and temperature is in Kelvin. The phases are assigned on the basis of the fraction of native contacts Q whose scale is on the right. The boundary between the NBA and UBA corresponds to $Q = 0.5$. (B) Fraction of unfolded molecules upon T -quench (blue) and C -jump (red). The T -quench simulations are performed with $C = 0.1$ mM, and T is quenched from 400 K to 250 K. For the C -jump T is fixed at 250 K, and the ion concentration is increased from 0.001 mM to 0.1 mM. The solid lines are fits to the data accounting for flux to the NBA through two dominant channels and are given by $P_U(t) = \Phi_1 \exp(-t/\tau_1) + \Phi_2 \exp(-t/\tau_2)$. The values for Φ_1 and Φ_2 for C -jump folding are 0.47 (0.50) and 0.35 (0.36), respectively, where the numbers in parentheses are for T -quench folding. Similarly, the values of τ_1 and τ_2 are 4.3 ms (4.9 ms) and 2.0 ms (3.0 ms), respectively. (C) Folding trajectories for two dominant pathways by which hTR folds. The contact formation involving S_2 , L_2 , and S_1 , are in blue, green, and red, respectively. The fraction Φ_{2L1} (Φ_{1L2}) corresponds to the formation of S_2 (S_1) followed by L_2 (S_2) and finally S_1 (S_2). The left panel is for C -jump folding, and the right is for T -quench folding.

we classify as tertiary interactions. Recently, Cho, Pincus, and Thirumalai (hereafter referred to as CPT) showed that the thermodynamic properties of hTR pseudoknot could be accurately predicted using the TIS model.³⁴ The folding transition was highly cooperative with a large value of the dimensionless cooperative measure.^{34,47}

To calculate the phase diagram of the hTR pseudoknot in the $[C, T]$ plane, we performed a series of coarse-grained MD simulations using the TIS model in the low friction underdamped limit.³⁹ Multiple simulations were performed over a broad range of ion concentrations (0.001 mM to 0.1 mM) and temperatures (150 K to 400 K). During the time course of each trajectory, we monitored the degree of folding by Q , the fraction of native contacts (See Models and Methods for a formal definition). The ensemble average, $\langle Q \rangle$, is used as an order parameter in the calculation of the phase diagram (Figure 2A). From a thermodynamic perspective there are two distinct states that correspond to the unfolded (red) and folded (blue) states with a boundary between the two states. However, there is a region (shown as green in Figure 2A), which perhaps should be classified as intermediate states. The locus of points separating the folded

and unfolded states can be approximately described as $T_m \approx C_m^\alpha$ with $\alpha \approx 0.3$.

Kinetics of hTR Folding. CPT showed that the related VPK pseudoknot folds in ~ 5 ms upon temperature quench. The measured folding times in a recent temperature jump experiment¹⁸ are in near quantitative agreement with our earlier predictions, which further validates the efficacy of the TIS model in predicting the overall folding times. With this as an additional validation of the TIS model and the Brownian dynamics protocol, we carried out two kinds of folding simulations, one in which folding is triggered by a temperature quench and the other by an ion concentration jump. The folding times were estimated using $\tau_F = \int P_U(s) ds$ where $P_U(t)$ is the probability that the molecule has not folded at time t . We computed $P_U(t) = \int_0^t P_{FP}(s) ds$ from the distribution of first passage times $P_{FP}(s)$ obtained from a hundred independent trajectories. The first passage time for each trajectory is the time at which a molecule starting from an unfolded state reaches the folded state for the first time. We assume that the folded state is reached if Q is 0.8.

In the CPT study it was established that upon temperature quench hTR folded by parallel pathways in a few milliseconds.

Here, we carried out T -quench and ion-jump folding kinetics of hTR so that a detailed comparison between the two methods used to initiate folding can be made. Figure 2B shows the time dependence of $P_U(t)$ when folding is initiated by ion-jump and T -quench. In both cases, $P_U(t)$ is well fit using $P_U(t) = \Phi_1 \exp(-t/\tau_1) + \Phi_2 \exp(-t/\tau_2)$ where Φ_1 and Φ_2 are the amplitudes representing flux of the two dominant channels to the folded state. Our results are in accord with previous work⁴⁸ showing the use of a master equation that pseudoknots fold by biphasic kinetics. The values of Φ_1 , Φ_2 , τ_1 , and τ_2 are given in the caption to Figure 2. Using extensive structural analysis CPT showed that, for T -quench folding, the biexponential fits for $P_U(t)$ imply that there are two dominant parallel (and a few subdominant) routes to folding. A similar interpretation holds well for folding induced by jumping the ion concentration (see below). Interestingly, the global parameters obtained from $P_U(t)$ fits characterizing both the T -quench and ion-jump folding are similar. Thus, the differences in folding mechanisms on time scales exceeding about one millisecond do not significantly depend on the differences between the initial conformations at high temperatures and low ion concentrations. As detailed below, the effect of the initial conditions are lost on time scales that are comparable to the overall collapse time, τ_c .

Parallel Routes to Folding. In both T -quench and C-jump folding, we find that there are two dominant pathways that account for flux in excess of 80% to the native state (Figure 2B). In addition, there are four minor pathways to the folded state. The time dependence of the fraction of native contacts $Q(t)$ for two sample trajectories, one for C-jump and the other for T -quench, illustrates the distinct routes to the native state (Figure 2C). In pathway I, S_2 forms first, which subsequently nucleates interactions with L_2 . Subsequently, S_1 forms, leading to tertiary structure formation. In pathway II, the order of formation is folding of S_1 followed by interactions involving L_2 , and finally consolidation of S_2 and the native structure. The amplitudes for these pathways using T -quench and C-jump folding are similar (Figure 2C), which suggests that the overall folding mechanisms do not change significantly. However, on time scales that are shorter than τ_c , the nature of the populated structures are different, depending on the protocol used to trigger folding (see below).

Link between Collapse and Folding. In order to establish the link between collapse and folding transitions we computed the time-dependent changes in $\langle R_g(t) \rangle$ upon initiation of folding. The normalized radius of gyration,

$$S_g(t) = \frac{\langle R_g(t) \rangle - R_g^N}{\langle R_g(0) \rangle - R_g^N} \quad (5)$$

where $\langle R_g(t) \rangle$ is the average (over an ensemble of trajectories) radius of gyration at time t , and R_g^N (~ 2.0 nm) is the size of hTR in the folded state as shown in Figure 3A. Upon ion-jump and T -quench, $S_g(t)$ decays in two major steps. The fast time constant (see the fits in the caption to Figure 3A) is 5 times shorter when collapse occurs by ion-jump. In both cases the bulk of the reduction in the radius of gyration occurs during the initial rapid step. The value of $\langle R_g(0) \rangle$ (~ 4.0 nm) corresponding to the thermally equilibrated ensemble is considerably smaller (~ 7.6 nm) than that at low C (Figure 3A). Interestingly, the degree of collapse measured by $S_g(t)$ is far greater upon C-jump than by T -quench. For example, upon C-jump $S_g(t) \approx 0.2$ at $t \approx 1$ ms, whereas during T -quench, even when $t \approx 14$ ms, $S_g(t)$ is above 0.2 (Figure 3A). Nevertheless, for a fixed S_g the absolute

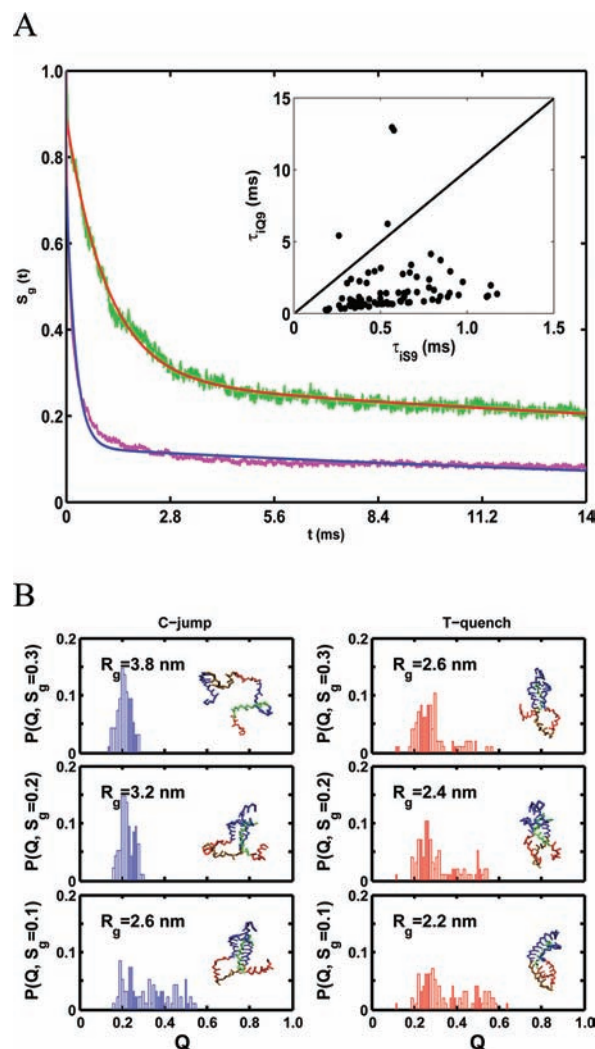


Figure 3. Chain compaction and folding. (A) Extent of chain compaction (eq 5) as a function of t following T -quench (green) and C-jump (purple). The red curve, a fit to the data in green, is given by $S_g(t) = 0.61 \exp(-t/\tau_{c1}) + 0.28 \exp(-t/\tau_{c2})$ where τ_{c1} and τ_{c2} are 1.1 and 44.1 ms, respectively. The corresponding fit to the purple curve is $S_g(t) = 0.61 \exp(-t/\tau_{c1}) + 0.13 \exp(-t/\tau_{c2})$ where τ_{c1} and τ_{c2} are 0.24 and 25.6 ms, respectively. Because $\langle R_g(0) \rangle$ for C-jump is much greater than for the thermally equilibrated ensemble, the absolute value of $R_g(t)$ is greater when folding is initiated by C-jump than by T -quench. The inset shows a plot of t_{iQ9} as a function of t_{iS9} where t_{iQ9} and t_{iS9} are the times when 90% of the native contact and chain compaction is reached. The line is the expected result if folding and collapse occur simultaneously. The plot reveals that in all trajectories collapse precedes folding. (B) Distribution of fraction of native contacts obtained at various stages of chain compaction. The left panel is for C-jump and the right panel is for T -quench. The values of R_g are also indicated. In the folded state $R_g^N \approx 2.0$ nm.

values of $\langle R_g(t) \rangle$ are less at all times upon T -quench than by C-jump (Figure 3B), which is a consequence of the dramatic difference in the values of $\langle R_g(0) \rangle$.

Although average values of folding and collapse time scales are instructive in ascertaining the link between folding and collapse, a more precise picture emerges by examining the behavior of individual molecules. We plot t_{iQ9} and t_{iS9} for each folding trajectory i where t_{iQ9} (t_{iS9}) denotes the time at which 90% of

folding (collapse) has occurred. As before, folding is assessed by Q . If folding and collapse occur nearly simultaneously, then the $[t_{iQ}, t_{iS}]$ plot should be linear with a slope of near unity as indicated by the line in the inset to Figure 3A. However, the inset in Figure 3A shows that in all the trajectories collapse occurs on time scales that are considerably shorter than the first passage time. Although on an average the folding time is about an order of magnitude greater than the collapse time, there are great variations in the behavior of individual molecules. A substantial number of points are clustered around the lower left of the $[t_{iQ}, t_{iS}]$ plot. For this subset of molecules the collapse times are about a factor of (5–10) smaller than the folding time. Collapse here may be considered to be specific in the sense that they produce nativelylike compact structures. However, the ensemble of hTR may be topologically frustrated, leading to the formation of a few non-native conformations. This finding provides a structural interpretation for the prediction made in an earlier study.⁴⁸ For a large number of molecules there is a clear separation in collapse and folding time scales, which is likely a consequence of incomplete formation of the necessary interactions to nucleate the folded conformation. The partitioning of the initial pool of molecules into two distinct populations with very different folding behavior is consistent with earlier predictions based on KPM, and is in accord with a variety of experiments.

The extent to which nativelylike structures form as hTR collapses can be obtained by plotting the distribution of the fraction of native contacts that are formed when hTR has compacted to varying fractions of the starting size. In the panels on the left side of Figure 3B we show $P(Q, t = t_{iSk})$, which is the distribution of Q calculated at various stages of chain compaction following a jump in C . The corresponding distribution upon T -jump is on the right side in Figure 3B. Surprisingly, both these distributions are far from being nativelylike, which implies that, despite acquiring compact structures, the folding (as measured by Q) is incomplete. We also find that at this level of compaction the secondary structures (S_1 or S_2) are not fully formed. The lack of nativelylike character even after considerable compaction is due to formation of incorrect contacts due to topological frustration. Even after nearly 90% compaction the ensemble of structures after C -jump and T -quench remain as only near-native globules. This implies that rearrangement to the native stage in a large fraction of molecules occurs at times that are longer than the collapse time.

Differences between C-Jump and T-Quench Folding. The results in Figure 2C show that hTR folds by two dominant pathways upon a single jump in C or a T -quench. Therefore, the differences between the two protocols must occur only on times that are comparable to collapse time scales. Such a difference is expected because the initial conformations at high T and low C are dramatically different. Because the order parameter to characterize the state of the finite-sized RNA and proteins is not unique we calculated the distribution of six quantities, which describe the structures at $t = 0$. For each of these order parameters (see Figure 4A) we show the distributions of the order parameters calculated from 100 independent conformations from which folding is initiated. The panels shaded in yellow in Figure 4A correspond to a fixed high T and the lower panels are for low C . The distribution functions $P(R_{ee})$ and $P(R_g)$ show that the structures at low C are considerably more extended than the equilibrium structures at high T . Similarly, these structures are less nativelylike than the counterparts generated at

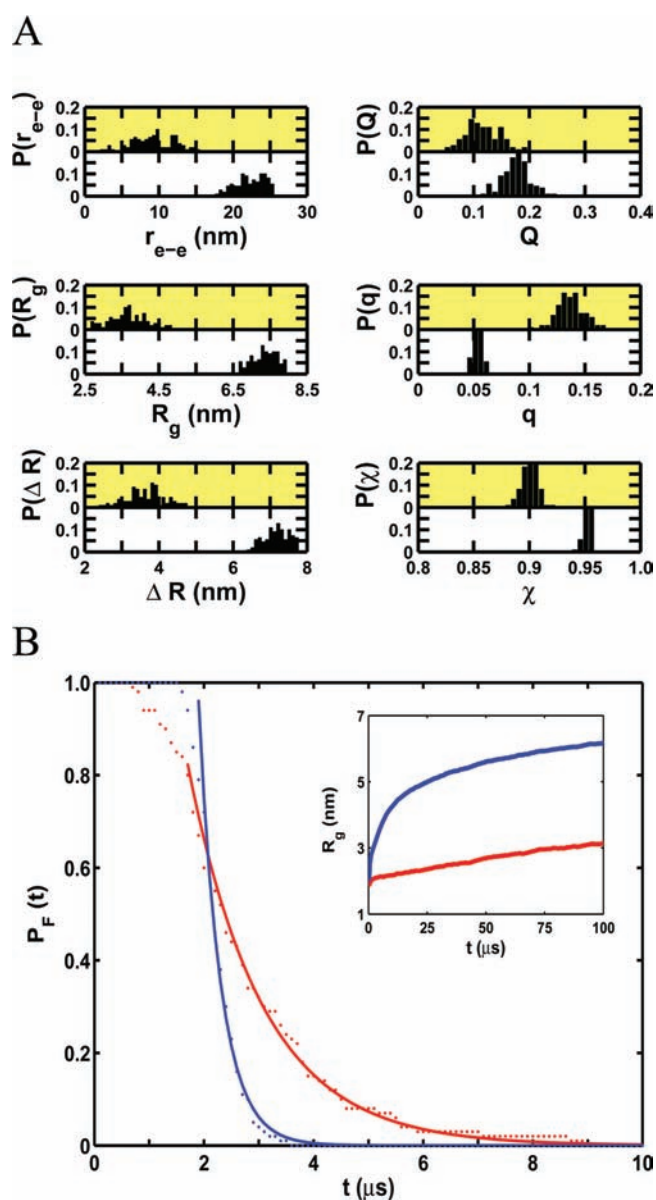


Figure 4. Characterizing the unfolded ensemble and unfolding kinetics. (A) The ensemble of unfolded states are described by six order parameters: end-to-end distance (R_{ee}), R_g , root-mean square deviation (ΔR) from the native state, contact q -score (eq 3), structural overlap function χ (eq 4), respectively. The distribution functions are computed from the ensemble of 100 equilibrated conformations from which the folding simulations are initiated. Yellow figures are the results at $T = 400$ K and $C = 0.1$ mM, and those in white correspond to $T = 250$ K and $C = 0.001$ mM. The unfolded states at low C and high T are similar according to the Q , q -score, and χ . However, there are significant differences between them according to the other three parameters indicating that different probes are likely to report distinct aspects of hTR. (B) Blue (red) corresponds to the fraction of molecules that remains folded ($P_F(t)$) as a function of t following an increase in C (T). The fits can be represented as $P_F(t) \approx e^{-(t)/(\tau_U)}$ where $\tau_U = 0.4 \mu\text{s}$ (blue) and $\tau_U = 1.4 \mu\text{s}$ (red). The inset shows swelling measured by $R_g(t)$ following an increase in C (blue) and T (red).

high T (see $P(\Delta R)$ in Figure 4A). Because all the structures from which folding commences are more compact at high T than at low C the initial routes explored on time scales on the order of

τ_C are different. These differences are clearly manifested in the rates of hTR collapse (see Figure 3A). The detailed structural attributes of hTR (see the right panels in Figure 4A) clearly show that initial structures have very little resemblance to the folded state. It is clear from the distribution functions in Figure 4A that the initial structures at low C have much less overlap with the native state than the high temperature structures.

Kinetics of hTR Unfolding. To understand how hTR unfolds we performed simulations by decreasing the ion concentration and by raising the temperature. In both sets of simulations the initial 100 structures are obtained from equilibrating the conformations in the NBA. The fraction, $P_F(t)$, of molecules that remains folded at time t following a sudden decrease in C or increase in T is used to quantify the unfolding kinetics. In both cases $P_F(t)$ can be fit using an exponential function (unfolding is two-state-like) after a transient time (Figure 4B). The unfolding rate is greater when C is suddenly decreased compared to the thermal melting rate. The time constant for unfolding caused by a decrease in C is $\sim 0.2 \mu\text{s}$, which is about 6 times faster than the time for thermal melting (see caption in Figure 4B). The inset in Figure 4B shows that kinetics of swelling follow either an increase in T (red) or decrease in C (blue). The overall swelling rate is also greater when C is decreased compared to when T is increased. Interestingly, the swelling rate occurs at a slower rate than loss of native contacts (see the inset in Figure 4B).

Hidden States during Folding and Unfolding. Folding simulations following a single jump in C show that there are two dominant pathways (Figure 2C). To quantify if there are additional hidden states that are sampled as hTR folds or unfolds we performed simulations by jumping the ion concentrations from a low (high) to a high (low) value and interrupting the folding (unfolding) at predetermined values of $\Delta\tau_F$. This strategy, which is easy to execute in simulations, is similar to the use of ion-pulses employed in a recent single-molecule study of RNase P RNA.³¹ Repeated cycling between high and low C carried out multiple times for a single molecule is equivalent to generating a swarm of trajectories and interrupting their folding at predetermined values of $\Delta\tau_F$. From a computational perspective it is efficient to generate a large number of trajectories and interrupt their folding at varying $\Delta\tau_F$ rather than subject one molecule to multiple ion-pulses. For ergodic systems the two methods are equivalent. To illustrate the effect of using ion pulses (Figure 1A) we carried out a few simulations by periodically altering the concentration of ions. An example of a trajectory, which reaches the folded state ($Q \approx 1$) in a single ion jump but unfolds subsequently (Figure 1A), illustrates the heterogeneity of hTR folding. In this example, after the first decrease in C the molecule remains unfolded even after about 4 cycles of increase and decrease in ion concentration even though it is folded after the first ion jump. Only in the final C -jump is the native state reached.

For each $\Delta\tau_F$, we generated a swarm of 100 trajectories from which the routes to the folded states could be determined in terms of variables that characterize the hTR pseudoknot. In Figure 5 we show the distributions, $P(Q, \Delta\tau_F)$, at various values of $\Delta\tau_F$. Unfolding distributions (Figure 5A) show that loss of structure occurs continuously with $P(Q, \Delta\tau_U)$ changing synchronously from the values at $t = 0$ to the equilibrium distribution function. As t progresses, the hTR samples Q regions with decreasing similarity to the native state. From the unfolding trajectory in Figure 5C it is clear that structure is lost continuously as the folded state is destabilized. In contrast, in the folding distribution functions, $P(Q, \Delta\tau_F)$, there is evidence that even as

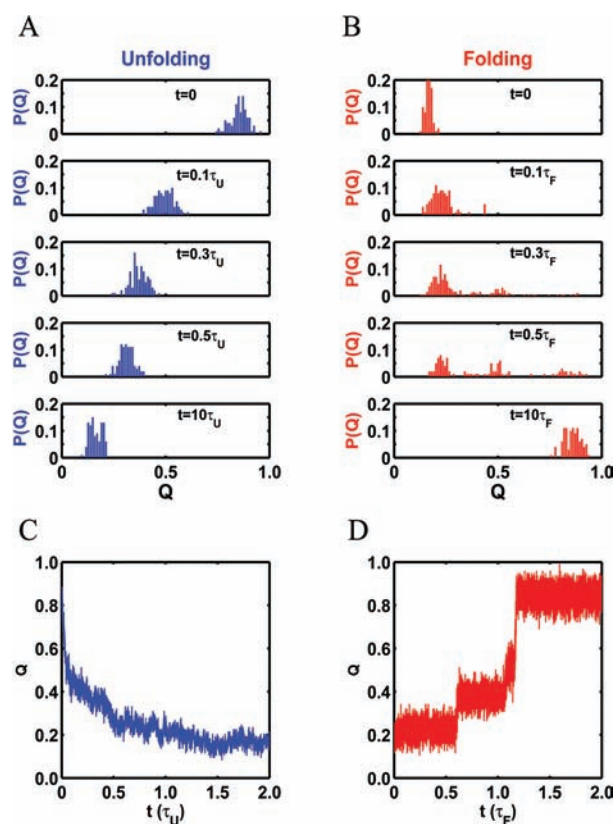


Figure 5. Approach to the folded and unfolded states. (A) On the left are shown the distribution of $P(Q, \Delta\tau_U)$ at various $\Delta\tau_U$ values (see Figure 1A) at which folding is interrupted following a decrease in C . (B) Same as in (A) except the distributions correspond to C -jump corresponding to folding simulations. (C) Decay of $Q(t)$ for a representative trajectory upon initiation of unfolding. (D) Same as in (C) except that this is a folding trajectory. Unfolding occurs in a single step (C), whereas the formation of the folded state takes place in discrete steps with great variations from trajectory to trajectory (see also top panel in Figure 1A).

t increases Q values that correspond to unfolded conformations are sampled (Figure 5B). We conclude that unfolding is not the reverse of folding under these conditions. The increase in the native content $Q(t)$ for a sample trajectory shows that the folded state is reached in steps. In contrast to unfolding the refolding process has greater heterogeneity. It should be pointed out that the nature of distribution functions could change as T is varied, and is likely to depend on the depth of jump, $(C_H - C_L)/(C_m)$ where C_H (C_L) are the high (low) ion concentrations and C_m is the ion concentration at which the folded and unfolded states coexist.

To explore the routes explored by hTR upon increasing or decreasing the ion concentration we calculated the distributions involving time-averaged values of the fraction of native contacts. We computed the distributions, $P_i(\bar{Q}, \Delta\tau_F = 0.1\tau_F)$ ($\tau_F \approx 3 \text{ ms}$), where i labels the trajectories, and

$$\bar{Q} = \frac{\int_0^{\Delta\tau_F} Q(s) ds}{\Delta\tau_F} \quad (6)$$

is the time average value of the fraction of native contacts over the interval $\Delta\tau_F$. The distributions for two folding trajectories following a jump in the ion concentration in Figure 6A are

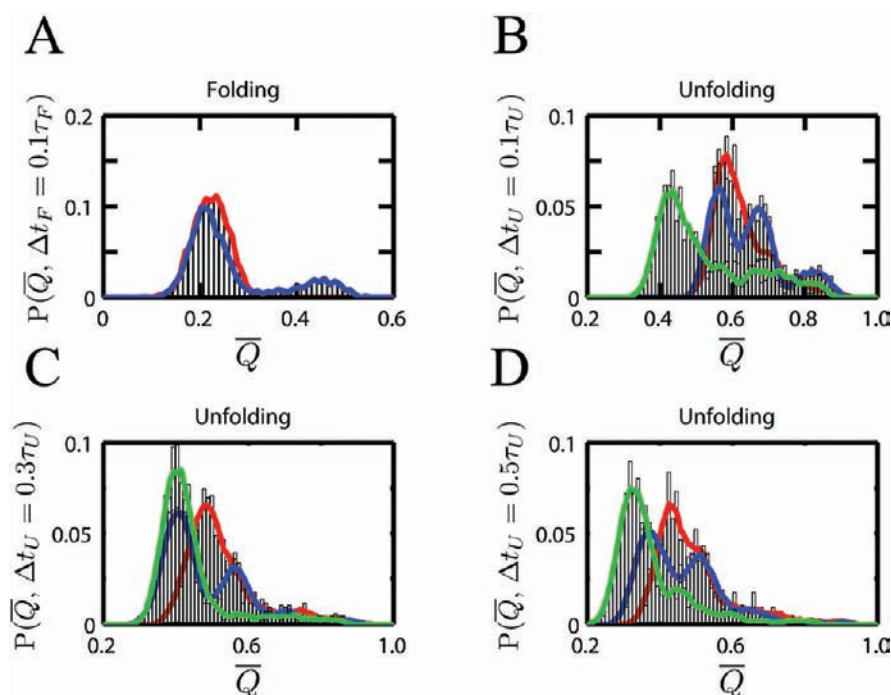


Figure 6. Folding and unfolding heterogeneities upon interrupted folding. (A) Distribution of \bar{Q} (eq 6) for two molecules ($\Delta\tau_F = 0.1\tau_F$) shortly after folding is initiated by C -jump. The ensemble of conformations partition into two basins shown in blue and red at early stages of folding. (B–D) Same as (A) except these correspond to distributions of \bar{Q} obtained at various $\Delta\tau_U$ values following C -increase. The solid colors are for four molecules.

similar. Not surprisingly, they differ dramatically from the equilibrium fluctuations under folding conditions because $\Delta\tau_F$ is not long enough to reach the folded state. Comparison of the distribution of Q under native conditions and the nonequilibrium $P(\bar{Q}, \Delta\tau_F = 0.1\tau_F)$ distributions in Figure 6A shows that native structures are not even partially formed when $\Delta\tau_F = 0.1\tau_F$. It should be emphasized that on the time scale $0.1\Delta\tau_F$ there is substantial reduction in R_g (Figure 3A). In other words, although increase in the ion concentration leads to the formation of collapsed structures, the folding reaction is incomplete.

The results for interrupted unfolding by decreasing the ion concentration (see Figure 1A) show that unfolding is highly heterogeneous for all values of $\Delta\tau_U$. Histograms $P(\bar{Q}, \Delta\tau_U)$ for $\Delta\tau_U = 0.1, 0.3$, and $0.5\tau_U$ ($\approx 3\mu\text{s}$) for 3–4 molecules are vastly different, which implies that at each $\Delta\tau_U$ there is considerable pathway diversity as hTR unfolds. During the unfolding process RNA goes from a low-entropy folded state to a high-entropy unfolded state. Thus, there are considerable fluctuations in the ensemble of unfolded states, which result in greater pathway diversity. Interestingly, for all values of $\Delta\tau_U$ and even $\Delta\tau_U = 0.5\tau_U$ hTR samples values of $\bar{Q} \approx 0.4$, which is larger than the average of $Q \approx 0.2$ at long times (see Figure 5A). Parts B–D in Figure (6), with peaks at multiple \bar{Q} values, show that even as $\Delta\tau_U$ is increased from $0.1\tau_U$ to $0.5\tau_U$ there are several routes through which flux is channeled from the folded to unfolded state. Despite the high entropy of the unfolded state we find that there are only a small (4–5) number of populated routes to the unfolded state. Indeed, the ensemble of unfolded structures can be grouped into a small number of clusters based on the Q values. The conformations in each cluster are structurally related.

Flux to the Folded State Involves Multiple Channels. As $(\Delta\tau_F)/(\tau_F)$ is increased beyond ~ 0.1 the conformations that are sampled as hTR folds can be partitioned into ~ 6 major clusters

(see Figure 7 for $\Delta\tau_F = 0.5\tau_F$). The distributions $P(\bar{Q}, \Delta\tau_F)$ in Figure 7 show the acquisition of native structure in each of the clusters. Somewhat surprisingly, even when $(\Delta\tau_F)/(\tau_F) \approx 1$ there is a substantial probability that RNA samples conformations with \bar{Q} that are closer to the unfolded state (see panels 1, 2, and 5 in Figure (7)). We find that the patterns in Figure 7 are similar at lower and higher values of $\Delta\tau_F$ except that the number of molecules in each cluster changes. In Figure 8 we show that various states are gradually populated as the folding reaction proceeds upon C -jump. Even at $t = 2\tau_F$ the folded state is not the most populated, which implies that equilibrium is only reached at times that far exceed τ_F . These observations show that the approach to the folded state under folding conditions, which occurs through multiple channels, involves interconversion among a small number of dominant states that are repeatedly visited as hTR continues to become natively like. Only when $(\Delta\tau_F)/(\tau_F)$ far exceeds unity do we find that the equilibrium distribution (last panel in Figure 5B) is reached.

DISCUSSION

The predicted phase diagram in the $[C, T]$ plane of the hTR can be understood as a transition between the folded and unfolded states as C is increased or T is lowered. From a thermodynamic perspective it appears that in the T and C range studied here hTR folds in an apparent two-state manner. Because of the simplicity of the model, which has allowed us to explore a wide range of conditions, the predicted phase diagram may not be quantitatively accurate. For example, the predicted melting temperatures are likely to be low. We note, however, that CPT³⁴ have shown that the melting temperature at high monovalent salt concentration is in good accord with experiments. Given that this genre of models have been remarkably successful in explaining and anticipating a range of scenarios for RNA folding we expect

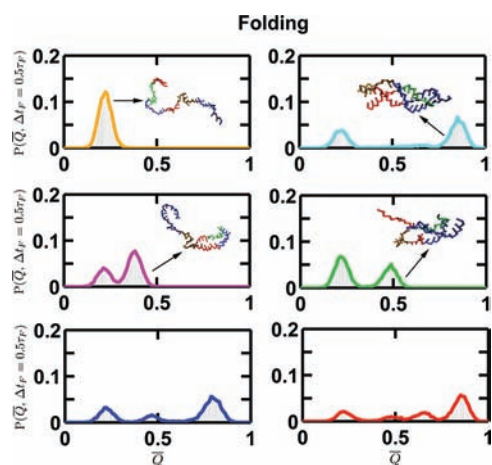


Figure 7. Hidden states during folding induced by C-jump. The ensemble of folded conformations at time $\Delta\tau_F = 0.5\tau_F$ at which folding is interrupted by partition into six distinct basins as assessed by distributions of \bar{Q} . Similar distributions are obtained at other values of $\Delta\tau_F > 0.1\tau_F$. The folded state is reached by interconversion among these basins as $\Delta\tau_F$ increases. Folding landscape of hTR can be viewed as a sparsely connected network of a small but distinct number of states.

that the qualitative predictions will hold well. In particular, the prediction that $T_m \approx C_m^\alpha$ is amenable to experimental test. The phase diagram will depend on the nature of counterions. In particular, it is possible that size and shape⁴⁹ can alter the predicted phase boundary. Capturing these important effects will require more refined treatments of electrostatic interactions⁵⁰ than attempted here.

The multistage compaction of RNA,²⁸ which was predicted using theory of polyelectrolyte collapse,⁵¹ is remarkably similar to recent time-resolved small angle X-ray scattering (SAXS) experiments on the longer *Azoarcus* ribozyme.⁵² Theoretical arguments predict that in general ion-induced compaction must occur in three stages with the first one being a very fast process representing substantial neutralization of the charged phosphate groups.²⁸ Typically, this process, which is perhaps diffusion controlled, is too fast to be resolved in SAXS experiments.⁵² Because the current version of the TIS model does not explicitly model ions the fast process of counterion condensation is not captured. In the second step, which occurs in about $\tau_c \approx 0.2$ ms for ion-jump, substantial compaction (greater than about 70%) occurs, just as observed in time-resolved SAXS in experiments. Surprisingly, the ratio $(\tau_c)/(\tau_f)$ for hTR in our simulations and *Azoarcus* ribozyme is similar,⁵² which reinforces the theoretical prediction that the collapse transition in the earliest stages of folding is a generic aspect of RNA folding driven primarily by nonspecific polyelectrolyte effects. Thus, τ_c should be largely determined by the number of nucleotides,^{28,51} which in turn determines the overall bare charge on RNA.

An interesting finding in our work is that the rate of hTR compaction under conditions used in the present simulations is greater when folding is initiated by ion-jump than by temperature-quench. Although demonstrated here in the context of hTR folding, this result is valid in general for the following reasons. It is now firmly established using theory and experiments that upon an increase in ion concentration the earliest event in folding is substantial neutralization (>80%, depending on the valence of counterions) of the charges on phosphate groups by nonspecific counterion condensation. This event drastically decreases the extent of repulsion between the phosphate groups, and poises

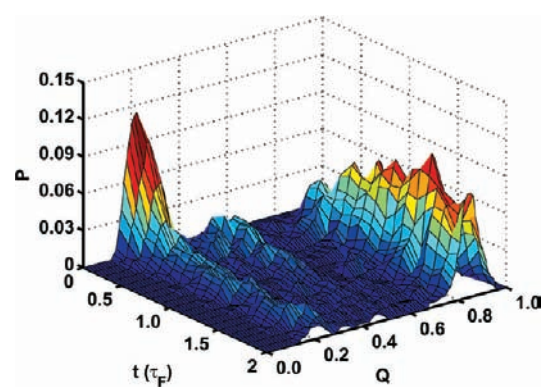


Figure 8. Dynamic picture of the folding landscape upon C-jump. Distribution of Q at various times expressed in terms of $P(Q,t)$ where time is measured in units of the folding time. The emergence of the various basins of attraction that are repeatedly sampled are revealed as peaks in $P(Q,t)$. Even when projected along a highly coarse-grained variable Q the folding landscape of hTR is rugged. A similar picture emerges when folding is initiated by T -quench as shown in the Table of Contents figure.

RNA to become compact on a time scale that is proportional to the number of nucleotides. In contrast, folding initiated by T -quench with high C does not lead to a large decrease in electrostatic repulsion because only the dielectric constant changes without altering the net charge on the phosphate groups. Thus, upon T -quench RNA folding is controlled to a greater extent by non-bonded short-range tertiary interactions, which become favorable only when RNA becomes compact enough for them to be effective. This key qualitative prediction can be experimentally checked by performing time-resolved SAXS experiments by varying both C and T . The phase diagram (Figure 2A) shows that the interplay between thermal fluctuations and electrostatic interactions is likely to be more subtle at intermediate values of T and C .

The existence of multiple channels through which flux to the folded state flows was first shown in the context of protein folding by Klimov and Thirumalai⁵³ and has subsequently been confirmed by others using different methods.⁵⁴ It is surprising that despite the complexity of the RNA folding landscape there are only a finite number of sparsely kinetically connected channels that carry flux to the native state. By sparse, we mean that the structures in the small number of clusters interconvert among each other, becoming increasingly native-like without creating (destroying) additional (existing) clusters. Thus, during the folding process the ensemble of unfolded hTR repeatedly visits only a finite number of states, which is consistent with conclusions reached on the basis of inherent structure calculations for a designed four-helix-bundle protein.⁵⁵ We should emphasize that these conclusions depend on the choice of the observable used in assessing the number of productive channels in the folding reaction. We suspect that the existence of multiple pathways through which RNA (or for that matter proteins) reach the NBA is independent of the reaction coordinate. More importantly, by manipulating external conditions (preincubating RNA in the presence of a modest ion concentration) one can change the number of pathways.

CONCLUSIONS

In order to dissect the complexity of RNA folding we have carried out extensive simulations of pseudoknot-forming hTR

using coarse-grained simulations. Coarse-grained models, which capture the essence of the folding problem,⁵⁶ have proven to be useful not only in interpreting experiments but also in anticipating their outcomes. Our work suggests that a new class of experiments in which folding is interrupted at various times when combined with CG simulations are powerful for monitoring the regions of the folding landscapes that cannot otherwise be easily explored. Although our results were obtained using the hTR pseudo knot as a case study, many of the conclusions are expected to be valid in general because they rely on the interplay of forces that govern folding of all RNA molecules. For example, the generic features of the phase diagram should be valid for other RNA molecules whose folding thermodynamics can be described using the two-state approximation. Many of the surprising findings, including the demonstration that a large fraction of compact structures could be far from being native and the demonstration that the folding mechanisms triggered by temperature quench and a jump in the ion concentration are different, warrant further scrutiny.

AUTHOR INFORMATION

Corresponding Author

thirum@umd.edu

Present Addresses

[#]Departments of Physics and Computer Science, Wake Forest University, Winston-Salem, North Carolina 27109, United States.

ACKNOWLEDGMENT

We appreciate the helpful comments by Shaon Chakraborty. We are grateful to the National Science Foundation (CHE 09-14033) for supporting this research. S.S.C. appreciates support from Wake Forest Science Research Fund.

REFERENCES

- (1) Treiber, D. K.; Williamson, J. R. *Curr. Opin. Struct. Biol.* **1999**, *9*, 339–345.
- (2) Chen, S. J.; Dill, K. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 646–651.
- (3) Thirumalai, D.; Hyeon, C. *Biochemistry* **2005**, *44*, 4957–4970.
- (4) Sosnick, T. R. *Protein Sci.* **2008**, *17*, 1308–1318.
- (5) Woodson, S. A. *Annu. Rev. Biophys.* **2010**, *39*, 61–77.
- (6) Chen, S.-J. *Annu. Rev. Biophys.* **2008**, *37*, 197–214.
- (7) Guo, Z. Y.; Thirumalai, D.; Honeycutt, J. D. *J. Chem. Phys.* **1992**, *97*, 525–535.
- (8) Goldstein, R. A.; Lutheyschulten, Z. A.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9029–9033.
- (9) Proctor, D. J.; Ma, H. R.; Kierzek, E.; Kierzek, R.; Gruebele, M.; Bevilacqua, P. C. *Biochemistry* **2004**, *43*, 14004–14014.
- (10) Sarkar, K.; Nguyen, D. A.; Gruebele, M. *RNA* **2010**, *16*, 2427–2434.
- (11) Ma, H. R.; Proctor, D. J.; Kierzek, E.; Kierzek, R.; Bevilacqua, P. C.; Gruebele, M. *J. Am. Chem. Soc.* **2006**, *128*, 1523–1530.
- (12) Hyeon, C.; Thirumalai, D. *J. Am. Chem. Soc.* **2008**, *130*, 1538–1539.
- (13) Zarrinkar, P. P.; Williamson, J. R. *Science* **1994**, *265*, 918–924.
- (14) Fang, X. W.; Pan, T.; Sosnick, T. R. *Nat. Struct. Biol.* **1999**, *6*, 1091–1095.
- (15) Pan, T.; Sosnick, T. R. *Nat. Struct. Biol.* **1997**, *4*, 931–938.
- (16) Pan, J.; Thirumalai, D.; Woodson, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *273*, 7–13.
- (17) Pan, J.; Thirumalai, D.; Woodson, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6149–6154.
- (18) Narayanan, R.; Velmurugu, Y.; Kuznetsov, S.; Ansari, A. *J. Am. Chem. Soc.* **2011**, *133*, 18767–18774.
- (19) Treiber, D. K.; Rook, M. S.; Zarrinkar, P. P.; Williamson, J. R. *Science* **1998**, *279*, 1943–1946.
- (20) Sclavi, B.; Sullivan, M.; Chance, M. R.; Brenowitz, M.; Woodson, S. A. *Science* **1998**, *279*, 1940–1943.
- (21) Hyeon, C.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6789–6794.
- (22) Lin, J. C.; Thirumalai, D. *J. Am. Chem. Soc.* **2008**, *130*, 14080–14081.
- (23) Feng, J.; Walter, N. G.; Brooks, C. L. *J. Am. Chem. Soc.* **2008**, *133*, 4196–4199.
- (24) Whitford, P. C.; Schug, A.; Saunders, J.; Hennelly, S. P.; Onuchic, J. N.; Sanbonmatsu, K. Y. *Biophys. J.* **2009**, *96*, L7–L9.
- (25) Thirumalai, D.; Woodson, S. A. *Acc. Chem. Res.* **1996**, *29*, 433–439.
- (26) Sosnick, T. R.; Pan, T. *Curr. Opin. Struct. Biol.* **2003**, *13*, 309–316.
- (27) Zhuang, X.; Bartley, L. E.; Babcock, H. P.; Russell, R.; Ha, T.; Herschlag, D.; Chu, S. *Science* **2000**, *288*, 2048–2051.
- (28) Thirumalai, D.; Lee, N.; Woodson, S. A.; Klimov, D. K. *Annu. Rev. Phys. Chem.* **2001**, *52*, 751–762.
- (29) Zhuang, X. W. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 399–414.
- (30) Rueda, D.; Bokinsky, G.; Rhodes, M. M.; Rust, M. J.; Zhuang, X. W.; Walter, N. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10066–10071.
- (31) Qu, X. H.; Smith, G. J.; Lee, K. T.; Sosnick, T. R.; Pan, T.; Scherer, N. F. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6602–6607.
- (32) Barsegov, V.; Thirumalai, D. *Phys. Rev. Lett.* **2005**, *95*, 168302–4.
- (33) Garcia-Manyes, S.; Dougan, L.; Badilla, C. L.; Brujic, J.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10534–10539.
- (34) Cho, S. S.; Pincus, D. L.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 17349–17354.
- (35) Walter, A. E.; Turner, D. H.; Kim, J.; Lyttle, M. H.; Muller, P.; Mathews, D. H.; Zuker, M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 9218–9222.
- (36) Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911–940.
- (37) Dima, R. I.; Hyeon, C.; Thirumalai, D. *J. Mol. Biol.* **2005**, *347*, 53–69.
- (38) Stammler, S. N.; Cao, S.; Chen, S.-J.; Giedroc, D. P. *RNA* **2011**, *17*, 1747–1759.
- (39) Honeycutt, J. D.; Thirumalai, D. *Biopolymers* **1992**, *32*, 695–709.
- (40) Veitshans, T.; Klimov, D.; Thirumalai, D. *Fold. Des.* **1997**, *2*, 1–22.
- (41) Cho, S.; Levy, Y.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 586–591.
- (42) Camacho, C. J.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6369–6372.
- (43) Theimer, C. A.; Blois, C. A.; Feigon, J. *Mol. Cell* **2005**, *17*, 671–682.
- (44) Theimer, C. A.; Finger, L. D.; Trantirek, L.; Feigon, J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 449–454.
- (45) Theimer, C. A.; Feigon, J. *Curr. Opin. Struct. Biol.* **2006**, *16*, 307–318.
- (46) Zuker, M. *Nucleic Acid Res.* **2003**, *31*, 3406–3415.
- (47) Klimov, D. K.; Thirumalai, D. *Folding Des.* **1998**, *3*, 127–139.
- (48) Cao, S.; Chen, S.-J. *J. Mol. Biol.* **2007**, *367*, 909–924.
- (49) Koculi, E.; Hyeon, C.; Thirumalai, D.; Woodson, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *129*, 2676–2682.
- (50) Tan, Z.-J.; Chen, S.-J. *Biophys. J.* **2011**, *101*, 176–187.
- (51) Lee, N.; Thirumalai, D. *J. Chem. Phys.* **2000**, *113*, 5126–5129.
- (52) Roh, J. H.; Guo, L. A.; Kilburn, J. D.; Briber, R. M.; Irving, T.; Woodson, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *132*, 10148–10154.
- (53) Klimov, D. K.; Thirumalai, D. *J. Mol. Biol.* **2005**, *353*, 1171–1186.
- (54) Noe, F.; Schuetter, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19011–19016.
- (55) Guo, Z.; Thirumalai, D. *J. Mol. Biol.* **1996**, *263*, 323–343.
- (56) Hyeon, C.; Thirumalai, D. *Nat. Commun.* **2011**, *2*, 487.