doi:10.1016/S0022-2836(03)00854-4 J. Mol. Biol. (2003) 332, 311–319

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

COMMUNICATION

The Fastest Global Events in RNA Folding: Electrostatic Relaxation and Tertiary Collapse of the Tetrahymena Ribozyme

Rhiju Das[1,2], Lisa W. Kwok[3], Ian S. Millett[4], Yu Bai[2,5], Thalia T. Mills[3] Jaby Jacob[6,7], Gregory S. Maskel[3], Soenke Seifert[8] Simon G. J. Mochrie[9], P. Thiyagarajan[6]*, Sebastian Doniach[1]* Lois Pollack[3]* and Daniel Herschlag[2,4,5]*

[1]*Department of Physics, Stanford University, Stanford, CA 94305-4060, USA*

[2]*Department of Biochemistry School of Medicine, Beckman Center Room B400, Stanford University, Stanford, CA 94305-5307, USA*

[3]*School of Applied and Engineering Physics, Cornell University, Ithaca NY 14853, USA*

[4]*Department of Chemistry Stanford University, Stanford, CA 94305, USA*

[5]*Biophysics Program, Stanford University, Stanford, CA 94305 USA*

[6]*Intense Pulsed Neutron Source Argonne National Laboratory 9700 South Cass Avenue, Argonne IL 60439, USA*

[7]*Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637,USA*

[8]*Chemistry Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439 USA*

[9]*Department of Physics, Yale University, New Haven, CT 06520 USA*

**Corresponding authors*

Large RNAs can collapse into compact conformations well before the stable formation of the tertiary contacts that define their final folds. This study identifies likely physical mechanisms driving these early compaction events in RNA folding. We have employed time-resolved small-angle X-ray scattering to monitor the fastest global shape changes of the Tetrahymena ribozyme under different ionic conditions and with RNA mutations that remove long-range tertiary contacts. A partial collapse in each of the folding time-courses occurs within tens of milliseconds with either monovalent or divalent cations. Combined with comparison to predictions from structural models, this observation suggests a relaxation of the RNA to a more compact but denatured conformational ensemble in response to enhanced electrostatic screening at higher ionic concentrations. Further, the results provide evidence against counterion-correlation-mediated attraction between RNA double helices, a recently proposed model for early collapse. A previous study revealed a second 100 ms phase of collapse to a globular state. Surprisingly, we find that progression to this second early folding intermediate requires RNA sequence motifs that eventually mediate native long-range tertiary interactions, even though these regions of the RNA were observed to be solvent-accessible in previous footprinting studies under similar conditions. These results help delineate an analogy between the early conformational changes in RNA folding and the "burst phase" changes and molten globule formation in protein folding.

Keywords: RNA folding; time-resolved small-angle X-ray scattering; molten globule; electrostatic relaxation; burst phase

Supplementary data associated with this article can be found at doi: 10.1016/S0022-2836(03)00854-4

Abbreviations used: trSAXS, time-resolved small-angle X-ray scattering.

E-mail addresses of the corresponding authors: thiyaga@anl.gov; lois@ccmr.cornell.edu; doniach@drizzle.stanford.edu; herschla@cmgm.stanford.edu

RNA molecules fold into specific, intricate structures to carry out numerous biological functions.[1] A deep understanding of RNA folding is thus a prerequisite for explaining the workings of RNA in current biology and throughout evolution, especially given RNA's likely role as the dominant

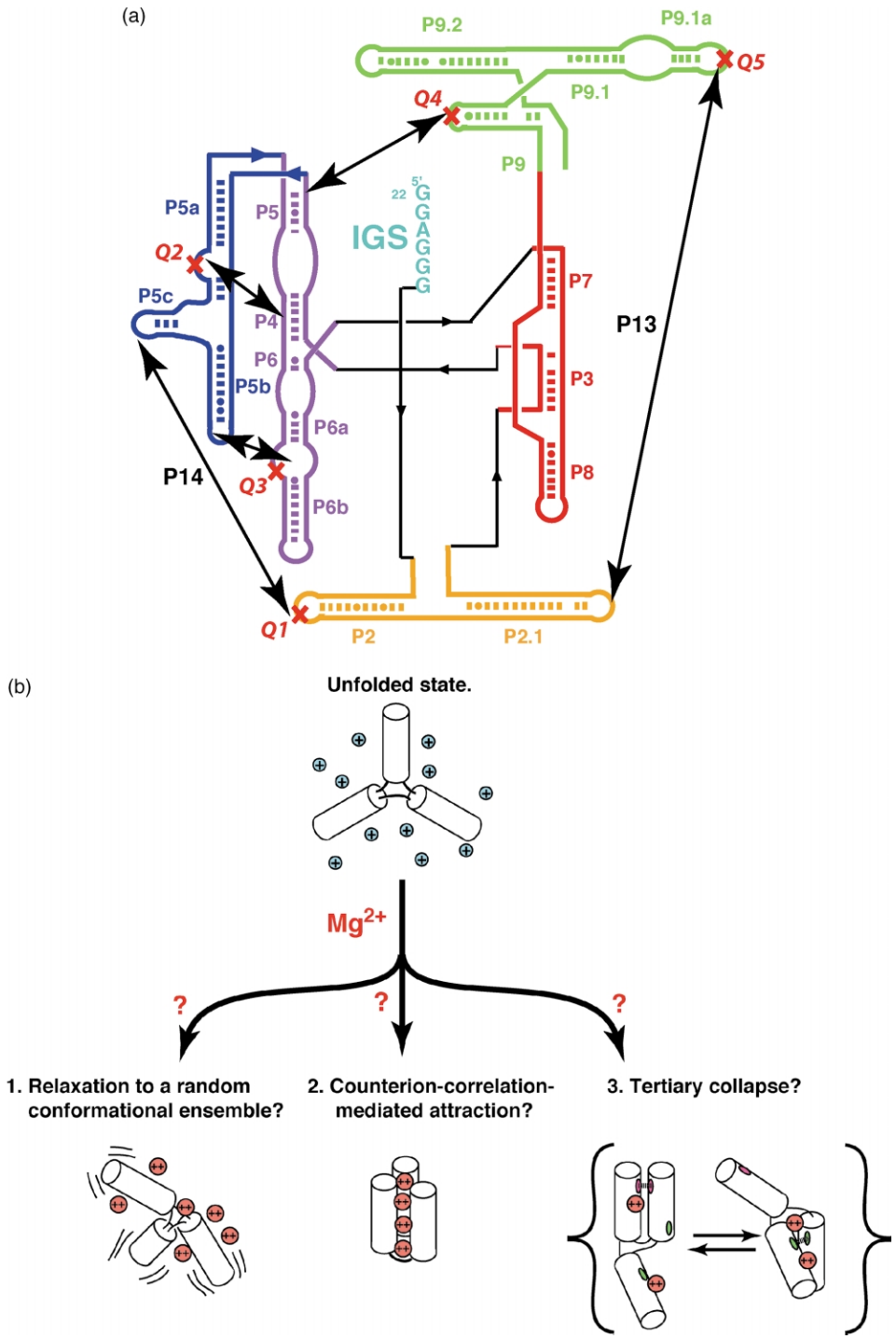

Figure 1. (a) The Tetrahymena ribozyme. Native secondary structure of the Tetrahymena ribozyme,[46] with arrows connecting sequences identified as making long-range tertiary contacts. The ribozyme was prepared in the L-21 *Sca* I form[47] by *in vitro* transcription and column purification as described.[48] Locations of sequence changes made in the quintuple-knockout mutant are depicted as crosses†. Folding of this mutant is highly destabilized, with much lower catalytic activity than the wild-type ribozyme (relative second-order rate constant $[k_{cat}/K_M]^G < 10^{-5}$ for the reaction E·CCCUCUAAAAA + G → E·CCCUCU + GAAAAA with 10 mM or 50 mM Mg^{2+} at 25 °C, assayed as described[49]), and little or no protection from hydroxyl radicals at 10 mM and 100 mM Mg^{2+} relative to no Mg^{2+} and compared to the wild-type ribozyme (R.D. & D.H., unpublished results). (b) Cartoon illustrations of the physical mechanisms for RNA collapse tested in this study, with cylinders representing negatively charged RNA double helices.

† The mutations in the quintuple-knockout molecule, implemented with the Stratagene QuikChange protocol, are the following: *Q1* (L2 → UNCG), GGCAUGCACCU (39–49) to CUUCGGU; *Q2* (A-rich-bulge → U), AAUAAG (183–188) to UUUUUU; *Q3* (J6a/b → base-pairs), UAAG (224–227) to AUA; *Q4* (L9 → UNCG), AAU (324–326) to C; *Q5* (L9.1 → UNCG), GAUGCAAC (346–353) to CUUCGG.

biomolecule in the early stages of life.[2] Furthermore, comparison of folding mechanisms of two rather different biopolymers, RNA and proteins, will likely illuminate general physical principles of biopolymer folding. In recent years, a bounty of solved structures of natural and man-made RNAs[3] as well as biophysical elucidation of folding pathways of large ribozymes[4] have advanced a description of RNA folding as a quasi-hierarchical process, often hindered by kinetically stable misfolded intermediates. Nevertheless, from a physical point of view, an outstanding challenge remains: to characterize the fundamental forces that guide RNAs along their folding pathways to their native structures.

A model RNA studied by several laboratories is the group I ribozyme from *Tetrahymena thermophila*, the best characterized large RNA with respect to catalytic function[5] and three-dimensional structure.[6–8] The native fold of this ~400 nucleotide ribozyme consists of several double-helical segments whose global arrangement is mediated by sequence-specific tertiary contacts between single-stranded loops and bulges (Figure 1(a)).[6–8] The ribozyme's conformation and catalysis are intimately coupled to metal ions,[9,10] with folding experiments usually initiated by the addition of divalent magnesium ions.[4] The majority of cations bound to the ribozyme are not localized to specific RNA sites but instead form a highly dynamic counterion "atmosphere",[11] whose properties are currently under intense theoretical scrutiny[12,13] and are also a strong focus of the present experimental study.

Recent discoveries, permitted by advances in time-resolved small-angle X-ray scattering (trSAXS), have uncovered a new physical puzzle in RNA folding. Upon addition of Mg^{2+}, the ribozyme rapidly compacts from an extended state with pre-formed double helices to a globular state of nearly native dimensions well before formation of any stable tertiary contacts.[14–16] A similar result has been reported for another large RNA, the catalytic domain of the ribonuclease P RNA from *Bacillus subtilis*.[17] The studies with the Tetrahymena ribozyme were able to resolve two phases of collapse, with time constants of ~10 ms and ~100 ms, both shorter than the one second time constant measured for the first protections from solution hydroxyl radical cleavage.[18] Ionic and temperature conditions of the SAXS and hydroxyl radical footprinting experiments were similar though not identical†; we have thus adopted a working model for this study wherein this early collapse occurs without concomitant solvent protections.[15]

At first glance, these collapsed RNA states might be considered analogous to the molten globule states that appear in the early folding stages of numerous proteins, stabilized by hydrophobic interactions.[19] However, in a large RNA, the nonpolar faces of the nucleotide bases are already largely buried into double helices prior to compaction, excluding a generic RNA "hydrophobic collapse". In further contrast to protein collapse, the massive negative charge of the RNA backbone imposes a strong electrostatic penalty to compaction. What, then, is the physical origin of rapid RNA compaction?

We present initial tests of three classes of physical mechanisms for the early collapse of a large RNA before the apparent formation of solvent-protected tertiary contacts;[15] each model is illustrated in cartoon form in Figure 1(b). The models, which are not necessarily exclusive of one another, are described below, along with their respective predictions that are tested herein.

Model 1. Electrostatic relaxation. The counterion atmosphere of Mg^{2+} will more effectively screen the strong Coulomb repulsion that pushes apart RNA segments than the diffuse atmosphere assembled from the low concentrations of monovalent ions in the starting solution.[13,20] The resulting screening would allow thermal fluctuations to relax the initially extended RNA molecule into a more compact but still unstructured conformational ensemble. According to this model, high concentrations of monovalent ions are predicted to induce an electrostatic relaxation similar to that observed with much lower concentrations of multivalent ions.[13,20]

Model 2. Counterion-correlation-mediated attraction. Sophisticated theoretical work predicts that multivalent ions condensed around double helices can induce a net attractive force between helices, thereby providing a natural explanation for RNA collapse.[12,21–27] Simplified model simulations of strongly charged helices show counterions ordering into lattices between the helix charges at very low temperature;[24] the inter-helical attractive force may be regarded as due to the salt-bridges each counterion makes with neighboring phosphate groups on separate RNA segments. At physiological temperatures, the attractive force is induced by correlations between thermally fluctuating counterion clouds that allow favorable electrostatic interactions of individual cations with multiple helices while minimizing cation–cation

† SAXS and footprinting studies have been conducted at slightly different temperatures (42 °C and 15–37 °C, respectively) and ionic conditions ($[Na^+] = 10$ mM and 20 mM, respectively; $[Mg^{2+}] = 10$ mM in both studies). Differences in temperature do not appear to affect compaction timescales, based on indistinguishable second phase time constants at 15 °C, 25 °C, and 37 °C in this and a previous SAXS study.[15] Effects of monovalent ionic differences, however, remain to be tested. Increasing monovalent salt concentration to higher concentrations has been shown to affect initial states and folding rates of several large RNAs.[15,16,50–53] In particular, a recent study of Tetrahymena ribozyme folding at the much higher Na^+ concentration of 200 mM displayed detectable protections on timescales of 100 ms and less, although none of these protections was as strong as in the native state.[53]

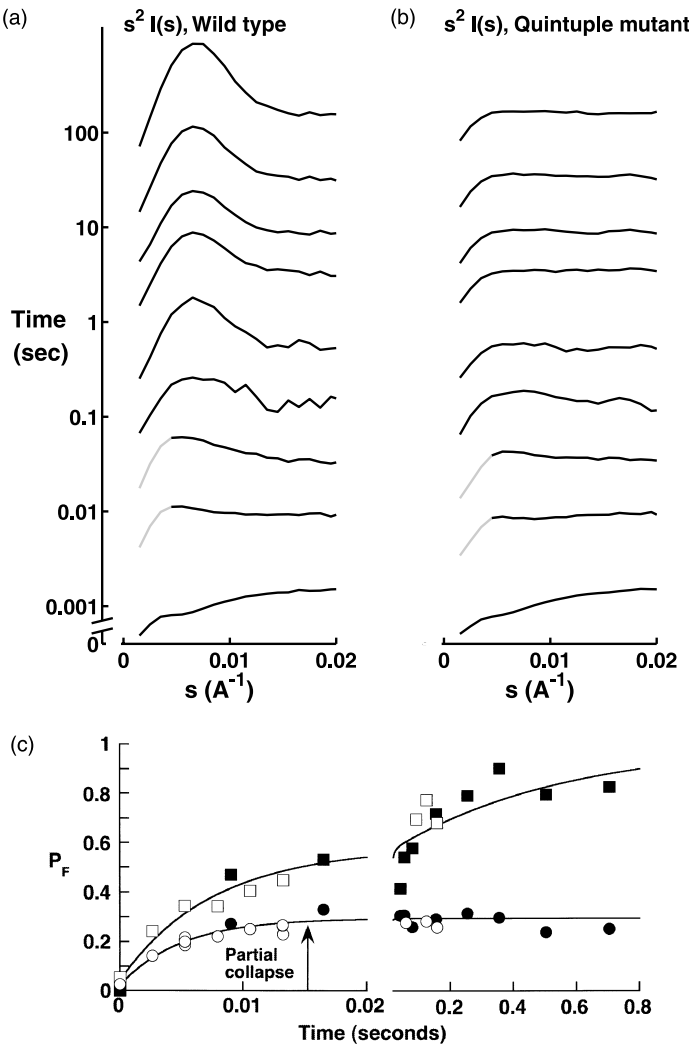

Figure 2. Global shape changes during folding of the Tetrahymena ribozyme initiated by the addition of Mg^{2+}. SAXS profiles obtained for (a) the wild-type molecule and (b) the tertiary-contact-knockout mutant are displayed as Kratky plots ($s^2I(s)$ *versus* s, where $s = 2 \times \sin(\theta/2)/\lambda$, with θ, the scattering angle, and λ, the X-ray wavelength), a format in which more compact globular states exhibit stronger low-scattering-angle peaks.[30] Selected timepoints are shown; all measured profiles are given in the Supplementary Material. Folding experiments were performed at 25 °C, with 50 mM Na–Mops (pH 7.0) (20 mM Na^+), and with added ions as noted. Timepoints from 15 ms to ten minutes were acquired in a Bio-Logic SFM 400 stopped-flow mixer at beamline 12-ID (BESSRC CAT) at the Advanced Photon Source (APS), mixing equal volumes of a 4 mg/ml (30 μM) RNA solution and a folding buffer with 30 mM $MgCl_2$ ($[Mg^{2+}]_{final,total} = 15$ mM; $[Mg^{2+}]_{final,free} \sim 12$ mM).[15] Very fast timepoints, 2–150 ms, were acquired at beamline 8-ID at the APS with microfabricated continuous-flow mixers by flowing 4 mg/ml of RNA through the mixer inlet channel to mix with excess solution containing 10 mM $MgCl_2$.[15] Repeated measurements with lower RNA concentrations and analysis of scattering intensity extrapolated to zero scattering angle indicate that aggregation is negligible for timepoints faster than one second (see the Supplementary Material). The lowest scattering angles attained in the stopped-flow and continuous-flow measurements were not the same; to aid visual comparison, the low s region of the latter profiles have been extended (gray curves) with a quadratic Guinier extrapolation (see the Supplementary Material). (c) Two-component fits of time-courses. As in a previous study, singular value decomposition reveals that all profiles can be described within noise as the linear combination of two basis curves.[15] However, if the scattering profile of a physical state happens to be a linear combination of the other two, such a decomposition is consistent with the presence of three (or more) physical states.[15] Each measured Kratky profile was projected onto two basis profiles acquired for the unfolded (U) and folded (F) wild-type ribozyme. The fractional weights of F in the projections (P_F) for the wild-type ribozyme (open and filled squares) and the quintuple-knockout mutant (open and filled circles) can be fit most simply by a double-exponential for the wild-type ($P_F = a_1(1 - e^{-t/\tau_1}) + a_2(1 - e^{-t/\tau_2})$; $a_1 = 0.39(\pm 0.05)$, $\tau_1 = 3(\pm 1)$ ms, $a_2 = 0.43(\pm 0.05)$, $\tau_2 = 100(\pm 30)$ ms) and by a single-exponential for the mutant ($a_1 = 0.28(\pm 0.01)$, $\tau_1 = 4(\pm 1)$ ms)†. Again, the ability to fit two phases with two basis profiles in no way implied the absence of a third state. Filled symbols are from stopped-flow measurements, and open symbols are from continuous-flow measurements; the data sets from the different mixers demonstrate good agreement for overlapping timepoints. Radius-of-gyration (R_g) estimates, based on Guinier fits of the data at the lowest s ranges, yield results in qualitative agreement with two-component fits, but have larger errors (see the Supplementary Material).

† Very rapid collapse of a sub-domain with a time constant faster than 10 ms in the wild-type ribozyme may account for the slightly larger fitted amplitude of the partial-collapse phase in the wild-type time-course compared to the tertiary-contact-knockout mutant. Indeed, time-resolved footprinting data on the P5abc sub-domain, both isolated and incorporated into the P4–P6 domain, suggest that P5abc collapses with a time constant of <30 ms.[16]

repulsions.[12,21–27] This effect can be considered a classical analog of the van der Waals attraction mediated by quantum correlations between the electron clouds of atoms.[23] This "counterion-correlation-mediated attraction" is a strong candidate for the driving force that condenses very long double-stranded DNAs into toroidal aggregates on the addition of multivalent cations.[26] However, the magnitude of such an attractive force has not

been quantified for large RNAs, whose flexibility at bulges and junctions compared to continuous DNA double helices may allow the interaction of multiple helical segments and may increase the attractive force.[15,28] Monovalent ions alone are predicted to display insignificant correlations and to therefore not induce this attraction.[22–26]

Model 3. Tertiary collapse. Different RNA molecules are stabilized in different compact states by different subsets of tertiary contacts, promoted by Mg^{2+}, but the involved RNA elements are buried in a sufficiently small fraction of the overall population so as to not give detectable protection from solution hydroxyl radicals in footprinting experiments.[15,16] Such contacts may be native or non-native; we present an initial test here of a specific version of this model in which the contacts necessary for tertiary collapse are mediated by the same loops and bulges that are involved in native long-range tertiary contacts. In this specific model for compaction, a mutant unable to form any of the native long-range tertiary contacts would be predicted to not undergo collapse upon addition of Mg^{2+}.

The predictions from each of these mechanisms for global RNA compaction have been tested by obtaining trSAXS folding time-courses initiated with different ionic conditions for the wild-type Tetrahymena ribozyme and for a "quintuple-knockout" mutant RNA unable to form the long-range tertiary contacts critical for stabilizing the native structure (see Figure 1(a)).

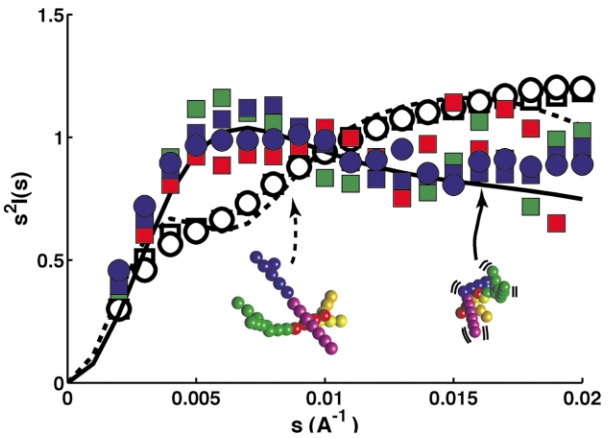

Figure 3. SAXS profiles of unfolded and electrostatically relaxed RNA states. Kratky plots are shown for the unfolded states of the wild-type and quintuple mutant ribozymes (open squares and circles, respectively; nearly indistinguishable) and for early timepoints (15 ms) of the time-courses of the wild-type ribozyme (blue squares) and the quintuple-knockout mutant (blue circles) in 10 mM Mg^{2+}, and of wild-type time-courses in 1 M Na^+ (green squares) and 1 M $N(CH_3)_4^+$ (red squares). Experimental profiles are compared to theoretical curves predicted from a coarse-grained model of the ribozyme with double helices extended away from each other (broken line) and with helices allowed to sample random conformations (continuous line).[15] Kratky profiles have been normalized to the same integrated area to aid visual comparison.

Testing models for collapse

Early partial collapse

We first discuss a striking feature that arose in all of the RNA folding time-courses herein: a rapid compaction to a partially collapsed state that occurred on a timescale of 10 ms. Figure 2(a) displays trSAXS folding time-courses for the wild-type Tetrahymena ribozyme and a quintuple-knockout mutant of the ribozyme, each initiated by the addition of 10 mM Mg^{2+}. The wild-type and mutant molecules both undergo initial compaction events with time constants that are the same within error ($3(\pm 1)$ ms and $4(\pm 1)$ ms; Figure 2(c)) to states with similar global structures (Figure 3, blue squares and circles), providing no indication that long-range tertiary interactions are required for this process (model 3). In addition, counterion-correlation-mediated attraction between RNA helices (model 2) is predicted to be negligible in the presence of monovalent ions alone,[22–26,29] especially larger ions.[24] However, partial collapse of the RNA is induced by high concentrations of Na^+ and the bulkier $N(CH_3)_4^+$ (Figure 3, green and red squares, respectively), suggesting that this novel physical force is not the explanation for the early partial collapse.

In contrast, the data are consistent with the earliest RNA partial collapse being an electrostatic relaxation of the initially extended RNA structure (model 1) due to enhanced charge screening. In the initial conditions of our experiments (20 mM Na^+), Coulomb repulsion between RNA double helices will not be screened effectively by the monovalent atmosphere.[13,20] The initial RNA structure should therefore consist of double helices splayed away from each other. Such a structural model is in agreement with the SAXS profile of the unfolded state[14,15] (Figure 3, broken line and open symbols). Addition of either multivalent ions or ~1 M monovalent ions is expected to effectively screen inter-helical repulsion, thereby allowing the initial conformation to relax to a more dynamic conformational ensemble.[13,20] The observed 15 ms timepoints from each of these time-courses are indeed consistent with the SAXS profile predicted for a random conformational ensemble (Figure 3, continuous line). There is thus no support for an attractive force between RNA helical segments, and a simple electrostatic relaxation can account for the initial partial collapse events.

Second phase of collapse

Our previous studies also identified a second phase of ribozyme collapse to a globular state with a time constant of ~100 ms.[15] This second phase can be seen in the Kratky plots of Figure 2(a) and the fit in Figure 2(c) (squares). However,

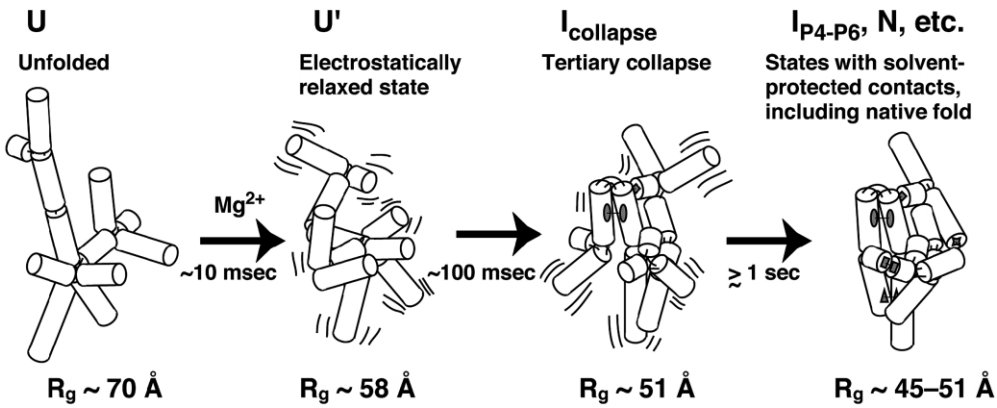

Figure 4. Physical model of the rapid compaction pathway of the Tetrahymena ribozyme. Gray shapes on cylinders represent long-range tertiary contacts between RNA double helices. Estimated radius-of-gyration values (see the Supplementary Material) are given below each state.

this phase is absent in the folding pathway of the quintuple-knockout mutant (Figure 2(b) and (c)). This result implies that the second phase is promoted by at least one RNA sequence motif associated with a native long-range tertiary contact (model 3), even though all of these contacts appear to be solvent accessible on the 100 ms timescale in hydroxyl radical footprinting experiments carried out under similar conditions (see the footnote on page 313).[15,18]

Proposed pathway for rapid compaction in a large RNA

The experiments herein identify electrostatic screening and tertiary collapse as the likely physical mechanisms involved in the earliest global folding events of the Tetrahymena ribozyme. These conclusions are summarized in the schematic pathway shown in Figure 4†. The generic nature of these mechanisms suggests that our results may be relevant to the early folding pathways and collapsed intermediates of other large structured RNAs.

Rapid partial collapse through electrostatic relaxation

Upon addition of 10 mM Mg^{2+}, enhanced screening of helix–helix Coulomb repulsion by divalent ions relaxes the ribozyme from the initial extended state with double helices splayed away from each other (Figure 4(a)) to a partially collapsed self-avoiding conformational ensemble more appropriate to the new ionic conditions (Figure 4(b)). The conformational change does not require any native long-range tertiary contacts‡. It is therefore likely to be an obligatory step in the *in vitro* folding of all large RNAs, when initiated by addition of physiological concentrations of divalent ions to initial "artificial" solution conditions of low ionic strength.

This RNA relaxation appears analogous to the sub-millisecond conformational changes seen in the folding of several proteins initiated by dilution of chemical denaturants. Such "burst phase" protein changes have been proposed to be a readjustment of the chemically denatured conformational ensemble (a random coil state) to a more compact, denatured state appropriate to the new solvent condition, although the extent of structure formation in such protein states is currently under extensive debate.[30–34] Regardless,

† A previous study[15] noted two possible compaction pathways in 10 mM Mg^{2+}: the sequential model shown in Figure 4 and a scenario in which two initial populations of the unfolded ribozyme that do not interconvert on the experimental timescale undergo compaction with different time constants of ~10 ms and ~100 ms. This second model predicts that the "partial-collapse" SAXS profile at the intermediate time of ~30 ms corresponds to a mixed population of a compact globule state and an unstable unfolded state, instead of the single population predicted to occur in the sequential model. However, the quintuple-knockout mutant collapses to a stable state in 10 mM Mg^{2+} (Figure 2(b) and (c)) with a profile indistinguishable from the wild-type partially collapsed state. Furthermore, identical partially collapsed SAXS profiles are obtained for the quintuple-knockout mutant in 100 mM Mg^{2+}, high concentrations (1 M) of monovalent ions, and low concentrations (1–10 mM) of $Co(NH_3)_6^{3+}$ and the polyamines spermidine^{3+} and spermine^{4+} (results not shown). These results most simply suggest that the partially collapsed state observed for the wild-type ribozyme corresponds to a single population, favoring the sequential collapse through a relaxed state (Figure 4) over the alternative scenario with two parallel populations.

‡ A question remains of whether all the RNA helix–helix junctions in this relaxed ensemble are truly random or whether some or all of the junctions maintain a local structure. Such locally folded junctions might be expected to give significant protections to hydroxyl radical cleavage,[54] which may be detectable in high-precision footprinting experiments with improved band fitting methods.[53]

the RNA relaxation contrasts with the protein readjustment in several ways: (i) the likely physical mechanism, electrostatic screening *versus* hydrophobic collapse; (ii) the presence of pre-formed secondary structure in the "unfolded" RNA state but not in most proteins; and (iii) the expected change in the number of accessible global conformations, with an increase in RNA flexibility after relaxation *versus* a decrease in protein conformational freedom after collapse.

Despite the apparent simplicity of the electrostatic relaxation mechanism, the timescale of RNA collapse (10 ms) appears to be much slower than Mg^{2+} association or any fundamental "speed limit" for conformational changes of nucleic acids. For example, formation of model hairpins from single-stranded DNA has been timed at 1–10 μs[35] and folding of a three-way RNA junction at 100 μs[36] under similar ionic conditions. We therefore suggest that the junctions connecting the helices of unfolded RNAs in low ionic strength solutions are not freely mobile loops. Rather, these junctions may be structured, stabilized by hydrogen bonds and base-stacking interactions that are compatible with the RNA's electrostatically extended conformation. The breaking of these initial interactions then slows the rate of the earliest global conformation changes of the ribozyme to a few milliseconds, rather than tens of microseconds, or less. The frustration of even these earliest global changes in a large RNA is consistent with the general tendency of RNA to promiscuously form stable, non-native structures observed in numerous experiments probing different steps of RNA folding[4,37–41] and in strong contrast to protein motion.

We anticipate that future investigations of both the global and local structure of this RNA relaxed folding intermediate will be greatly simplified by the isolation here of a thermodynamically stable analog of this state, the tertiary-contact-knockout mutant, whose folding progresses no further than the relaxed state (Figure 3).

Tertiary collapse

After electrostatic relaxation, the ribozyme collapses on a 100 ms timescale to a globular state (Figure 4(c)), which rearranges on slower timescales to form stable tertiary structure (Figure 4(d)), as monitored by protections of the tertiary contacts from solution radicals.[14,15,18] A generic attractive inter-helical force, such as a counterion-correlation-mediated attraction,[12,21–26] is unlikely to stabilize the 100 ms state, since such a globular state is not attained by the quintuple-knockout mutant. Rather, sequence motifs required for the formation of native long-range RNA tertiary contacts are found here to be critical for formation of the 100 ms state.

Such sequence motifs, despite their importance in this 100 ms "tertiary collapse", appear to be unprotected from solvent in hydroxyl radical footprinting studies with similar, but not identical, temperature and ionic conditions (see the footnote on page 313).[15,18] Further studies, probing Tetrahymena ribozyme folding by both footprinting and trSAXS under identical conditions and with additional mutant forms of the ribozyme, will be necessary to test our working model that collapse occurs without concomitant protection from solvent. According to this model, tertiary collapse leads to an ensemble of compact ribozyme conformations, in which tertiary contact formation is unsynchronized or transient, so that no individual long-range tertiary contact is buried from solvent in the majority of the conformations. If such sub-states do not interconvert at subsequent times, ribozyme folding would best be described as involving a large number of pathways. If the compact sub-states do interchange quickly, it would be reasonable to denote this intermediate ensemble an RNA molten globule state, just as conformational ensembles of proteins with near-native compactness but without stable tertiary structure have been termed molten globules.[19,42,43] The contacts stabilizing such an RNA molten globule state may well be native, in analogy with findings of native-like folding topology in protein molten globules.[44,45]

In summary, our results lead to three major conclusions.

(1) The fastest phase of ribozyme collapse (3–4 ms) is induced not only by the divalent ion Mg^{2+} but also by Na^+ and $N(CH_3)_4^+$, implying that counterion-correlation-mediated attraction (model 2) does not explain this faster phase.

(2) The fastest phase of ribozyme collapse (3–4 ms) occurs in a highly destabilized quintuple-knockout mutant as well as in the wild-type ribozyme. Therefore, long-range tertiary contacts (model 3) are not required for this phase. In combination with (1), above, and structural modeling, electrostatic relaxation (model 1) emerges as the favored explanation for the fastest phase of ribozyme compaction.

(3) The quintuple-knockout mutant does not display the subsequent phase of collapse (~100 ms) that is observed for the wild-type ribozyme, suggesting that at least one of the sequence motifs involved in native long-range tertiary contacts is required for this phase. Footprinting studies suggest that no region of the ribozyme is strongly protected from solvent after this tertiary collapse, leading to a working model of the tertiary collapsed state as an ensemble of compact intermediates, either formed along multiple folding pathways or comprising an RNA molten globule state.

Combining global shape information from SAXS and information about individual nucleotides from hydroxyl radical footprinting, obtained for Tetrahymena ribozyme constructs with individually mutated tertiary contacts, promises to further illuminate the structures and underlying energetics of these earliest intermediates in RNA folding.

Acknowledgements

We thank Harold Gibson, Kazuki Ito, Guy Jennings, Larry Lurio, Tobin Sosnick, Gil Toombes, and Hiro Tsuruta for experimental assistance; and Michael Brenowitz, Mark Engelhardt, Rick Russell, and the members of the Herschlag laboratory for helpful discussions. We acknowledge funding from NASA (NAG8-1178 to L.P.), the NSF (to L.P. through Cornell's Nanobiotechnology Center; and graduate fellowship to R.D.), the NIH (GM49243 to D.H.; training grant GM08267 to L.W.K.; NS40132 to I.S.M. & S.D.), the Packard Foundation Interdisciplinary Science Program (99-8327 to P.T.), Stanford Graduate Fellowships (Y.B. and R.D.), and the US DOE (under contract no. W-31-109-Eng-38 to the APS, beamlines 8-ID and 12-ID, University of Chicago, and IPNS). These experiments made use of the Cornell Nanofabrication Facility, supported by the NSF, Cornell University, and the Facility's industrial affiliates. Initial static SAXS measurements on the quintuple-knockout mutant were carried out at beamline 4–2 of the Stanford Synchrotron Radiation Laboratory, supported by the NIH and the DOE.

References

1. Moore, P. B. (1999). The RNA folding problem. In *The RNA World* (Atkins, J. F., ed.), 2nd edit., pp. 381–401, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Joyce, G. F. & Orgel, L. E. (1999). Prospects for understanding the origin of the RNA world. In *The RNA World* (Gesteland, R. F., Cech, T. R. & Atkins, J. F., eds), 2nd edit., pp. 49–77, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
3. Doudna, J. A. & Cech, T. R. (2002). The chemical repertoire of natural ribozymes. *Nature*, **418**, 222–228.
4. Treiber, D. K. & Williamson, J. R. (2001). Beyond kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **11**, 309–314.
5. Narlikar, G. J. & Herschlag, D. (1997). Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes. *Annu. Rev. Biochem.* **66**, 19–59.
6. Lehnert, V., Jaeger, L., Michel, F. & Westhof, E. (1996). New loop–loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem. Biol.* **3**, 993–1009.
7. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E. *et al.* (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.
8. Golden, B. L., Gooding, A. R., Podell, E. R. & Cech, T. R. (1998). A preorganized active site in the crystal structure of the Tetrahymena ribozyme. *Science*, **282**, 259–264.
9. Shan, S., Kravchuk, A. V., Piccirilli, J. A. & Herschlag, D. (2001). Defining the catalytic metal ion interactions in the Tetrahymena ribozyme reaction. *Biochemistry*, **40**, 5161–5171.
10. Feig, A. L. & Uhlenbeck, O. C. (1999). The role of metal ions in RNA biochemistry. In *The RNA World* (Gesteland, R. F., Cech, T. R. & Atkins, J. F., eds), 2nd edit., pp. 287–302, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
11. Anderson, C. F. M. T. & Record, J. (1990). Ion distributions around DNA and other cylindrical polyions: theoretical descriptions and physical implications. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 423–465.
12. Grosberg, A. Y., Nguyen, T. T. & Shklovskii, B. I. (2002). The physics of charge inversion in chemical and biological systems. *Rev. Mod. Phys.* **74**, 329–345.
13. Misra, V. K. & Draper, D. E. (2002). The linkage between magnesium binding and RNA folding. *J. Mol. Biol.* **317**, 507–521.
14. Russell, R., Millett, I. S., Doniach, S. & Herschlag, D. (2000). Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nature Struct. Biol.* **7**, 367–370.
15. Russell, R., Millett, I. S., Tate, M. W., Kwok, L. W. B., Nakatani, B., Gruner, S. M. *et al.* (2002). Rapid compaction during RNA folding. *Proc. Natl Acad. Sci. USA*, **99**, 4266–4271.
16. Deras, M. L., Brenowitz, M., Ralston, C. Y., Chance, M. R. & Woodson, S. A. (2000). Folding mechanism of the Tetrahymena ribozyme P4-P6 domain. *Biochemistry*, **39**, 10975–10985.
17. Fang, X. W., Thiyagarajan, P., Sosnick, T. R. & Pan, T. (2002). The rate-limiting step in the folding of a large ribozyme without kinetic traps. *Proc. Natl Acad. Sci. USA*, **99**, 8518–8523.
18. Sclavi, B., Sullivan, M., Chance, M. R., Brenowitz, M. & Woodson, S. A. (1998). RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. *Science*, **279**, 1940–1943.
19. Chamberlain, A. K. & Marqusee, S. (2000). Comparison of equilibrium and kinetic approaches for determining protein folding mechanisms. *Advan. Protein Chem.* **53**, 283–328.
20. Heilman-Miller, S. L., Thirumalai, D. & Woodson, S. A. (2001). Role of counterion condensation in folding of the Tetrahymena ribozyme. I. Equilibrium stabilization by cations. *J. Mol. Biol.* **306**, 1157–1166.
21. Murthy, V. L. & Rose, G. D. (2000). Is counterion delocalization responsible for collapse in RNA folding? *Biochemistry*, **39**, 14365–14370.
22. Arenzon, J. J., Stilck, J. F. & Levin, Y. (1999). Simple model for attraction between like-charged polyions. *Eur. Phys. J. ser. B*, **12**, 79–82.
23. Ha, B.-Y. & Liu, A. J. (1997). Counterion-mediated attraction between two like-charged rods. *Phys. Rev. Letters*, **79**, 1289–1292.
24. Grønbech-Jensen, N., Mashl, R. J., Bruinsma, R. F. & Gelbart, W. M. (1997). Counterion-induced attraction between rigid polyelectrolytes. *Phys. Rev. Letters*, **78**, 2477–2480.
25. Lee, N. & Thirumalai, D. (1999). Kinetics of condensation of flexible polyelectrolytes in poor solvents: effects of solvent quality, valence and size of counterions. http://arXiv.org/cond-mat/9907199
26. Bloomfield, V. A. (1997). DNA condensation by multivalent cations. *Biopolymers*, **44**, 269–282.
27. Oosawa, F. (1968). Interaction between parallel rodlike macroions. *Biopolymers*, **6**, 1633–1647.
28. Podgornik, R. & Parsegian, V. A. (1998). Charge-fluctuation forces between rodlike polyelectrolytes: pairwise summability reexamined. *Phys. Rev. Letters*, **80**, 1560–1563.
29. Khan, M. O. & Jönsson, B. (1999). Electrostatic correlations fold DNA. *Biopolymers*, **39**, 121–125.

30. Doniach, S. (2001). Changes in biomolecular conformation seen by small angle X-ray scattering. *Chem. Rev.* **101**, 1763.
31. Krantz, B. A., Mayne, L., Rumbley, J., Englander, S. W. & Sosnick, T. R. (2002). Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *J. Mol. Biol.* **324**, 359–371.
32. Mayor, U., Guydosh, N. R., Johnson, C. M., Grossmann, J. G., Sato, S., Jas, G. S. *et al.* (2003). The complete folding pathway of a protein from nanoseconds to microseconds. *Nature*, **421**, 863–867.
33. Akiyama, S., Takahashi, S., Kimura, T., Ishimori, K., Morishima, I., Nishikawa, Y. & Fujisawa, T. (2002). Conformational landscape of cytochrome c folding studied by microsecond-resolved small-angle X-ray scattering. *Proc. Natl Acad. Sci. USA*, **99**, 1329–1334.
34. Segel, D. J., Bachmann, A., Hofrichter, J., Hodgson, K. O., Doniach, S. & Kiefhaber, T. (1999). Characterization of transient intermediates in lysozyme folding with time-resolved small-angle X-ray scattering. *J. Mol. Biol.* **288**, 489–499.
35. Goddard, N. L., Bonnet, G., Krichevsky, O. & Libchaber, A. (2000). Sequence dependent rigidity of single stranded DNA. *Phys. Rev. Letters*, **85**, 2400–2403.
36. Kim, H. D., Nienhaus, G. U., Ha, T., Orr, J. W., Williamson, J. R. & Chu, S. (2002). Mg^{2+}-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules. *Proc. Natl Acad. Sci. USA*, **99**, 4284–4289.
37. Bartley, L. E., Zhuang, X., Das, R., Chu, S. & Herschlag, D. (2003). Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. *J. Mol. Biol.* **328**, 1011–1026.
38. Herschlag, D. (1995). RNA chaperones and the RNA folding problem. *J. Biol. Chem.* **270**, 20871–20874.
39. Uhlenbeck, O. C. (1995). Keeping RNA happy. *RNA*, **1**, 4–6.
40. Gartland, W. J. & Sueoka, N. (1966). Two interconvertible forms of tryptophanyl sRNA in *E. coli*. *Proc. Natl Acad. Sci. USA*, **55**, 948–956.
41. Lindahl, T., Adams, A. & Fresco, J. R. (1966). Renaturation of transfer ribonucleic acids through site binding of magnesium. *Proc. Natl Acad. Sci. USA*, **55**, 941–948.
42. Ohgushi, M. & Wada, A. (1983). Molten-globule state: a compact form of globular proteins with mobile side-chains. *FEBS Letters*, **164**, 21–24.
43. Arai, M. & Kuwajima, K. (2000). Role of the molten globule state in protein folding. *Advan. Protein Chem.* **53**, 209–282.
44. Gillespie, J. R. & Shortle, D. (1997). Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.* **268**, 170–184.
45. Peng, Z. Y., Wu, L. C. & Kim, P. S. (1995). Local structural preferences in the alpha-lactalbumin molten globule. *Biochemistry*, **34**, 3248–3252.
46. Cech, T. R., Damberger, S. H. & Gutell, R. R. (1994). Representation of the secondary and tertiary structure of group I introns. *Nature Struct. Biol.* **1**, 273–280.
47. Zaug, A. J., Grosshans, C. A. & Cech, T. R. (1988). Sequence-specific endoribonuclease activity of the Tetrahymena ribozyme: enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme–substrate complexes. *Biochemistry*, **27**, 8924–8931.
48. Russell, R. & Herschlag, D. (1999). New pathways in folding of the Tetrahymena group I RNA enzyme. *J. Mol. Biol.* **291**, 1155–1167.
49. Karbstein, K., Carroll, K. S. & Herschlag, D. (2002). Probing the Tetrahymena group I ribozyme reaction in both directions. *Biochemistry*, **41**, 11171–11183.
50. Silverman, S. K., Deras, M. L., Woodson, S. A., Scaringe, S. A. & Cech, T. R. (2000). Multiple folding pathways for the P4-P6 RNA domain. *Biochemistry*, **39**, 12465–12475.
51. Russell, R., Zhuang, X., Babcock, H. P., Millett, I. S., Doniach, S., Chu, S. & Herschlag, D. (2002). Exploring the folding landscape of a structured RNA. *Proc. Natl Acad. Sci. USA*, **99**, 155–160.
52. Takamoto, K., He, Q., Morris, S., Chance, M. R. & Brenowitz, M. (2002). Monovalent cations mediate formation of native tertiary structure of the *Tetrahymena thermophila* ribozyme. *Nature Struct. Biol.* **9**, 928–933.
53. Uchida, T., Takamoto, K., He, Q., Chance, M. R. & Brenowitz, M. (2003). Multiple monovalent ion-dependent pathways for the folding of the L-21 *Tetrahymena thermophila* ribozyme. *J. Mol. Biol.* **328**, 463–478.
54. Churchill, M. E., Tullius, T. D., Kallenbach, N. R. & Seeman, N. C. (1988). A Holliday recombination intermediate is twofold symmetric. *Proc. Natl Acad. Sci. USA*, **85**, 4653–4656.

Edited by J. Doudna

(Received 28 March 2003; received in revised form 27 June 2003; accepted 30 June 2003)

Supplementary Material is available on Science Direct