Chapter 2
Theory of RNA Folding: From Hairpins to Ribozymes

D. Thirumalai(✉) and Changbong Hyeon

Abstract The rugged nature of the RNA folding landscape is determined by a number of conflicting interactions like repulsive electrostatic potential between the charges on the phosphate groups, constraints due to loop entropy, base stacking, and hydrogen bonding that operate on various length scales. As a result the kinetics of self-assembly of RNA is complex, but can be easily modulated by varying the concentrations, sizes, and shapes of the counterions. Here, we provide a theoretical description of RNA folding that is rooted in the energy landscape perspective and polyelectrolyte theory. A consequence of the rugged folding landscape is that, self-assembly of RNA into compact three-dimensional structures occurs by parallel routes, and is best described by the kinetic partitioning mechanism (KPM). According to KPM one fraction of molecules (Φ) folds rapidly while the remaining gets trapped in one of several competing basins of attraction. The partition factor Φ can be altered by point mutations as well as by changing the initial conditions such as ion concentration, size and valence of ions. We show that even hairpin formation, either by temperature or force quench, captures much of the features of folding of large RNA molecules. Despite the complexity of the folding process, we show that the KPM concepts from polyelectrolyte theory, and charge density of ions can be used to explain the stability, pathways and their diversity, and the plasticity of the transition state ensemble of RNA self-assembly.

2.1 Introduction

The landmark discovery that RNA molecules are ribozymes (RNA enzymes) (Guerriertakada et al. 1983; Kruger et al. 1982) has triggered an intense effort to decipher their folding mechanisms. In the intervening years an increasing repertoire of cellular functions has been associated with RNA (Doudna and Cech 2002). These

D. Thirumalai
Department of Chemistry and Biochemistry, University of Maryland, College Park, College Park, MD 20742, USA
e-mail: thirum@umd.edu

N.G. Walter et al. (eds.) *Non-Protein Coding RNAs*

doi: 10.1007/978-3-540-70840-7_2,

include their role in replication, translational regulation, viral propagation etc. Moreover, interactions of RNA with each other and with DNA and proteins are vital in many biological processes. Even, the central chemical activity of ribosomes, namely, the formation of the peptide bond in the biosynthesis of polypeptide chains by ribosomes near the peptidyl transfer center, involves only RNA, leading many to suggest that *ribosomes are ribozymes* (Nissen et al. 2000; Yusupov et al. 2001). The appreciation that RNA molecules play a major role in a number of cellular functions has made it important to establish the structure – function relationship. Thus, the need to understand, at the molecular level the ribozyme activity, inevitably leads to the question: How do RNA molecules fold?

In little over a decade great success has been achieved in an attempt to answer this question because of progress on a number of fronts. The number of experimentally determined high resolution RNA structures (Ban et al. 2000; Cate et al. 1996; Nissen et al. 2000; Yusupov et al. 2001) continues to increase which has enabled us to understand the interactions that stabilize the folded states. Single molecule (Ma et al. 2006; Onoa et al. 2003; Russell et al. 2002b; Woodside et al. 2006; Zhuang et al. 2000) and ensemble experiments (Zarrinkar and Williamson 1994; Koculi et al. 2006; Pan et al. 1999) using a variety of biophysical methods combined with theoretical techniques (Thirumalai and Woodson 1996; Thirumalai and Hyeon 2005) have led to a conceptual framework for predicting various processes by which RNA molecules fold.

There are two aspects to RNA folding. The first is the prediction of the folded structures from sequence (Hofacker 2003; Zuker and Stiegler 1981). The second problem concerns the mechanisms by which assembly of the three dimensional functionally competent structure forms, start from the unfolded conformations. In this chapter we describe the folding mechanisms from the energy landscape perspective with focus on the polyelectrolyte aspects of RNA.

At a first glance it might appear that the RNA folding problem should be simple at least in comparison to the better investigated problem of protein folding (Tinoco and Bustamante 1999). However, there are several reasons why RNA folding is a difficult problem.

1. The building blocks of RNA are the four nucleotides each with a base, ribose, and phosphate groups. The bases (two purines and two pyrimidines), that are chemically similar, interact with each other either through hydrogen bonding or base stacking. The secondary structural elements (helices, loops, bulges) are independently stable which gives the impression that the three dimensional assembly is built much the same way as complicated architecture using prefabricated building blocks. However, the difficulty arises not only because of the chemical similarity of the nucleotides but also due to the polyelectrolyte nature arising from the charged phosphate groups.
2. The bases, their ability to form hydrogen bonds through Watson–Crick (WC) pairing withstanding, are all hydrophobic. The uniformity of the hydrophilic backbone along with lack of diversity in the bases make RNA closer to a "homopolymer" than polypeptide chains (Thirumalai and Hyeon 2005). The "homopolymer" nature of nucleic acids results in RNA structures being able to adopt alternate structures i.e., the stability gap between the folded and the other

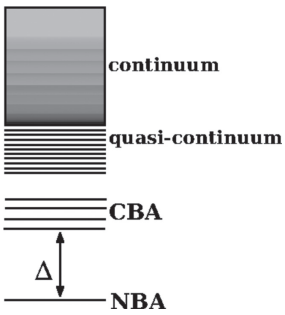

Fig. 2.1 View of the states of RNA as a free energy spectrum. The conformations in the NBA are separated from those in the competing basins of attraction (CBA) by the stability gap Δ. The structures in the CBA, while misfolded, can have many native-like features. Rapid folding without long pauses in the CBAs is likely if $\Delta/k_BT >> 1$. Figure adapted from (Guo et al. 1992)

misfolded structures is not large (Fig. 2.1). As a result, the energy landscape of RNA, even at the secondary structural level, is rugged containing many metastable conformations that serve as kinetic traps.

3. At some level, WC base pairing does simplify the prediction of RNA secondary structures. However, not all nucleotides are engaged in WC base pairing. Analysis of RNA secondary structures shows that the number of base-pairs (N_{BP}) varies with sequence length N as $N_{BP} = 0.27 \times N$. The linear growth of N_{BP} with N with slope 0.5 is expected if all the nucleotides are engaged in Watson–Crick base pairings. However, the slope is only 0.27 (Dima et al. 2005). This shows that 46% of the sequence, which is computed using $N_{BP}/N \approx (1-x)/2$, constitute non-pairing regions such as bulges, loops, dangling ends, and other motifs. The bulges and loops are important structural elements that glue the independent helices together to make the RNA structures compact.
4. Finally, the folding mechanisms can be greatly altered by changing the nature of counterions which makes it necessary to consider explicitly the polyelectrolyte nature of RNA. In particular, the important role of valence, shape and size of the counterions (Koculi et al. 2004, 2006, 2007) in modulating the secondary structures and possibly altering them during the course of tertiary structure formation, are difficult to predict (Chauhan and Woodson 2008; Thirumalai 1998; Wu and Tinoco 1998). The varying flexibilities of different regions of RNA, the homopolymer character of the building blocks, the key role of counterions in the folding process, and the presence of alternate structures render RNA folding a challenging problem.

2.2 Structural Characteristics of RNA

Determination of the size, shape, flexibility, and base-pairs statistics in RNA native structures, is important in understanding the nature of packing in folded structures and also in elucidating interaction between RNA and DNA or proteins. Analysis of

the RNA native structures available in the Protein Data Bank (PDB) can be used to infer the general characteristics of the shapes and flexibility of folded RNA.

Native Structures are Compact: If RNA structures are compact then their volumes are expected to scale as $V \sim R_G^3 \sim a^3 N$, where R_G is the radius of gyration, a is an effective monomer length. More generally, Flory showed that $R_G \sim aN^{\nu}$ where the Flory exponent $\nu = 1/3$ for maximally compact structures, $\nu = 1/2$ for polymers in Θ-condition, and $\nu = 3/5$ for flexible polymers in good solvents. As RNA is a polyelectrolyte valence, shape, and concentration (C) of counterions can alter solvent quality, and hence R_G. At low C, RNA is expanded and the transition to a compact structure occurs only when C exceeds the midpoint of the unfolded to folded transition.

Computation of the sizes of RNA structures using the PDB coordinates reveals that R_G, follows the Flory scaling law, namely, $R_G = a_N N^{1/3}$ Å (Hyeon et al. 2006). The pre-factor, a_N = 5.5 Å, corresponds approximately to the average distance between the phosphate groups (≈5.8 Å) along the ribose-phosphate backbone. For a given N, the approximate volume of RNA is larger than that of proteins whose R_G scales as $R_G = 3.1\ N^{1/3}$ Å (Dima and Thirumalai 2004; Hyeon et al. 2006). In other words, RNA molecules are more loosely packed than proteins, which are probably linked to their folding being dependent on accommodation of counterions to form compact structures. The difference is due to the larger size of the nucleotides compared to amino acids and the nature of interactions that stabilize the folded states of RNA and proteins.

Folded RNAs are Prolate Ellipsoids: Even though folded RNAs are compact, as assessed by R_G, substantial deviations from sphericity have been found. When the shape of RNA molecules is characterized by the asphericity Δ and the shape parameters S that are computed using the eigenvalues of the moment of inertia tensor (Aronovitz and Nelson 1986; Hyeon et al. 2006), we find that a large fraction of folded RNA structures are aspherical and the distribution of S values shows that RNA molecules are prolate. The prolate ellipsoid shape of RNA renders their diffusion intrinsically anisotropic. The observed difference between shapes of RNAs and globular proteins is primarily due to the nature of interactions that stabilize the folded structures of RNA and proteins. Packing in RNA is not only determined by the favorable interactions between nucleotides but also by counter-ion mediated long-range interactions. The volume excluded by counterions affects packing, and consequently the shape of RNA structures.

Persistence Length of RNA shows Similarity to Polyelectrolytes. From the polymer perspective, flexibility of RNA is best assessed by its persistence length, l_p, and its dependence on the changes in ionic strength. The overall compact RNA structure is formed by gluing together flexible (loops and bulges) and stiff helical regions. Despite the potential variations in the flexibility it is useful to obtain estimates of the global l_p. The total persistence length of RNA may be written as $l_p = l_p^0 + l_p^{el}$ where l_p^0 is the intrinsic persistence length and l_p^{el} is the electrostatic contribution. If RNA were a polyelectrolyte then $l_p^{el} = l_B/4\kappa^2 A^2$ where the Bjerrum length $l_B = e^2/4\pi\varepsilon k_B T$ (e is the unit of charge, ε is the dielectric constant, k_B is the Boltzmann constant, and T is the temperature), for monovalent couterions $\kappa^{-2} = 8\pi l_B I$ (I is the ionic strength), and A is the average distance between the charges (Odijk 1977; Skolnick and Fixman

1977). The l_p values can be obtained from the distance distribution functions, which, for folded RNA molecules, can be directly computed using the PDB coordinates. The persistence length of the folded RNA can be extracted by fitting, for $r/R_G > 1$, the distance distribution function $P(r)$, which is computed using the coordinates of the folded RNA, to the wormlike chain model $P_{WLC}(r) \sim \exp\{-1/(1-(l_p r/R_G^2)^2)\}$ (Caliskan et al. 2005; Hyeon et al. 2006). The persistence length is scale-dependent and varies as $l_p = 1.5\, N^{0.33}$ Å (Hyeon et al. 2006). The dependence of l_p on N implies that the average length of helices with stacks should increases as N grows.

In principle, as the counterion concentration decreases the changes in l_p can be secured by obtaining $P(r)$ using Small Angle X-ray Scattering (SAXS) experiments. To date, SAXS data is available for only a few RNA molecules (*Azoarcus* ribozyme (Rangan et al. 2004), RNase P (Fang et al. 2002), and *Tetrahymena* ribozyme (Russell et al. 2002a)). Surprisingly, analysis of $P(r)$ for *Azoarcus* ribozyme and RNase *P* showed that the distance distribution function is well fit using $P_{WLC}(r)$ for the WLC model. As the concentration of Mg^{2+} and Na^{+} decreases l_p increases (Caliskan et al. 2005) for *Azoarcus* ribozyme, $l_p \sim 21$ Å in the unfolded state, and $l_p \sim 10$ Å in the compact folded state. It is noteworthy that $l_p\, \kappa^{-2}$ which is predicted for polyelectrolytes (Odijk 1977; Skolnick and Fixman 1977) do not have globally compact folds like RNA molecules. Thus, not only does l_p change dramatically as RNA folds, but it also exhibits the characteristics of polyelectrolytes especially at low ionic strength. Thus, how the polyelectrolyte problem is solved in RNA remains a key problem.

2.3 Rugged Folding Landscape and the Kinetic Partitioning Mechanism

The observed multiple folding routes and the associated heterogeneity of folding pathways can be anticipated from the energy landscape perspective (Thirumalai and Woodson 1996). The states for RNA (or for proteins for that matter) can be represented as a free energy spectrum (Guo et al. 1992). If the free energy gap (Δ in Fig. 2.1) is large, then trapping in one of the many Competing Basins of Attraction (CBAs) is not very probable. The presence of many alternate structures implies that the stability gap (especially when scaled by N) for RNA is not very large. As a result, RNA folding landscape is rugged (Fig. 2.2a), and is characterized by the presence of multiple minima that are separated by free energy barriers of varying heights.

The rugged nature of the energy landscape arises due to the presence of several competing interactions. Favorable hydrophobic stacking, and tertiary interactions favor chain compaction while the negatively charged interactions are better accommodated by extended structures. As a result RNA molecules are "frustrated" because not all interactions involving a given nucleotide can be simultaneously satisfied. In addition, the polyelectrolyte nature of RNA also induces topological frustration. The formation of stable secondary structures is largely driven by interactions on "local" scales in which the persistence length is comparable to the Debye screening

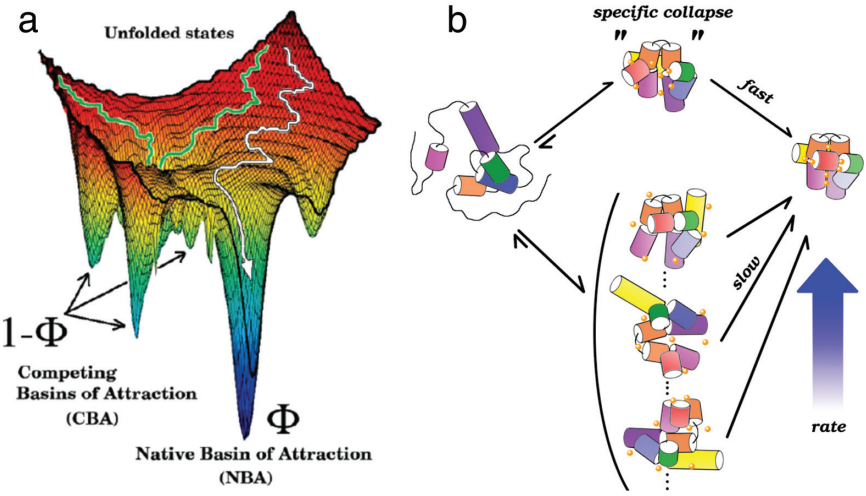

Fig. 2.2 (**a**) Schematic sketch of the rugged folding landscape of RNA. Conformational entropy and electrostatic repulsion between the phosphate groups favor the high free energy unfolded structures at low ionic strength. Under folding conditions a fraction of molecules (Φ) reach the NBA directly. A sketch of a trajectory for a fast track molecule that starts in a region of the energy landscape and which connects directly to the NBA is given in white. Trajectories (shown in green) that begin in other regions of the energy landscape can be kinetically trapped in the CBAs with probability ($1-\Phi$). The low dimensional representation of the complex energy landscape suggests that the initial conditions, which can be changed by counterions, stretching force, or denaturants, can alter the folding pathways. (**b**) Representation of RNA folding by KPM. Based on theory it is suggested that the fast track molecules specifically collapse into near native-like structures that rearrange to the native state without being trapped in the CBA. In contrast, the slow track molecules collapse to one of the manifold of misfolded structures. The collapse time scale, that depends on the nature of ions, for fast and slow track molecules, is similar. A spectrum of rates determine the transition from the CBAs to NBA (See figure insert for colour reproduction)

length. Compact folded structures result from the packing of locally formed secondary structures. Because there are multiple ways of assembling the stable secondary structures, several misfolded compact tertiary structures can form readily.

The incompatibility of the metastable misfolded structures that may share many of the correct secondary structures and the global stable fold, result in topological frustration. The folded structure may be thought of as the least frustrated and hence the most stable. From the perspective of topological frustration it follows that even the secondary structures can rearrange in the course of forming the global fold as was demonstrated in the context of P5abc formation (Wu and Tinoco 1998). In other words, organization of tertiary interactions might force the correct formation of even the secondary structures, as illustrated sometime ago using P5abc and more recently in the case of tertiary structure formation of a self-splicing group I intron in *Azoarcus* pre-tRNA (Chauhan and Woodson 2008).

The kinetic consequence of the rugged energy landscape is that folding is greatly impeded by long pausing in the CBAs. The structures in the CBAs could have many

native-like features that make them long-lived under folding conditions. The diversity in the folding trajectories that leads to the kinetic partitioning mechanism (KPM) is best illustrated using the sketch of the energy landscape (Fig. 2.2b). Under folding conditions (excess Mg^{2+}) the heterogeneous population of unfolded molecules navigates the rugged energy landscape in search of the NBA (Fig. 2.2a). A fraction (Φ) of unfolded molecules reaches the NBA rapidly without being trapped in any of the CBAs (Fig. 2.2b). The precise value of Φ depends on the sequence as well as external conditions, and is an indicator of the size of the NBA that in turn is determined by the extent to which a given sequence under specific ionic condition is frustrated. The remaining fraction, $(1-\Phi)$, gets kinetically trapped in one of the many CBAs. The transitions from the CBAs to the NBA might require large conformational changes, and hence involve overcoming substantial free energy barriers. Consequently, the transition rate CBA → NBA might be extremely slow depending on the extent of structural rearrangement required to reach the folded state. Because there are many kinetic metastable states, several rate constants are needed to fully describe the CBA → NBA transition. Thus, with the multivalley structure of the free energy landscape, the initial ensemble of molecules kinetically partition into fast folders (Φ being their fraction) and slow folders. From the KPM it follows that the fraction of molecules that reach the NBA at time t is $f_{NBA} = 1-\Phi \exp(-k_F t)-\Sigma a_i \exp(-k_i t)$ where k_F is the rate of reaching the NBA from the unfolded conformations for the fast folders, k_i is the rate of transition from the ith CBA to the NBA, and a_i is the corresponding amplitude.

Experimental Evidence. In key experiments, Zarrinkar and Williamson showed that the slow folding of *Tetrahymena* ribozyme is due to the presence of multiple long-lived metastable intermediates (Zarrinkar and Williamson 1994). This ribozyme, which has become the workhorse of group I intron folding, is roughly made up of two subdomains containing paired (P) regions P4–P6 and P3–P7 (Fig. 2.3). Using kinetics of oligonucleotide hybridization, two discrete intermediates along the presumed hierarchical folding pathway was identified. One of them is $\mathbf{I_1}$ (folded P4–P6) and the other is $\mathbf{I_2}$ in which both the major subdomains are nearly formed. Thus, in this picture, RNA folds through well-defined intermediates some of which are dependent on Mg^{2+}. The rate-limiting step is the association of the two major subdomains.

The possibility that $\Phi < 1$ implies that folding of RNA, regardless of the complexity of the fold, must occur by parallel pathways as predicted by KPM. The key prediction of KPM is that there must be a direct pathway from Unfolded Basin of Attraction (UBA) to the NBA. The evidence that *Tetrahymena* ribozyme folds by KPM was first provided by Pan et al. using a combination of theory and experiments (Pan et al. 2000; Thirumalai and Woodson 1996). Using native gel assay to measure the time-dependent increase in the population of the NBA under folding conditions and theoretical estimates for the rate of fast track molecules it was shown that $\Phi \approx 0.08$ for the precursor RNA. Thus, about 8% of the initially unfolded molecules reach the NBA without being kinetically trapped while the majority of the misfolded molecules fold through multiple intermediates. The results by Pan et al. also showed that addition of urea can modestly accelerate the rates of

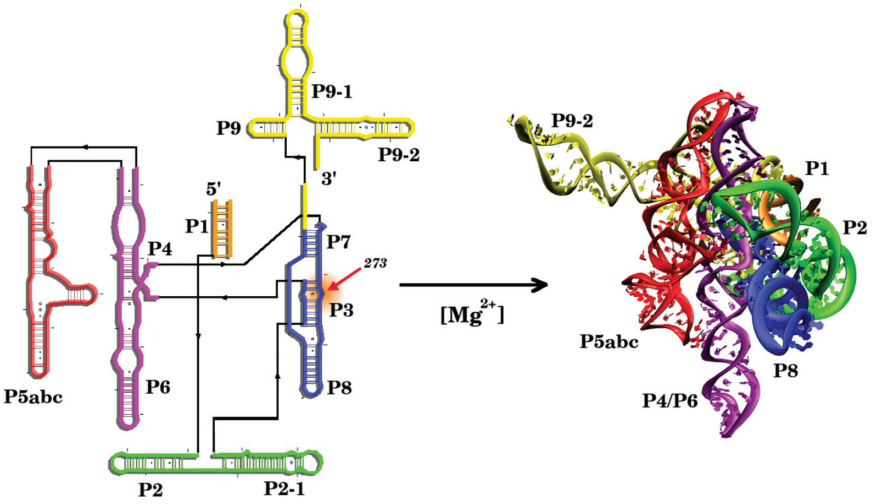

Fig. 2.3 Secondary structure of the most extensively studied group I intron from Tetrahymena. The secondary structure has a number of paired helices indicated by P1 through P9. Upon addition of excess Mg transition to compact tertiary structure, occurs (shown on the right) that is stabilized by the catalytic core formed by an interface involving the P5–P4–P6 and P3–P7–P8 helices. The structure of the independently folding P4–P6 domain is known in atomic detail (Cate et al. 1996). The structure on the right is a model proposed by Westhof and Michel (Lehnert et al. 1996) (See figure insert for colour reproduction)

escape from the misfolded conformations. Subsequent studies have used urea as an analytic probe of RNA stability (Sosnick and Pan 2003) in much the same way as it is done in protein folding studies. Another key prediction of the KPM is that point mutations can alter Φ. Remarkably, a single point mutation U273A in P3 increases Φ to about 80% (Pan et al. 2000). Thus, the mutation greatly reduces the kinetic possibility of being trapped in AltP3 that impedes folding of the wild type.

The most direct evidence for KPM was provided by using single molecule experiments that probes fluorescent energy transfer (FRET) efficiency (**E**) between two dyes attached to the 3′ and 5′ ends of the *Tetrahymena* ribozyme (Zhuang et al. 2000). The value of **E** is high ($\approx$1) in the NBA whereas in the UBA **E** is low because the dyes are, on an average, far apart. Thus, under various folding or unfolding conditions, time-dependent changes in **E** in the FRET signal can be used as a reporter of the folding reaction. Addition of excess Mg^{2+} to initially unfolded molecules initiates the folding process. Under folding conditions **E** increases and the time needed to reach high **E** for the first time is the first passage time, τ_{1i} for the *i*th RNA molecule. From the distribution of first passage times, $P_{FP}(t)$, for an ensemble (in practice 100 molecules will suffice) of unfolded molecules, the probability that a molecule remains unfolded at time t is $P_u(t) = 1-\int P_{FP}(s)\,ds$. Using the measured $P_{FP}(s)$ with single molecule FRET technique (Zhuang et al. 2000) the calculated $P_u(t)$ is best fit using a sum of two exponentials for the 400 nucleotide L-21 ribozyme (Thirumalai et al. 2001). The partition factor $\Phi \approx 0.06$. In other

words, only 6% of the molecules fold rapidly by fast track without being kinetically trapped. It is worth noting that Φ for both L-21 is similar to the estimate for the pre-RNA, which suggests that the folding trajectories for the fast track molecules are similar.

2.4 Hairpin Formation Occurs by Multiple Routes

The relatively small stability gap between the native state and alternate misfolded or native-like conformations (Fig. 2.1) suggests that the folding landscape of even hairpins with a simple loop and a stem is rugged. The possibility of misfolding, at the secondary structural level, was already established in the context of tRNA folding, over 40 years ago (Lindhal et al. 1966). As a result, hairpin formation, when examined in detail, need not follow the classical two-state kinetics. Indeed, a series of recent experiments show that the kinetics of hairpin formation in RNA or ssDNA is best described as a multi-step process (Jung and Van Orden 2006; Ma et al. 2006, 2007), thus challenging the conventional premise that small nucleic acid hairpins, fold in a two-state manner (Bloomfield et al. 2000; Tinoco et al. 2002; Turner et al. 1988).

The signatures of multi-state folding/unfolding are reflected in the kinetic data of ultra fast *T*-jump experiments that can discern the metastable intermediates. Multiple probes attached to the same molecule revealed that the folding is achieved through a series of dynamic steps that occur on vastly different time scales (Jung and Van Orden 2006; Ma et al. 2006, 2007). In contrast, single molecule force experiments (Liphardt et al. 2001; Woodside et al. 2006) showed that, when the ends of molecule are held at the transition mid-force (f_m), the hairpin stochastically hops between the two discrete values of end-to-end distance (R). The statistics of R exhibits a bimodal distribution without signatures of populated intermediates. However, when refolding is initiated by relaxing the applied force (f), metastable intermediates manifest themselves. By varying *f,* transitions from these misfolded structures to the folded structure can be facilitated – a process that is reminiscent of annealing by raising temperature.

To illustrate the consequences of the rugged folding landscape of nucleic acid hairpins, we simulated both thermodynamics and kinetics of RNA hairpin in detail by varying temperatures and mechanical forces using a coarse-grained Three Interaction Site (TIS) model (Hyeon and Thirumalai 2005, 2006). The TIS model simplifies the structural details of a nucleotide into the three coarse-grained interaction centers representing base, ribose, and phosphate group. Using the 22-nucleotide (nt) P5GA RNA hairpin (PDB ID: 1EOR) as a model system, we characterized the equilibrium ensemble of the RNA hairpin over the broad range of T and f conditions, and also simulated the relaxation dynamics of RNA hairpin under T and f-jump/quench conditions (Hyeon and Thirumalai 2008). The dynamics of RNA hairpins are monitored using two order parameters, i.e., the end-to-end distance (R) and the loop dihedral angles (φ) that can best describe the characteristics of the

molecule. Here, $\varphi = 1-\cos(\varphi_i-\varphi_i^0)$ where φ_i is the value of the ith dihedral angle in the GAAA tetraloop in the TIS representation of P5GA, and φ_i^0 is the corresponding value in the folded structure (Hyeon and Thirumalai 2008).

The equilibrium free energy surface expressed in terms of (R, φ) is characterized by two basins of attraction at the locus of critical points (T_m, f_m). Away from the critical condition, only one basin of attraction dominates. The free energy surface succinctly explains the origin of sharp bimodal transition between the folded and unfolded state when the RNA hairpin is subject to force. Thus, from thermodynamic consideration, hairpin formation can be described as a two-state system (see Hyeon and Thirumalai 2008 for details).

The refolding kinetics can be initiated by either a temperature (T) quench from high T to $T < T_m$ or by a force quench to $f < f_m$. Surprisingly, in both cases the kinetic folding pathways cannot be inferred from the free energy landscape. The RNA hairpin reaches the native state via multiple steps as observed in the recent kinetic experiments using high resolution T-jump experiments (Fig. 2.4). The expectation that kinetics can be gleaned from the free energy surface may be valid only if the RNA internal dynamics is rapid enough to establish quasi-equilibrium. For refolding induced by f or T-quench, such an assumption apparently breaks down. We find that the folding trajectories of different molecules are distinct which implies that there is diversity in the folding routes (Fig. 2.4).

The time-dependent changes in the order parameters R and ϕ show differences in folding pathways between T-quench and f-quench refolding. The ensemble of initially unfolded structures prepared by stretching the hairpin differs greatly from the thermally unfolded conformations. The initial ensemble of fully extended conformations, generated by forced-unfolding, is narrow and structurally homogeneous. The various conformations largely differ in the internal degrees of freedom while the overall end-to-end distance is large. Thus, the first step in the hairpin formation from the initially stretched conformations is the tetra-loop formation (Fig. 2.4), corresponding to the slow nucleation stage. Subsequent to the nucleation step the zipping of remaining base pairs leads to hairpin formation. Thus, hairpin is formed by this classic mechanism when folding is initiated by f-quench.

In contrast, upon T-quench, refolding commences from a broad thermal ensemble of unfolded conformations. As a result, nucleation can originate from regions other than near the tetra-loop. Consequently, the pathway diversity is larger when hairpin formation is initiated by T-quench rather than f-quench. The differences in the folding mechanism between these two methods are entirely due to the variations in the initial conformations. Just as folding trajectories in the self-assembly of ribozymes can be altered by pre-incubation with Na^+, here the routes to hairpin formation can be precisely controlled by applying mechanical force.

The simulations show that the complexity of energy landscape observed in ribozyme experiments is already reflected in the formation of simple RNA hairpin (Chen and Dill 2000; Thirumalai and Hyeon 2005; Treiber and Williamson 2001; Woodson 2005). Exploring the details of the heterogeneous kinetics requires multiple probes that control the conformations in the ensemble of unfolded states.

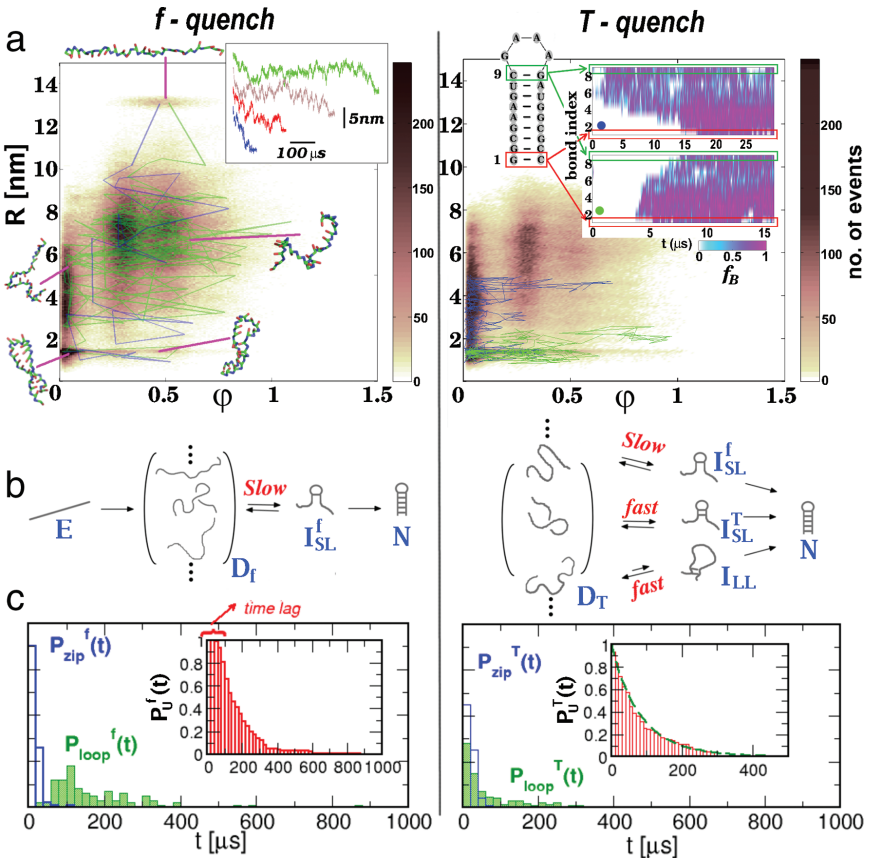

Fig. 2.4 Kinetic analysis of the refolding trajectories upon f-quench and T-quench. (**a**) Conformational space navigated by the refolding trajectories projected onto the (R, φ) plane. The trajectories of individual molecules are overlapped onto the (R, φ) plane. The corresponding trajectories monitored using a single parameter are shown in the insets. (**b**) Summary of the pathways to the NBA inferred from the dynamics depicted in (**a**). (**c**) Statistical analysis of refolding kinetics. The refolding time for each molecule is decomposed into looping and zipping time as $\tau_{FP} = \tau_{loop} + \tau_{zip}$. The fraction of unfolded molecules ($P_u(t)=1-\int_0^t d\tau P_{FP}(\tau)$ where $P_{FP}(\tau)$ is the refolding or first passage time distribution) is plotted in the inset. The probability of the hairpin remaining unfolded upon f-quench $P_u^f(t)$ shows a lag phase (left hand side of C) suggesting the presence of an intermediate, while $P_u^T(t)$ is well fit using $P_u^T(t)=0.4\exp(-t/62\,\mu s) + 0.6\exp(-t/100\mu s)$ (See figure insert for colour reproduction)

2.5 Ion–RNA Interactions Affects Stability, Pathway Diversity and Transition States

To fold, RNA must overcome the large electrostatic repulsion between the negatively charged phosphate groups. At high temperatures ion–RNA interactions are weak, and the gain in translational entropy makes the ions disperse homogeneously in solution without condensing onto RNA. As a result RNA is relatively extended

with $R_G \sim aN^{\nu}$ ($\nu \approx 1$). A naive estimate of the electrostatic repulsion is $E_R \approx (Ne)^2/\varepsilon R_G \approx Nk_BT(l_B/a)$ (Thirumalai et al. 2001). Since $(l_B/b) > 1$ it follows that $E_R/k_BT >> 1$ even when N is small. Therefore, under folding conditions substantial softening of the electrostatic interactions must be achieved through the screening of the electrostatic repulsion or counterion condensation.

Although a complete theoretical treatment of the interaction of counterions and RNA (or other polyelectrolytes for that matter) is lacking, the qualitative aspects of RNA–ion interactions can be understood using the Manning picture (Manning 1978). Charge neutralization is thought to result from the condensation of counterions onto the charged polyanion resulting in overall minimization of the free energy of RNA. Because folded RNA is aperiodic with irregular grooves the electrostatic potential is non-uniform. As a result, the condensed ions can be grouped into distinct classes. Examination of crystal structures of RNA, biophysical and theoretical analysis shows that ions in the vicinity of the strong electrostatic RNA molecule can be considered as (a) diffuse ions that are localized within the volume of RNA or (b) discrete ions that interact specifically with certain sites in the folded structure (Draper 2004).

The theories based on the Manning picture as well as solution to the non-linear Poisson–Boltzmann (NLPB) equation (Draper 2004) show that bulk of the charge neutralization is due to the non-specific association of the diffuse ions on RNA (Heilman-Miller et al. 2001). Counterion-condensation occurs at low temperatures because the loss in the translational entropy of the ions (viewed as unstructured species) is compensated by a gain in the association energy between ions and RNA. As a result of the condensation of the ions there is a substantial reduction in the overall average charge per phosphate group. For highly charged rod ion, condensation occurs if $l_B/A > 1/Z$ where Z is the counterion valence, and A, the distance between charges which is about 3 Å for poly A and 1.3 Å for A-form double helix. The estimate based on charged rods also provides a useful measure of the charge renormalization for RNA.

A few key consequences of the Manning theory follow by treating the condensed and free (in solution) counterions as two equilibrium phases. The chemical potential of the free ions is $\mu_F = -k_BT \log \phi$ where ϕ is the volume fraction of the counterions, while the chemical potential of the diffuse condensed ions is $\mu_C = Ne_RZk_BT \times (l_B/R_G)$ where N is the number of nucleotides, e_R ($<e$) is the fraction of net charge on the phosphate upon condensation of the ions, and Z is the valence of counterion. By equating $\mu_F = \mu_C$ we obtain $Ne_R = -(R_G/l_BZ)\log \phi$. Using an appropriate R_G value for *Tetrahymena* ribozyme it turns out that nearly 90% of the charge is neutralized with $\phi = 0.01$ for monovalent ions. A similar value is obtained for $[CO(NH3)_6]^{3+}$ with $\phi \approx 10^{-6}$ (Heilman-Miller et al. 2001).

The relationship between the extent of charge neutralization and the size of the compact structures lead to several qualitative predictions. (a) Multivalent cations are more efficient than monovalent ions in reducing *Ne*. As a result the concentration of ions needed for inducing compact RNA structures decreases as Z increases. (b) Compared to monovalent ions RNA structures formed by multivalent ions are more compact with R_G decreasing as $\approx 1/Z^2$. (c) The more compact misfolded structures have lower free energies than those formed in the presence of monovalent

ions. As a result the folding rates from the compact misfolded structures to the native state should decrease as Z increases. Below we briefly discuss experiments that have provided support to these predictions.

Stability and Valence. From the simple theoretical picture we infer that as the valence increases the efficiency of inducing compact conformation must also increase. Indeed, experiments show that the midpoint of transition (C_m) decreases as Z increases. In particular, for *Tetrahymena* ribozyme C_m = 0.46 M in Na^+ whereas in $[CO(NH3)_6]^{3+}$, C_m = 12 μM (Heilman-Miller et al. 2001). Cooperativity of the folding transition, and hence the free energy of stability of the native state typically increases with Z although other factors such as size and shape of ions also play an important role (see below). The crucial role played by the ion valence can be understood by considering the electrostatic attraction between the residual charge on the counterions and the renormalized total charge (Ne_R) on RNA after ion condensation. In the monovalent case there is only a weak dipole–dipole interaction between ion pairs that form when Z = 1 ion interacts with the negatively charged phosphate groups. In contrast, multivalent ions induce attractive bridging interactions between regions of RNA that are well separated. Such long range interactions ($\propto 1/r$) in the presence of multivalent ions stabilize compact structures more effectively than monovalent ions because they can bridge two or more phosphate groups.

Charge Density of Ions and RNA Stability. The extent of compaction depends on size as well, and we expect variations in stability at a fixed Z but differing size. Excluded volume interactions between condensed counterions result in spatial correlations that position any two metal ions at distances greater than the sum of their ionic radii. As a result ion size not only determines the distance of closest approach to the negatively charged phosphate group but also affects the spatial location of other diffuse ions. The effect of ion–ion correlation is difficult to include in theoretical treatments within the Manning picture and the NLPB approach (see however Ha and Thirumalai 2003; Tan and Chen 2005; Chen 2008). However, qualitative effects of correlations between condensed ions, due to excluded volume, on RNA stability can be obtained using simple arguments.

As both Z and the volume (V) of the ions contribute to the distribution of ions around the RNA, they also both determine the nature of the counterion-induced compact states of RNA. Thus, to a first approximation, the natural variable that should control RNA stability is the charge density $\zeta = Ze/V$ where Ze is the charge and V is the volume of the cation. In accordance with this expectation, the changes in stability of in *Tetrahymena* ribozyme in various Group II metal ions (Mg^{2+}, Ca^{2+}, Ba^{2+}, and Sr^{2+}) showed a remarkable linear variation with ζ (Fig. 2.5). The extent of stability is largest for ions with the largest ζ (smallest V). Brownian dynamics simulations showed that this effect could be captured solely by non-specific interaction of ions with polyelectrolytes (Fig. 2.5) in the absence of any site-specific ion–RNA interactions. These findings and similar variations of stability in different sized diamines (Koculi et al. 2004) show that (a) the bulk of the stability arises from non-specific association of ions with RNA, and (b) stability can be greatly altered by valence, shape, and size of the counterions.

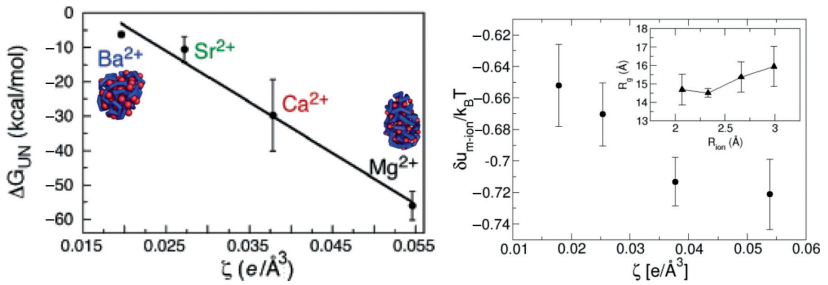

Fig. 2.5 (*Left*) Stability of *Tetrahymena* L-21Sca ribozyme (ΔG_{UF}) vs. cation charge density (ζ). (*Right*) The Brownian Dynamics simulations of polyelectrolyte collapse for the average stabilization energy between monomer of polyelectrolyte and ion as a function of ζ. The remarkable linear dependence on the left is captured by ion-induced collapse of flexible polyelectrolytes. The radius of gyration of the collapsed polyelectrolyte ($N = 120$) is plotted as a function of group II metalion size in the inset (See figure insert for colour reproduction)

Diversity of Folding Routes Depends on Initial Conditions and ζ. Single molecule FRET experiments indicate that the time-dependent changes ($\mathbf{E}(t)$) in FRET efficiencies vary from molecule to molecule (Zhuang et al. 2000). While the averages over an ensemble of such single molecule measurements are consistent with bulk experiments, the substantial variations might be indicative of inherent pathway diversity. Indeed, a consequence of the KPM is that the partition factor, Φ, can be altered not only by sequence variations but can also be changed by altering initial conditions. Analysis of single molecule FRET data of L-21 Sca I construct of the group I intron shows that preincubation in excess Na^+ before initiating folding alters Φ. Monovalent ions induce compaction in the initial unfolded ensemble of structures. As a result, folding commences from a region of the rugged energy landscape that restricts the starting unfolding ensemble which differs from the more expanded ensemble of structures in the absence of Na^+. Thus, upon initiation of folding the ribozyme assembles via different routes. The partition factor is a global measure of the pathway diversity because Φ is proportional to the number of molecules that fold rapidly through a restricted channel in the folding landscape (Fig. 2.2) without being kinetically trapped. Thus, the nature of pathways traversed depends critically on the starting RNA structures that can be manipulated by pre-incubation with monovalent cations.

Varying ζ can also change the diversity of folding routes. Just as pre-incubation with Na^+ leads to a more compact ensemble of initial structures the extent of collapsed structures can be altered by varying ζ. As a result, ions of differing ζ can modulate the diversity of folding routes. From a suitable generalization of the Manning picture it follows that as ζ increases, the extent of compaction of RNA increases. From the folding landscape (Fig. 2.6) it follows that the low ζ-ensemble of structures $\{\mathbf{I}_{NS}\}$ is less compact than those formed in ions with high ζ. As a result the higher the entropy associated with low ζ ions, the number of conformations in the $\{\mathbf{I}_{NS}\}$ ensemble is greater than in ions with high ζ. Thus, we expect that folding pathway diversity should increase when the ion charge density is low – prediction that can be tested using single molecule measurements.

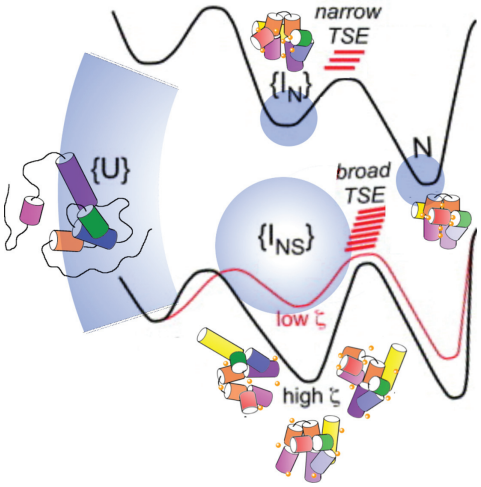

Fig. 2.6 Coupling of diversity of the folding pathways and heterogeneity of the transition state structures of *Tetrahymena* ribozyme to the charge density of counterions (ζ). The majority of the ribozyme folds through intermediates in which the core P3 helix is replaced by a non-native helix alt-P3. The pathway diversity increases as ζ decreases (lower part of the figure). For the fraction ($(1-\Phi)$) slow track molecules, the transition state ensemble (TSE) along the $\{U\}\rightarrow\{I_{NS}\}\rightarrow N$ pathway becomes broader and less structured as the ζ decreases. Thus, the system is more dynamic in polyamines (low ζ) than Mg^{2+} (high ζ). The fast track molecules that fold via $\{U\}\rightarrow\{I_N\}\rightarrow N$ (Φ) (upper part of the figure) first form specifically collapsed compact structure that becomes increasingly native-like as the folding reaction proceeds. For $\{U\}\rightarrow\{I_N\}\rightarrow N$ we suggest that the TSE is narrow with little structural heterogeneity. The pathway diversity is expected to increase as ζ decreases (See figure insert for colour reproduction)

Role of ζ on the Plasticity of the Transition State Ensemble (TSE). In contrast to protein folding much less is known about the nature of TSE in self-assembly of large RNA molecules. Only recently has there been concerted efforts to decipher the nature of TSE in RNA folding (Bokinsky et al. 2003; Fang et al. 2002; Koculi et al. 2006). Although it is tempting to propose a very general picture of the TSE or the rate-limiting step in RNA self-assembly, it should be kept in mind that, just as the nature of intermediates in RNA folding can be easily altered by changing the properties of ions so too can the location and plasticity of the TSE (Koculi et al. 2006). Using single molecule FRET experiments of docking and undocking in hairpin ribozyme it has been suggested that the TSE is compact in Mg^{2+}, and perhaps share much of the structural characteristics of collapsed native-like intermediates. By using a combination of biophysical methods it has been suggested that the TSE for the C domain of RNase P involves reorganization of metal-ion binding sites late in the folding process.

A much more general analysis of the TSE and its variations requires studies that alter the valence, size, and shape of the ions. In a recent study, the variations in folding kinetics and TSE movements were probed using concentrations of polyamines ($^{+}H_3N(CH_2)_nNH_3^{+}$ with n = 2–5) as a natural perturbation of the RNA

folding landscape. Several key observations were made: (a) The TSE is much broader in polyamines than the Mg^{2+} which has a larger ζ. (b) Ions with larger ζ give rise to higher free energy barriers to folding. (c) By using the Tanford β parameter it was surmised that the average location of the TSE is closer to the native state (β closer to unity) when folded in ions with small ζ. These observations allowed us to describe the general changes in the TSE as ζ is changed (Fig. 2.6). At low ζ the $\{\mathbf{I}_{NS}\}$ ensemble is less compact with higher entropy than when ζ is high. Thus, the free energy barriers are largely determined by entropy changes at low ζ. By contrast, when ζ is high (n = 2 in polyamine for example) the $\{\mathbf{I}_{NS}\}$ ensemble is more compact, and the free energy barrier separating the intermediates and the NBA is largely enthalpic. Thus, by modulating the charge density of ions one can modulate the interplay between entropy and enthalpy and control the very nature of transition state structures in RNA folding.

2.6 Folding Rates, free Energy Barriers and Kramers Prefactor for RNA

Besides the nature of ions, sequence length (N) also plays an important role in determining the folding time, τ_F. Given that non-coding RNAs are "evolved" heteropolymers, it is not surprising that the N should play a crucial role in controlling the folding rate (k_F). For minimally frustrated sequences, $\log(k_F/k_0) \sim \alpha \log N$ with $\alpha \approx 4$ (Thirumalai 1995) at $T < T_F$ where the prefactor k_0 can be obtained using Kramers' theory. Because biopolymers are topologically frustrated, there is residual roughness even in two-state folders. As a result, the folding kinetics characterized with a single barrier crossing event follows the relation

$$\log (\tau_F/\tau_o) = \Delta G^{\ddagger}_{UF}/k_B T$$

where the prefactor τ_0, that is often estimated using transition state theory (TST), has to be determined using Kramers theory. If we assume a Gaussian distribution for the free energy barriers with dispersion $< (\Delta G^{\ddagger}_{UF})^2 > \sim N$ then $\Delta G^{\ddagger}_{UF}/k_B T \sim N^{\beta}$ with $\beta = 1/2$ (Thirumalai 1995; Thirumalai and Hyeon 2005). Other arguments predict that $\beta = 2/3$ (Finkelstein and Badretdinov 1997; Wolynes 1997). The sublinear scaling of the effective barrier height with N naturally explains both rapid folding (kinetics) and marginal stability (thermodynamics) of single domain proteins and RNA.

In contrast to proteins (Li et al. 2004), the number of experiments for RNA molecules that report τ_F as a function of N is small; hence, the variation of k_F with N has not been examined. Experiments on hairpin formation in oligonucleotides and helix-coil transition theories already showed that k_F must be sensitive to N. We have analyzed the N dependence on RNA folding kinetics using the available data from the literature (Thirumalai and Hyeon 2005). Here, we extend these calculations using a slightly larger dataset. Surprisingly, the rates that vary over 7 orders of

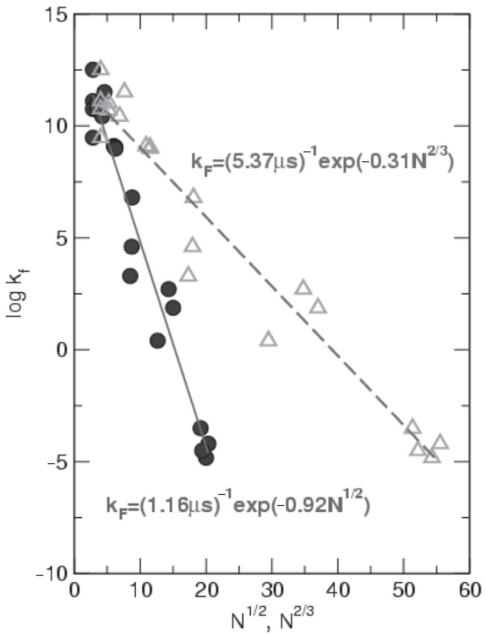

Fig. 2.7 Dependence of RNA folding rates as a function of N, the number of nucleotides. Fits of log k_f as a function of N^{β} with $\beta = 1/2$ or $\beta = 2/3$ are also shown

magnitude depend on N as predicted by theory. The correlation coefficients for both values of β are in excess of 0.9. In contrast to proteins, the predicted N dependence of k_F is more closely obeyed (Finkelstein and Badretdinov 1997; Galzitskaya et al. 2003, 2004). Using the results in Fig. 2.7, the difficult-to-measure prefactor τ_0, which should be estimated by using Kramers' theory, can be calculated. From the scaling plots in Fig. 2.7 we find that $\tau_0 \approx 1.2\,\mu s$ for $\beta = 1/2$ and $\tau_0 \approx 5.4\,\mu s$ for $\beta = 2/3$. Both these estimates for the RNA folding prefactor are nearly *six orders of magnitude larger than the TST value* ($=h/k_BT \sim 0.2\,ps$). The large value of τ_0 implies that the effective free energy barriers from the measurements of rates alone using TST prefactor, overestimates the activation free energies by ~15 k_BT. The TST prefactor is applicable only if breakage of a single bond is involved at the transition state. While this may be appropriate for gas phase reactions it cannot describe folding that is determined by collective events. The prefactor τ_0 represents the time scale in which folding can occur without barriers, i.e., by diffusion limited process. An estimate for the most elementary event in folding (for example base pairing in RNA) leads to the Kramers' estimate of ~1 μs for τ_0. Our estimate is in accord with the typical base pairing rate (Porschke and Eigen 1971; Porschke et al. 1973).

2.7 Conclusions

A number of factors, such as the lack of diversity of the building blocks, sequence variations, polyelectrolyte character of the phosphate backbone, and the subtle roles played by the ions, contribute to the complexity of RNA folding. The interplay of

these factors are evident in the emergence of astounding variety of structures with each fold having both regions of flexibility and rigidity – features that lend themselves to RNA molecules being able to execute wide-ranging cellular functions. However, from a biophysical perspective the following features make it hard to provide a molecular understanding of RNA folding. (a) It would seem that the constraint of Watson–Crick base pairing and the inherent stability of RNA secondary structures would make RNA folding relatively simpler than the protein-folding problem. However, nearly half of the base pairs are involved in non-WC structures, which makes it difficult to predict even the RNA secondary structures especially when the number of nucleotides exceeds about 50. (b) The inherent complexity of RNA folding kinetics can be better appreciated by comparisons to the better-studied protein folding problem. To a large extent, folded proteins are stabilized by favorable interactions between hydrophobic residues that are buried in the interior. The interactions between all the residues are short-range, and are in the order of the size of the residues themselves (~6 Å). In contrast, the ranges of interactions between the nucleotides or the structural motifs that drive RNA folding vary greatly. The ion-mediated interactions occur on the persistence length scale that varies from about (1–2) nm depending on the ion concentration. Other interactions using hydrogen bonds between the bases and stacking interaction that stabilize various elements of the RNA structure are shorter range in distance. The interplay of the interactions on distinct length scales that can be altered by changing valence and size of ions gives rise to multiple scenarios for folding. Despite these difficulties it is remarkable that, at some global level, the principles based on KPM, polyelectrolyte theory, and ion–RNA interactions allow us to qualitatively rationalize many puzzling aspects of RNA folding. Developments in single molecule experiments and novel theoretical tools will be needed to quantitatively understand the richness of RNA folding.

Acknowledgments One of us (DT) is grateful to Sarah A. Woodson for pleasurable collaboration on all aspects of RNA folding for over 12 years. We are pleased to acknowledge useful discussions with her and Eda Koculi on ion-RNA interactions. This work was supported in part by a grant from the National Science Foundation (CHE 05-14056).

References

Aronovitz JA, Nelson DR (1986) Universal features of polymer shapes. J Phys 47(9):1445–1456

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution. Science 289(5481):905–920

Bloomfield VA, Crothers DM, Tinoco I Jr (2000) Nucleic acids, structures, properties and functions. University Science Books, Sausalito, CA

Bokinsky G, Rueda D, Misra VK, Rhodes MM, Gordus A, Babcock HP et al. (2003) Single-molecule transition-state analysis of RNA folding. Proc Natl Acad Sci U S A 100(16):9302–9307

Caliskan G, Hyeon C, Perez-Salas U, Briber RM, Woodson SA, Thirumalai D (2005) Persistence length changes dramatically as RNA folds. Phys Rev Lett 95(26):268–303

Cate JH, Gooding AR, Podell E, Zhou KH, Golden BL, Kundrot CE et al. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. Science 273(5282):1678–1685

Chauhan S, Woodson SA (2008) Tertiary interactions determine the accuracy of RNA folding. J Am Chem Soc 130(4):1296–1303
Chen SJ (2008) RNA folding: conformational statistics, folding kinetics, and ion electrostatics. Annu Rev Biophys Biomol Struct 37:197–214
Chen SJ, Dill KA (2000) RNA folding energy landscapes. Proc Natl Acad Sci U S A 97(2):646–651
Dima RI, Thirumalai D (2004) Asymmetry in the shapes of folded and denatured states of proteins. J Phys Chem B 108:6564–6570
Dima RI, Hyeon C, Thirumalai D (2005) Extracting stacking interaction parameters for RNA from the data set of native structures. J Mol Biol 347(1):53–69
Draper DE (2004) A guide to ions and RNA structure RNA 10:335–343
Doudna J, Cech T (2002) The chemical repertoire of natural ribozymes. Nature 418:222–228
Fang XW, Thiyagarajan P, Sosnick TR, Pan T (2002) The rate-limiting step in the folding of a large ribozyme without kinetic traps. Proc Natl Acad Sci U S A 99(13):8518–8523
Finkelstein AV, Badretdinov AY (1997) Rate of protein folding near the point of thermodynamic equilibrium between the coil and the most stable chain fold. Fold Des 2(2):115–121
Galzitskaya OV, Garbuzynskiy SO, Ivankov DN, Finkelstein AV (2003) Chain length is the main determinant of the folding rate for proteins with three-state folding kinetics. Proteins 51(2):162–166
Guerriertakada C, Gardiner K, Marsh T, Pace N, Altman S (1983) The RNA moiety of ribonuclease-P is the catalytic subunit of the enzyme. Cell 35(3):849–857
Guo Z, Honeycutt JD, Thirumalai D (1992) Folding kinetics of proteins: a model study. J Chem Phys 97(1):525–535
Ha BY, Thirumalai D (2003) Bending rigidity of stiff polyelectrolyte chains: a single chain and bundle of multichains. Macromolecules 46:9658–9666
Heilman-Miller SL, Thirumalai D, Woodson SA (2001) Role of counterion-condensation in folding of the *Tetrahymena* ribozyme. I. Equilibrium stabilization by cations. J Mol Biol 306:1157–1166
Hofacker I (2003) Vienna RNA secondary structure server. Nucleic Acids Res 31 3429–3431
Hyeon C, Thirumalai D (2005) Mechanical unfolding of RNA hairpins. Proc Natl Acad Sci U S A 102(19):6789–6794
Hyeon C, Thirumalai D (2006) Forced-unfolding and force-quench refolding of RNA hairpins. Biophys J 90(10):3410–3427
Hyeon C, Thirumalai D (2008) Multiple probes are required to explore and control the rugged energy landscape of RNA hairpins. J Am Chem Soc 130:1538–1539
Hyeon C, Dima RI, Thirumalai D (2006) Size, shape, and flexibility of RNA structures. J Chem Phys 125(19):194905
Jung JY, Van Orden A (2006) A three-state mechanism for DNA hairpin folding characterized by multiparameter fluorescence fluctuation spectroscopy. J Am Chem Soc 128(4):1240–1249
Koculi E, Lee NK, Thirumalai D, Woodson SA (2004) Folding of the *Tetrahymena ribozyme* by polyamines: importance of counterion valence and size. J Mol Biol 341(1):27–36
Koculi E, Thirumalai D, Woodson SA (2006) Counterion charge density determines the position and plasticity of RNA folding transition states. J Mol Biol 359(2):446–454
Koculi E, Hyeon C, Thirumalai D, Woodson SA (2007) Charge density of divalent metal cations determines RNA stability. J Am Chem Soc 129(9):2676–2682
Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR (1982) Self-splicing RNA-auto-excision and auto-cyclization of the ribosomal- RNA intervening sequence of *Tetrahymena*. Cell 31(1):147–157
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 273:1678–1685
Li MS, Klimov DK, Thirumalai D (2004) Thermal denaturation and folding rates of single domain proteins: size matters. Polymer 45(2):573–579
Lindhal T, Adams A, Fresco JR (1966) Renaturation of transfer ribonucleic acids through site binding of magnesium. Proc Natl Acad Sci U S A 55:941–948

Liphardt J, Onoa B, Smith SB, Tinoco I Jr, Bustamante C (2001) Reversible unfolding of single RNA molecules by mechanical force. Science 292(5517):733–737
Ma HR, Proctor DJ, Kierzek E, Kierzek R, Bevilacqua PC, Gruebele M (2006) Exploring the energy landscape of a small RNA hairpin. J Am Chem Soc 128(5):1523–1530
Ma HR, Wan CZ, Wu AG, Zewail AH (2007) DNA folding and melting observed in real time redefine the energy landscape. Proc Natl Acad Sci USA 104(3):712–716
Manning GS (1978) The molecular theory of polyelectrolyte solutions with applications to the electrostatic Properties of polynucleotides. Q Rev Biophys 11:179–246
Nissen P, Hansen J, Ban N, Moore PB, Steitz TA (2000) The structural basis of ribosome activity in peptide bond synthesis. Science 289:920–930
Odijk T, (1977) Polyelectrolytes near rod limit. J Polym Sci Polym Phys 15:477–483
Onoa B, Dumont S, Liphardt J, Smith SB, Tinoco I Jr, Bustamante C (2003) Identifying kinetic barriers to mechanical unfolding of the *T-thermophila* ribozyme. Science 299(5614): 1892–1895
Pan J, Thirumalai D, Woodson SA (1999) Magnesium-dependent folding of self-splicing RNA: exploring the link between cooperativity, thermodynamics, and kinetics. Proc Natl Acad Sci U S A 96:6149–6154
Pan J, Deras ML, Woodson SA (2000) Fast folding of a ribozyme by stabilizing core interactions: evidence for multiple folding pathways in RNA. J Mol Biol 296:133–144
Porschke D, Eigen M (1971) Co-operative non-enzymic base recognition 3. Kinetics of helix-coil transition of oligoribouridylic acid system and of oligoriboadenylic acid alone at acidic pH. J Mol Biol 62(2):361–381
Porschke D, Uhlenbec O, Martin FH (1973) Thermodynamics and kinetics of helix-coil transition of oligomers containing GC base pairs. Biopolymers 12(6):1313–1335
Rangan P, Masquida B, Westhof E, Woodson SA (2004) Architecture and folding mechanism of the Azoarcus group I pre-tRNA. J Mol Biol 339(1):41–51
Russell R, Millett IS, Tate MW, Kwok LW, Nakatani B, et al. (2002a) Rapid compaction during RNA folding. Proc Natl Acad Sci U S A 99:4266–4271
Russell R, Zhuang X, Babcock HP, Millett IS, Doniach S, Chu S, Herschlag D (2002b) Exploring the folding landscape of a structured RNA. Proc Natl Acad Sci U S A 99(1):155–160
Skolnick J, Fixman M (1977) Electrostatic persistence length of a wormlike polyelectrolyte. Macromolecules 10:944–948
Sosnick TR, Pan T (2003) RNA folding: models and perspectives. Curr Opin Struct Biol 13(3):309–316
Tan ZJ, Chen SJ (2005) Electrostatic correlations and fluctuations for ion binding to a finite length polyelectrolyte. J Chem Phys 122:044903
Thirumalai D (1995) From minimal models to real proteins: time scales for protein folding kinetics. J Phys I France 5:1457–1467
Thirumalai D (1998) Native secondary structure formation in RNA may be slave to tertiary folding. Proc Natl Acad Sci U S A 95:11506–11508
Thirumalai D, Hyeon C (2005) RNA and protein folding: common themes and variations. Biochemistry 44(13):4957–4970
Thirumalai D, Lee NK, Woodson SA, Klimor DK (2001) Early events is RNA folding. Ann Rev Phys chem 52:751–762
Thirumalai D, Woodson SA (1996) Kinetics of folding of proteins and RNA. Acc Chem Res 29:433–439
Tinoco I Jr, Bustamante C (1999) How RNA folds. J Mol Biol 293(2):271–281
Tinoco I Jr, Sauer K, Wang JC, Puglisi JD (2002) Physical chemistry principles and applications in biological sciences. Prentice-Hall, Englewood Cliffs, NJ
Treiber DK, Williamson JR (2001) Beyond kinetic traps in RNA folding. Curr Opin Struct Biol 11(3):309–314
Turner DH, Sugimoto N, Freier SM (1988) RNA structure prediction. Ann Rev Biophys Chem 17:167–192

Wolynes PG (1997) Folding funnels and energy landscapes of larger proteins within the capillarity approximation. Proc Natl Acad Sci U S A 94(12):6170–6175

Woodside MT, Anthony PC, Behnke-Parks WM, Larizadeh K, Herschlag D, Block SM (2006) Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. Science 314(5801):1001–1004

Woodson SA (2005) Structure and assembly of group I introns. Curr Opin Struct Biol 15(3):324–330

Wu M, Tinoco I Jr (1998) RNA folding causes secondary structure rearrangement. Proc Natl Acad Sci U S A 95:11555–11560

Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JHD et al. (2001) Crystal structure of the ribosome at 5.5 angstrom resolution. Science 292(5518):883–896

Zarrinkar PP, Williamson JR (1994) Kinetic intermediates in RNA folding. Science 265(5174):918–924

Zhuang Z, Bartley L, Babcock A, Russell R, Ha T, Herschlag D, Chu S (2000) A single-molecule study of RNA catalysis and folding. Science 288:2048–2051

Zuker M, Stiegler P (1981) Optimal computer folding of larger RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res 293:271–281

http://www.springer.com/978-3-540-70833-9

Non-Protein Coding RNAs
Walter, N.; Woodson, S.A.; Batey, R.T. (Eds.)
2009, XI, 398 p., Hardcover
ISBN: 978-3-540-70833-9