BIOCHEMISTRY

 Volume 44, Number 13 April 5, 2005

Current Topics

RNA and Protein Folding: Common Themes and Variations†

D. Thirumalai*,‡,§ and Changbong Hyeon‡

Biophysics Program, and Department of Chemistry and Biochemistry, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

Received December 21, 2004; Revised Manuscript Received February 17, 2005

ABSTRACT: Visualizing the navigation of an ensemble of unfolded molecules through the bumpy energy landscape in search of the native state gives a pictorial view of biomolecular folding. This picture, when combined with concepts in polymer theory, provides a unified theory of RNA and protein folding. Just as for proteins, the major folding free energy barrier for RNA scales sublinearly with the number of nucleotides, which allows us to extract the elusive prefactor for RNA folding. Several folding scenarios can be anticipated by considering variations in the energy landscape that depend on sequence, native topology, and external conditions. RNA and protein folding mechanism can be described by the kinetic partitioning mechanism (KPM) according to which a fraction (Φ) of molecules reaches the native state directly, whereas the remaining fraction gets kinetically trapped in metastable conformations. For two-state folders $\Phi \approx 1$. Molecular chaperones are recruited to assist protein folding whenever Φ is small. We show that the iterative annealing mechanism, introduced to describe chaperonin-mediated folding, can be generalized to understand protein-assisted RNA folding. The major differences between the folding of proteins and RNA arise in the early stages of folding. For RNA, folding can only begin after the polyelectrolyte problem is solved, whereas protein collapse requires burial of hydrophobic residues. Cross-fertilization of ideas between the two fields should lead to an understanding of how RNA and proteins solve their folding problems.

Proteins (*1*) and RNA molecules (*2*) are enzymes that catalyze biochemical reactions. The enzymology of proteins has been studied for a number of decades (*1*). In contrast, the need to understand, in molecular detail, the catalytic activity of RNA molecules earnestly began only after the pioneering discovery of self-splicing activity of pre-ribosomal RNA of the ciliate *Tetrahymena* (*3*−*6*). With this discovery the ribozyme era was launched. In the intervening years, RNA enzymes (ribozymes) have become ubiquitous. These developments and the linkage of protein misfolding to neurodegenerative disease (*7*, *8*) have made it urgent to investigate all aspects of biomolecular folding.

Similarly, the discoveries of the increasingly diverse role RNA plays in biology (*2*) requires a conceptual framework for understanding the folding kinetics of RNA molecules (*2*, *9*). To function as successful catalysts, ribozymes and proteins adopt well-defined three-dimensional structures either spontaneously or in association with other biomolecules. Thus, the need to understand the enzymatic activities of proteins and ribozymes inevitably leads to the question: How do these molecules fold? The quest to answer this question has led to major experimental (*10*−*12*) and theoreti-

† This work was supported by grants from the National Science Foundation and the National Institute of Health (1R01GM067851-01).

* To whom correspondence should be addressed. Phone: 301-405-4803; fax: 301-314-9404; e-mail: thirum@glue.umd.edu.

‡ Biophysics Program.

§ Department of Chemistry and Biochemistry.

10.1021/bi047314+ CCC: $30.25
Published on Web 03/11/2005

cal developments over the last 15 years. In the past decade alone novel experimental methods (*11*, *13*) have allowed us to probe folding events from the microsecond time scale. More recently, single molecule experiments (*14*−*16*) have monitored the folding of one protein or ribozyme (*17*, *18*) at a time thus providing a detailed picture of their folding landscape. It remains only a matter of time before single molecule experiments will enable us to routinely watch enzymes execute their catalytic function (*19*). To some extent these developments were spurred by theoretical studies that emphasized the importance of visualizing protein folding in terms of the energy landscape (*20*−*24*). Although initially the statistical mechanical and polymer physics based perspectives were used to describe protein folding, more recently, similar approaches have been adopted to obtain details of RNA folding (*25*, *26*).

Some time ago, we argued that the folding of proteins and RNA can be described using a unified perspective (*25*, *27*). The resulting conceptual framework, the kinetic partitioning mechanism (KPM),[1] led to a number of testable predictions for folding of proteins and large RNA molecules (*25*). Many of these predictions have been confirmed for several RNA molecules using a variety of techniques (*17*, *28*). From a global perspective, there are a number of common principles that must govern the folding of proteins and RNA (*25*, *29*−*31*). Under unfolding conditions (high temperature (T) or presence of denaturants) even a moderate-sized polypeptide chain explores potentially a large number of unfolded conformations. Similarly, the number of unfolded conformations of RNA molecules at low ionic concentrations or high T is also exponentially large. Biomolecular folding is concerned with how the transition from an ensemble of such disordered structures to a well-defined native state occurs. The native states of RNA molecules have Watson−Crick pairs between complementary bases and a precise organization of the various structural motifs (hairpin, bulges, loops, etc.) (*32*). Polypeptide chains form dense close packed structures with discernible secondary structural features (α-helices, β-strands, and loops) (*33*). The goal of folding kinetics studies is to describe the pathways, nature of the intermediates, and the structural features of the transition state ensemble (TSE).

The interplay between theory and experiments has been fruitful in providing insights into the mechanisms of self-assembly of RNA (*34*) and proteins (*35*−*37*). The realization that certain global aspects of RNA folding can be understood using the tools originally developed to describe protein folding (*20*, *24*, *38*, *39*) has resulted in cross-fertilization between the two fields (*40*). Here, we describe the origins of the common themes that govern the folding of proteins and RNA and outline a number of unifying aspects that emerge at a global level in the description of their folding kinetics of RNA and proteins. The common themes also extend to chaperonin-assisted folding of proteins and protein-assisted assembly of RNA molecules. The iterative annealing mechanism, originally developed to describe the function of GroEL (*39*), can also be used for ribonucleoprotein assembly. There are crucial differences in their self-assembly as well. The initial events leading to chain compaction in RNA are driven by counterion-mediated collapse, whereas lowering of the free energy of the polypeptide chain by burying the hydrophobic residues is the main driving force in the collapse of proteins (*41*).

Diagram of States and Marginal Stability of Folded States.

(a) Proteins. The conformations of a polypeptide chain are determined not only by the sequence but also by external conditions (pH, ionic strength, T, and cosolutes). Although a majority of experiments are triggered by addition of a cosolute (urea and guanidinium hydrochloride), the interest is to obtain thermodynamics and kinetics at neutral pH and in the absence of denaturants. Under these conditions the "phases" of a monomeric polypeptide chain are determined only by T. At high temperatures the polypeptide chain resembles an unfolded (U) coil. At a characteristic temperature, T_θ, (the Flory temperature) the polypeptide chain undergoes a coil to globule (ensemble of collapsed states) transition. Finally, the transition to the native basin of attraction (NBA) occurs at the folding temperature T_F (also referred to as the melting temperature in the experimental literature). The simplest phase diagram of the polypeptide chain consists of three phases, namely, random coil ($T > T_\theta$), collapsed phase ($T_\theta < T < T_F$), and the folded state $T \leq T_F$. To distinguish between extended and collapsed states monomer density, ρ, (or equivalently the radius of gyration R_g) is a natural order parameter. When protein is extended $\rho \approx N/R_g^3$ is small, whereas $\rho \approx O(1)$ in the compact phase. Because $\rho \sim O(1)$ in both the compact and the native state, an additional order parameter is needed to characterize the NBA. The structural overlap function, χ, that measures the similarity to the native state is an order parameter that differentiates between compact non-native states ($\chi \neq 0$) and the native state ($\chi = 0$) (or fraction of native contacts Q (*42*, *43*). Thus, two order parameters (ρ, χ) are needed to describe the diagram of states of protein when temperature is varied. The collapse transition at T_θ is thought to be second order, whereas the folding transition is first order. For proteins that fold by an all-or-none transition $T_\theta \approx T_F$ so that efficient folding occurs near a tricritical point (*44*). If collapse and folding are nearly simultaneous (this requires measuring both R_g and fraction of molecules in the NBA) then $\sigma_{CT} = |T_\theta - T_F|/T_\theta$ is small (*44*). An example is protein L in which collapse and the acquisition of the native structure is synchronous (*45*). When σ_{CT} is different from zero folding occurs through long-lived intermediates which complicate folding kinetics. Thus, there is a natural link between thermodynamics and the associated folding mechanisms. At temperature much lower than T_F the polypeptide chain exhibits the characteristics of glasses so that folding becomes sluggish. Optimal folding sequences should maximize T_F/T_g (*46*), which is equivalent to minimizing σ_{CT} (*27*).

(b) RNA. Because RNA folding requires counterions to neutralize the negatively charged phosphate groups, it follows that the diagram of states is determined by T and the concentration, C, of counterions. The equilibrium phase diagram of even simple RNA molecules [such as tRNA (*47*) or P5abc domain] is rich in the (T, C) plane. In terms of the Bjerrum length, $l_B = e^2/4\pi\epsilon k_B T$ (e is the electron charge, ϵ is the dielectric constant of water, k_B is the Boltzmann

[1] Abbreviations: KPM, kinetic partitioning mechanism; NBA, native basin of attraction; CBA, competing basin of attraction; NC, nucleation collapse; MFN, multiple folding nuclei; TSE, transition state ensemble; IAM, iterative annealing mechanism; SP, substrate protein; CBP2, cytochrome *b* pre-mRNA processing protein 2; CYT-18, *Crassa* mitochondrial tyrosyl tRNA synthetase; RNP, ribonucleoprotein.

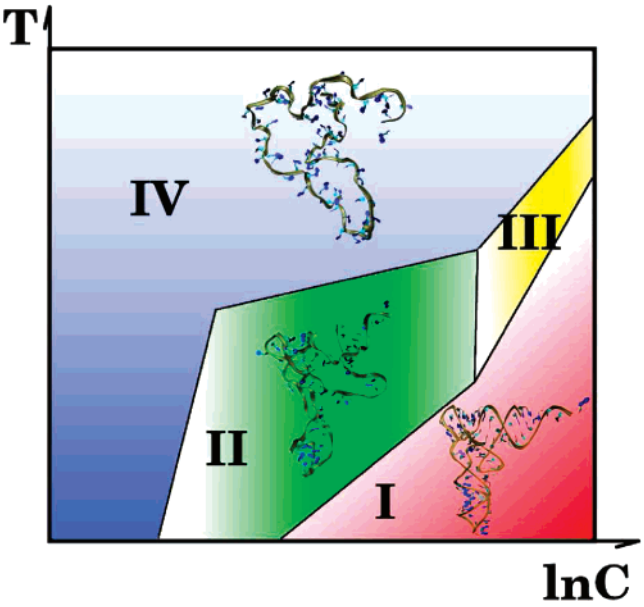

FIGURE 1: Proposed (T, C) phase diagram for a "simple" RNA. There are potentially four phases in the (T, C) plane. Phase I corresponds to the native state ($l_B/b > 1$, $C > C_m$). In phase II structures are stable ($l_B/b > 1$, $C < C_m$). Partially compact structures with fluctuating tertiary structure (no representative structure is shown) form in phase III ($l_B/b < 1$, $C > C_m$) At high T and low C (phase IV) RNA adopts fluctuating expanded single-stranded random coil structure ($l_B/b < 1$, $C \ll C_m$). The boundaries separating the phase are a guide to the eye.

constant), one can predict at least four phases. RNA adopts a single-stranded random coil conformation stabilized, perhaps, by a few stacking interaction when $l_B/b < 1$ and $C \ll C_m$ where C_m is the midpoint of the counterion concentration at which the transition to the folded state occurs. When $l_B/b > 1$ and $C > C_m$ the folded structure is stable. For $l_B/b > 1$ and $C < C_m$ stable secondary structures could form. The conformation of RNA in this phase is expected to be extended but less so than in the random coil state. It is difficult to predict the nature of structures when $l_B/b < 1$ and $C > C_m$. In all likelihood such structures could be somewhat compact and contain fluctuating tertiary interaction as well. These structures may be analogous to the molten globule phase in proteins (Figure 1).

In a remarkable series of papers in the early 1970s Crothers and co-workers (*48*−*50*) obtained, using careful experiments and theoretical analysis, the equilibrium phase diagram of tRNA in the (T,C) plane. Just as summarized above, they were able to identify four distinct phases. They conjectured that the structures in the region $l_B/b < 1$ and $C > C_m$ of tRNA resembles a cloverleaf or some variant. The precise determination of these structures may be difficult because of their marginal stability. For RNA sequences (tRNA or P5abc domain of the *Tetrahymena* ribozyme) the region ($l_B/b < 1$ and $C > C_m$) may be relatively small, so that upon decreasing l_B with $C > C_m$ the transition from the unfolded states with well-defined secondary structure to the NBA may appear "two-state like". However, the separation of energy scales stabilizing RNA structures results in RNA folding occurring over a wide range of time scales involving multitude of intermediates.

Folding Scenarios from an Energy Landscape Perspective. General Considerations. Proteins (*21*) and RNA (*25*, *51*) are stabilized by several competing interactions. Chain compaction is achieved by hydrophobic and/or stacking interactions, while hydrophilic and charged interactions are better accommodated by extended structures. Multiple competing interactions make RNA and proteins "frustrated" (*52*), i.e., that not all favorable interactions at specific residues or nucleotides can be simultaneously satisfied. As a result, the underlying energy landscape is rugged, i.e., there are multiple minima separated by free energy barriers of varying heights. In condensed matter physics energetically frustrated systems have been extensively studied (*53*). An example is the Ising spin glass (*53*) in which ferromagnetic interactions, which align the spins in the same direction, and antiferromagnetic interactions, which order neighboring spins in the opposite direction, are simultaneously present. The presence of conflicting interactions leads to the spin glass phase that is characterized by glassy dynamics which prevents transition to the ordered ferromagnetic phase. Similarly, heteropolymers or a randomly designed polynucleotides also exhibit all the hallmarks of glasses (*54*) and hence are not biologically competent. On the other hand, naturally occurring proteins and RNA, which are finite-sized, may have evolved to minimize such conflicts in the interactions that stabilize the native state. Protein sequences that have compatible interactions are said to be minimally frustrated (*46*, *52*). A similar notion has been introduced for rationalizing multidomain RNA folding (*51*).

Besides energetic frustration proteins and RNA experience "topological frustration" that arises from the polymeric nature of these molecules (*25*, *55*, *56*). Consider the formation of the native structure under folding conditions. The propensity to form secondary structures on local scales is driven by the tendency of hydrophobic residues to be in proximity (proteins) and the formation of hairpins on scales where the persistence length is comparable to the screening length (RNA). The packing of the secondary structures gives rise to the tertiary fold which in RNA results in a hierarchical assembly (see however refs *57* and *58*). Although the secondary structures are locally favorable the many distinct ways of packing them result in multiple tertiary non-native folds. The incompatibility between stable structures *formed on local length scales and the native fold is termed "topological frustration"* (*25*, *55*). Even if *energetic frustration is totally eliminated* chain connectivity renders proteins and RNA topologically frustrated.

The rugged nature of the energy landscape can be used to anticipate many scenarios for folding. Of particular interest are sequences that fold by an apparent "all-or-none" process under a range of external conditions. Because of the relative simplicity of their folding mechanisms, they have been extensively investigated in the protein folding field (*59*). In addition, the energy landscape description also leads to the KPM. The KPM provides a unifying perspective of RNA and protein folding. Here we will focus on these two scenarios. Folding without barriers (downhill folding) may also be possible (*46*, *60*, *61*).

Apparent Two-State Folders and the Nucleation-Collapse (NC) Mechanism. Two-state folders are minimally frustrated so that the energy landscape is relatively smooth with superimposed roughness. There are two dominant basins of attraction, namely, the basin corresponding to unfolded states and the NBA. Under folding conditions, the unfolded state population decays exponentially with the time constant τ_F.

In the early stages of folding, the radius of gyration of the chain decreases rapidly on the collapse time scale τ_c. For two-state folders τ_F/τ_c is on the order of unity ($O(1)$) so that collapse and folding are nearly simultaneous (*27*). By $O(1)$ it is meant that $0 < \tau_F/\tau_c < (5-10)$. Fast folding experiments on a few proteins, which monitor folding from about 50 μs onward, are in accord with these arguments (*62*, *63*).

Majority of the two-state folders reach the NBA by a nucleation-collapse (or condensation) (NC) mechanism (*55*, *64*−*67*). According to the NC model the acquisition of the native fold is preceded by the formation of one of the folding nuclei, in which a fraction of interactions that stabilize the native structure is present. The transition to the NBA is rapid once the folding nuclei are formed. In this sense, the folding reaction is similar to the gas−liquid transition (*38*). However, there are profound difference in the nature of the folding nuclei due to the topological restrictions. In proteins, systematic computations show that the folding nuclei have a mixture of local and nonlocal contacts (*24*, *66*, *68*). Nonlocal contacts are required to stabilize distant parts of the protein because the secondary structural elements are not independently stable.

It is difficult to decipher the nature of the folding nuclei even for simple two-state folders (*69*, *70*). Theoretical arguments and computations have shown that either there is an extended nucleus in which virtually all of the residues form their nativelike contacts with some probability in the TSE (*71*) or there are multiple folding nuclei (MFN), which argues for a number of smaller nuclei (*55*, *71*). According to the MFN model, in each molecule certain contacts form with high probability in the transition state (> 0.5 say) than others. Mutations of these high probability contacts would lead to a redistribution of population in the TSE without totally disrupting the folding process. The tolerance to mutations at many sites shows that the TSE itself should be broad and plastic (*72*, *73*). This suggests that, in general, there ought to be MFN with considerable heterogeneity in the TSE structures.

Much less is known about structures in the TSE in RNA. Several recent experiments suggest that TSE in RNA must also be heterogeneous (*17*, *28*, *74*, *75*). The formation of independently stable secondary structure at very low counterion concentration and subsequent assembly into tertiary fold are expected to make the nature of TSE different in RNA than in proteins. Because neutralization of charges on phosphate by counterion−condensation is a prerequisite for forming tertiary contacts TSE in RNA may be conformationally more restricted than in proteins (*58*).

Pathway Diversity and the Kinetic Partitioning Mechanism (KPM). Because of the ruggedness of the energy landscape navigation to the NBA is impeded by pauses in kinetic traps. The presence of kinetic traps is exacerbated, especially for large RNA molecules (see below) (*31*, *76*−*79*) and proteins with complex folds. In these systems, the alternate misfolded structures (*25*) or overstabilized parts of the native substructure (*31*) retard folding. The structures in the competing basins of attraction (CBAs) could have many nativelike features that make them long-lived under folding conditions. From the schematic sketch of the rugged energy landscape (Figure 2) the basic notions of KPM can be obtained. Imagine the navigation process in which an ensemble of unfolded molecules (U) begins to traverse the rugged energy landscape

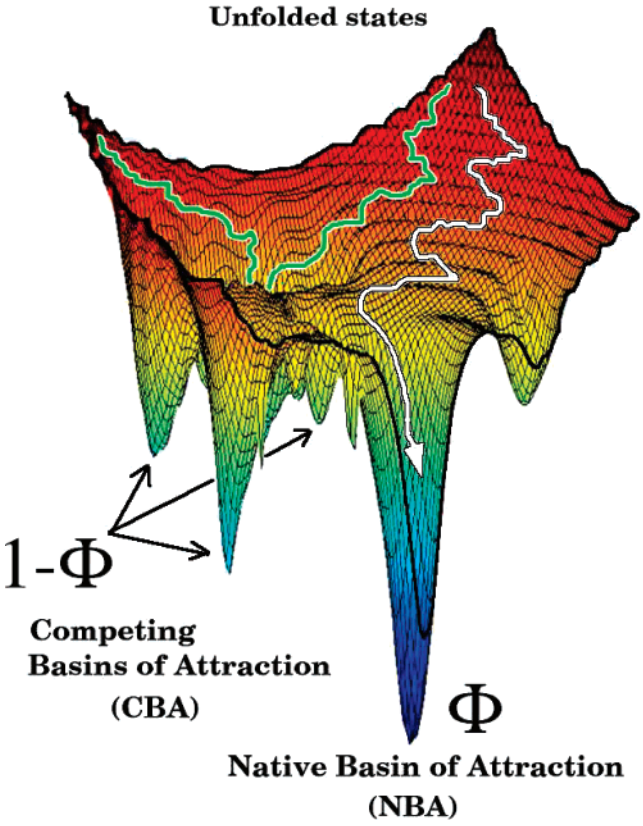

FIGURE 2: Schematic sketch of the rugged energy landscape underlying proteins and RNA that fold by the KPM. The entropically stabilized high free energy states are populated under unfolding conditions. Under folding conditions a fraction of molecules (Φ) reaches the NBA directly. A sketch of a trajectory for a fast track molecule that starts in a region of the energy landscape 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 ($1 - \Phi$). This small dimensional representation of the complex energy landscape suggests that the initial conditions, which can be changed by counterions, stretching force, or denaturants, can alter folding pathways.

in search of the NBA. The conformations in the U states are heterogeneous and span a range of vastly differing structures. A fraction Φ (referred to as the partition factor) can reach the NBA rapidly. These molecules fold rapidly without populating any discernible intermediates. The remaining fraction, $1 - \Phi$, is trapped in a manifold of discrete intermediates $\{I_{NS}\}$. Since the transitions from the CBAs to the NBA requires large-scale structural rearrangement for crossing the free energy barriers the folding of this class of molecules is slow. Thus, due to the multivalley structure of the free energy landscape the initial ensemble of molecules partitions into fast folders (Φ being their fraction) and slow folders. According to KPM, the fraction of molecules that reaches the native state $f_{NBA}(t) = 1 - \Phi e^{k_{fast}t} - \sum a_i e^{-k_i t}$ ($\Phi + \sum a_i = 1$) where k_{fast} is the rate for the fast process, k_i is the rate for transition from the discrete intermediates in the $\{I_{NS}\}$ ensemble to the NBA, and a_i is the corresponding amplitude.

The partition factor Φ, which gives the yield of fast track molecules, has been measured for a few biomolecules (Table 1). Because the energy landscape can be manipulated by mutation, addition of cosolvents, and counterions it follows that Φ also should respond to these changes. The value of

Table 1: Experimental Estimates of the Partition Factor Φ for Proteins and RNA that Fold by KPM

biomolecule	Φ[a]	τ_{fast}[b]
lysozyme	0.20−0.25[c]	∼50 ms
Tetrahymena ribozyme	0.06−0.10[d]	∼1 s[e]
U273A	0.80[f]	∼1 s
RuBisCo	0.05[g]	

[a] The fraction of fast track molecules that reach the NBA without getting kinetically trapped. [b] Time scale for the fraction of fast folding molecules. [c] Under the experimental conditions (acidic conditions and $T = 20$ °C) used by Kiefhaber (*142*) $\Phi = 0.20$, whereas $\Phi = 0.25$ at $T = 25$ °C and pH = 5.5 (*143*). At neutral pH Φ is expected to be unity. Thus, contradictory to reports in the literature (*144*), the mechanism of lysozyme folding can be altered by changing external conditions. These changes are also reflected in T_θ and T_f values which depend on sequence, pH, and salt concentration. [d] Value of Φ is taken from ref *80* for the precursor-RNA. Analysis of the single molecule data (*17*) (reported in ref *81*) for L-21 Sca I form of the group I intron gives $\Phi \approx 0.06$. [e] The time constant of ∼ 1 s may be the folding speed limit for the large group I intron. [f] The value of Φ for the U273A mutant of the group I intron is taken from ref *82*. [g] Data taken from ref *38*.

Φ for the *Tetrahymena* ribozyme (Figure 3), estimated from ensemble experiments and measured directly in single molecule studies, is about (6−10)% (*17*, *80*, *81*) (Table 1). Remarkably, a single point mutation U273A in P3 (see Figure 3) increases Φ to 80% (*82*). Similarly, analysis of single molecule FRET data L-21 Sca I construct of the group I intron shows that preincubation in excess monovalent cations can also alter Φ (*83*). From the KPM, it also follows that if Φ is small, so that a substantial fraction of molecules are trapped in the CBA for long times, then their transition to the NBA might be helped by chaperones (see below).

The fast track molecules fold by a NC-like mechanism with little pathway heterogeneity (*27*). The slow track molecules reach the NBA by a multistage process that involves formation of compact structures, search among the misfolded ensemble $\{I_{NS}\}$, and the transition to the NBA. Recent, single-molecule experiments of adenylate kinase trapped in liposomes (*14*) and surface-immobilized group I intron have provided direct evidence for the multistage assembly of biomolecules with complex native folds.

Folding Rates: Size Matters. Several factors contribute to folding rates of proteins and RNA. For proteins native topology, which is characterized by contact order (*84*, *85*), is important. The role of sequence length is often overlooked as a factor that influences the folding time τ_F. Given that proteins and RNA are "evolved" heteropolymers it is not surprising that N should play a crucial role in controlling k_F and the unfolding rate k_U. For minimally frustrated sequences, $\ln(k_F/k_F^0) \sim \alpha \ln N$ with $\alpha \approx 4$ at $T < T_F$ where the prefactor k_F^0 can be obtained using Kramers' theory (*27*, *86*). Because RNA and proteins are topologically frustrated, there is residual roughness even in two-state folders. As a consequence one finds that

$$\ln (\tau_F/\tau_F^0) \approx \Delta G_{UF}^{\ddagger}/k_B T \qquad (1)$$

where the effective free energy barrier $\Delta G_{UF}^{\ddagger}$ scales as $\Delta G_{UF}^{\ddagger}/k_B T \sim CN^{\beta}$ with $\beta < 1$. Using analogies to disordered systems, we (*27*) predicted that $\beta = 1/2$ while others have suggested that $\beta = 2/3$ (*64*, *87*). 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 perhaps of RNA (see below).

(a) Proteins. In two papers (*88*, *89*), the folding rates of 57 two- and three-state proteins have been analyzed using eq 1. The fits of $\ln k_F$ on N using the theoretically proposed models show that the effective free energy barrier indeed scales sublinearly with N. Both $\beta = 1/2$ or 2/3 fit the data equally well (*89*). The average values of the prefactor τ_F^0 for the two fits is in the range of 1 μs. The models yield an average value of $\tau_F^0 \gg \tau_{F,TST}^0 = h/k_B T$ (h is the Planck's constant). Recent experiments suggest $\tau_F^0 \approx 1$ μs (*90*).

(b) RNA. In contrast to proteins 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 rates using the available data in Table 2. Surprisingly, the rates that vary *over 7 orders*

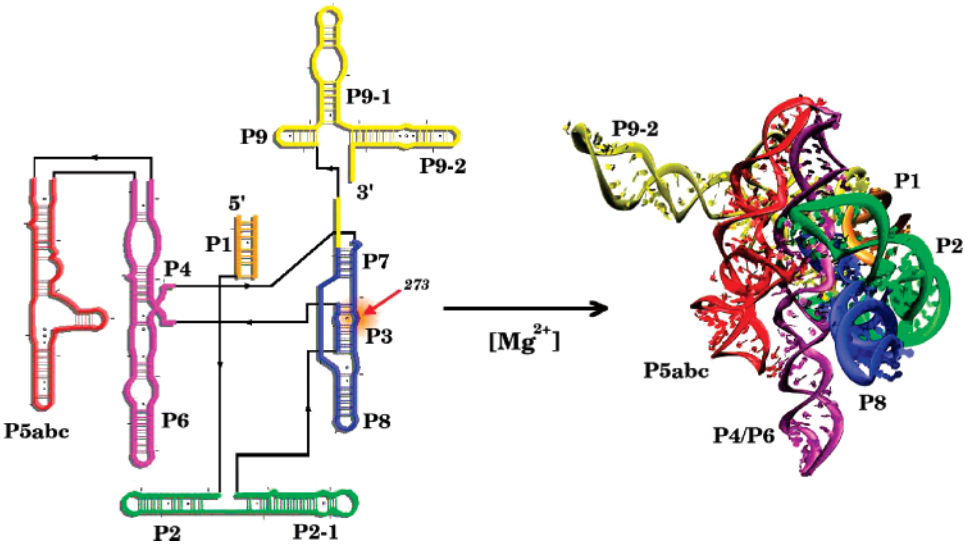

FIGURE 3: Secondary structure of the most extensively studied group I intron from *Tetrahymena*. The secondary structure has a number of paired helices indicated by P1 through P9. Upon addition of excess Mg^{2+} compact tertiary structure forms (shown on the right) by the catalytic core formed by an 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 (*152*). The structure on the right is a model proposed by Westhof and Michel (*153*).

Table 2: Folding Rates and Estimates of Free Energy Barrier for RNA[a]

	N[b]	k_f (s^{-1})[c]	$\Delta G_{UF}^{\ddagger KR}(k_BT)$[d]	$\Delta G_{UF}^{\ddagger TST}(k_BT)$[e]
extra-arm of $tRNA^{Ser}$ (yeast) (*145*)	21	1×10^5	3	18
pG-half of $tRNA^{Phe}$ (yeast) (*145*)	36	9×10^3	5	20
CCA-half of $tRNA^{Phe}$ (yeast) (*145*)	39	8.5×10^3	5	20
CCA-half of $tRNA^{Phe}$ (wheat) (*145*)	39	8×10^3	5	20
$tRNA^{Phe}$ (yeast) (*145*)	76	1×10^2	10	25
$tRNA^{Ala}$ (yeast) (*145*)	77	9×10^2	8	23
P5abc (*146*)	72	$28 \sim 50$	10	26
P4−P6 domain (*Tetrahymena* ribozyme) (*146*)	160	$1 \sim 2$	14	29
Azoarcus (*147*)	205	$10 \sim 20$	11	27
B. subtilis RNase P RNA catalytic domain (*148*)	225	6.5 ± 0.2	12	28
Ca.L-11 ribozyme (*149*)	368	0.03	18	33
E. coli RNase P RNA (*150*)	377	0.011 ± 0.001	19	34
B. subtilis RNase P RNA (*150*)	400	0.008 ± 0.002	19	34
Tetrahymena ribozyme (*151*)	414	0.02	18	34

[a] We do not include available data for pseudoknots. [b] N is the number of nucleotides. [c] k_f are the RNA folding rates. The rates were taken from the literature. The references are given in square brackets in the first column. [d] Free energy barriers are obtained using eq 1 with the prefactor $\tau_0^{KR} = 0.6$ μs. [e] Estimate of the transition state barrier using a value of the prefactor $\tau_0^{TST} = h/k_BT = 0.16$ ps.

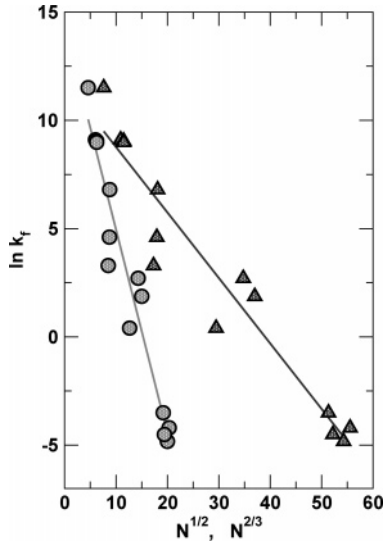

FIGURE 4: Fits of ln k_f as a function of N^β with $\beta = 1/2$ or 2/3 (ln $k_f = \ln 1/\tau_0 - CN^\beta$) where N is the number of nucleotides. Data points are taken from Table 2. For $\beta = 1/2$ (circles), $C = 0.94$, $\tau_0 = 0.58$ μs, and the correlation coefficient is 0.97. For $\beta = 2/3$ (triangles), $C = 0.30$, $\tau_0 = 7.5$ μs, and the correlation coefficient is 0.96. Both fits show that *the free barrier scales sublinearly with N*.

of magnitude depend on N (with $\beta = 1/2$ or 2/3) 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. This may be because the range of N in RNA changes by a factor of 20, whereas in the proteins that have so far been studied experimentally N varies over a much shorter range (*89*, *91*, *92*).

Using the results in Figure 4 the difficult-to-measure prefactor k_F^0, which should be estimated using Kramers' theory (*27*, *93*), can be calculated. Using the scaling plots in Figure 4 we find that $\tau_0^{KR} \approx 0.6$ μs for $\beta = 0.5$ and $\tau_0^{KR} \approx 8$ μs for $\beta = 2/3$. Both these estimates for the RNA folding prefactor are nearly *6 orders of magnitude larger than the transition state value*.

The large value of τ_0^{KR} implies that the effective free energy barriers from the measurements of rates alone using transition state prefactor overestimates the activation free energies. If the Kramers' prefactor is used the estimated barriers are considerably less (Table 2). The TST prefactor is only applicable if a single bond is involved in the transition state. While this may be appropriate for gas-phase reactions it cannot describe folding that is determined by collective events. The prefactor $k_F^0 \approx (\tau_F^{KR})^{-1}$ represents the rates at which folding can occur without barriers, i.e., by a diffusion-limited process. Using an estimate for the most elementary event in folding (formation of the most probable contact between amino acids or base pairing in RNA) we arrive at the Kramers' estimate of 1 μs for the prefactor. Our estimate is in accord with the typical base pairing rate (*94*, *95*).

Initial Events in RNA and Protein Folding. The most profound distinction between RNA and protein folding arises when considering the early events in the folding process. In both cases, the radius of gyration, R_g, decreases in the transition from the unfolded states to the native conformation. However, the nature of events that drives the collapse transition is *markedly different in the two problems*. In polypeptide chains the aversion of hydrophobic residues to water results in the chain forming compact structures, whereas in RNA collapse is mediated by nonspecific condensation of counterions (*96*−*98*).

(a) Collapse for Proteins. For certain proteins (*45*, *99*) and RNA, under "optimal" conditions (temperature, pH, counterions, etc.), the collapse transition may be "specific" (*100*). By specific, we mean that the mobile collapsed conformations contain many characteristics of the native state. Because these structures are near native, it is likely that the transition to the native state would occur rapidly after chain collapse. Using theoretical arguments, it was proposed that for sequences that reach the native conformations largely by the specific collapse process $\tau_F/\tau_c \sim O(1)$, i.e., τ_F does not exceed τ_c by more than an order of magnitude ($1 < \tau_F/\tau_c < 10$) (*27*). Just as predicted theoretically, the time scales separating collapse and folding in cytochrome *c* is small i.e., $\tau_F/\tau_C \sim O(1)$ (*101*). The temperature dependence of τ_c follows Arrhenius law that implies that there is substantial *enthalpy barrier early in the folding process* (*99*). The acquisition of the native structure and chain compaction occurs simultaneously. The fast folding experiments show that protein collapse occurs in multiple stages (*27*).

For proteins that fold by the KPM, chain collapse for a large fraction of molecules results in the formation of an ensemble of misfolded structures that contain a number of nonnative interactions. Such nonspecifically collapsed structures reach the native conformation by a three stage process (*44*). Because kinetic partitioning occurs early in the folding process, it is difficult to distinguish between specific and nonspecific collapse experimentally. Recent simulations have shown that there can be substantial non-native interactions on time scales prior to τ_c (*102*).

(b) Multiple Stages for RNA Collapse. The polyelectrolyte effects make chain compaction of even simple RNA molecules (tRNA for example) more complex than collapse of proteins. The phase diagrams in the (T, C) plane already suggests at least two intermediate phases can be populated in the transition from fully unfolded states (single strand) to the native conformation (Figure 1). In the intermediate concentration regime, $C_{cc} < C < C_m$, where C_{cc} is the concentration relevant for Manning condensation, we expect structures in which the electrostatic repulsions are muted but stable tertiary interactions are not formed. Subsequent increase in C results in the consolidation of the native structure. Minimally, RNA collapse must involve three distinct time scales. On the shortest time scale, τ_{cc}, counterions condense. Under folding conditions, τ_{cc} should be diffusion-limited so that $\tau_{cc} \approx (10^{-8}-10^{-6})$ s. In the second stage, local structures (perhaps secondary structures with mobile tertiary interactions) form because the electrostatic interactions are screened. The nature of the structures in this stage depends not only on the renormalized electrostatic interactions but also on other forces that stabilize the native state. This stage, which is characterized by time scale τ_c, results in the largest reduction in the radius of gyration (R_g) (*103*). The final stage involves consolidation of the tertiary interactions and a further decrease in R_g. In group I intron the compact structure that is formed at the end of the third stage is an ensemble with varying degree of native contact. Only a small fraction of the molecules has substantial native structure. Recent experiments have revealed the multi-stages in RNA collapse (*96, 97, 103, 104*)

Insight into the early events of RNA folding can be obtained by simulating collapse of strongly charged polyelectrolytes. Brownian dynamics simulations of polyelectrolytes (*105*) and theory (*106*) have shown that polyelectrolyte collapse occurs in three distinct stages just as in RNA. Using the analogy with counterion-dependent kinetics of collapse of polyelectrolytes we have estimated that $\tau_c \approx 0.4$ ms (*81*) which is consistent with experiments on L-21 Sca I form of the *Tetrahymena* ribozyme using time-resolved X-ray measurements (*103*).

ASSISTED FOLDING

Iterative Annealing Mechanism (IAM) for GroEL and RNA Chaperones. The KPM also suggests that molecular chaperones are sometimes required for biomolecular folding. Such cofactors are recruited when spontaneous folding under nonpermissive conditions does not lead to sufficient yield of the native state. Thus, whenever Φ is small, as is the case for in vitro folding of RuBisCo and *Tetrahymena* ribozyme, it is logical to expect that assisted folding may be required.

The *Escherichia coli* chaperonin machinery GroEL/GroES is known to assist folding of a number of cytosolic proteins. GroEL consists of two heptameric rings, stacked back-to-back, that communicate with one another and operate in a coordinated manner (Figure 5). The cofactor GroES, which has a complementary 7-fold symmetry, binds to GroEL only after ATP binding. The coordinated movements in GroEL triggered by binding of ATP, GroES, and SP allow the SP to reach the folded state (*39*, *107*). Although it has been appreciated for a long time that transient interactions with certain proteins help RNA fold (*108*), only recently several candidate RNA chaperones have been identified (*109*−*111*). Because the annealing mechanism of chaperonin system is much better understood, we first describe the essential steps in the interactions of GroEL/GroES system with the substrate protein (SP).

GroEL Nanomachine. The major role of the complete chaperonin system is to enhance the yield of the SPs. Only rarely is the folding rate increased by the chaperonins. Most often the rate is either unaffected or even decreased. To understand the global mechanism of GroEL-assisted protein folding it is sufficient to consider the single ring, which is the fundamental unit. In the hemicycle the single ring undergoes a series of four sequential events (Figure 5).

(a) Capture. The promiscuity of SP interactions with GroEL (*112*) suggests that various proteins in the non-native state must present a common binding motif. A structural feature common to almost all non-native state of proteins is an exposed hydrophobic surface that gets buried during the transition to the folded structure. A variety of experimental studies have shown that the favorable hydrophobic interaction is the major driving force for the recognition of SP by GroEL. Transient interaction with apical domain alone is sufficient to promote folding of nonstringent SPs (*113*). The mechanisms by which apical domain (referred to as minichaperone) alone can enable refolding (*114*) of a few SPs is similar to the action of a number of RNA chaperones (see Figure 6A) (*114*, *115*). However, the intact complete machinery is required for assisting SP, under nonpermissive conditions (*115*).

In the oligomeric construct of GroEL, the apical domain is repeated 7-fold so that the inner surface of the GroEL ring in the "resting" T state (which binds SP most tightly) presents a near continuous hydrophobic surface (Figure 5). This enables the exposed hydrophobic segment of the nonnative SP to interact favorably with the GroEL particle. Two conditions restrict the nature of SPs that can profit by interaction with GroEL. (i) The volume of GroEL in the T state is about 85 000 Å^3, and the diameter of the roughly cylindrical pore is 45 Å. Thus, only proteins with molecular mass less than about 30 kDa can be fully encapsulated in the cavity. Much larger proteins can also form stable complexes with GroEL. In such cases at least part of SP must be outside the cavity. (ii) If the stability of the SP−GroEL complex is less than a few k_BT the complex would not be stable enough for annealing action to commence. Similarly, the release of SPs from an exceedingly stable molecular complex would be too difficult. These restrictions impose a range of stability for the SP−GroEL complex for efficient function of the nanomachine (*116*).

The annealing action is intimately related to the domain movements that are triggered by binding of SP, ATP, and GroES. The T state has greater affinity for the SP than the R state. The transition to the R state, which is induced upon

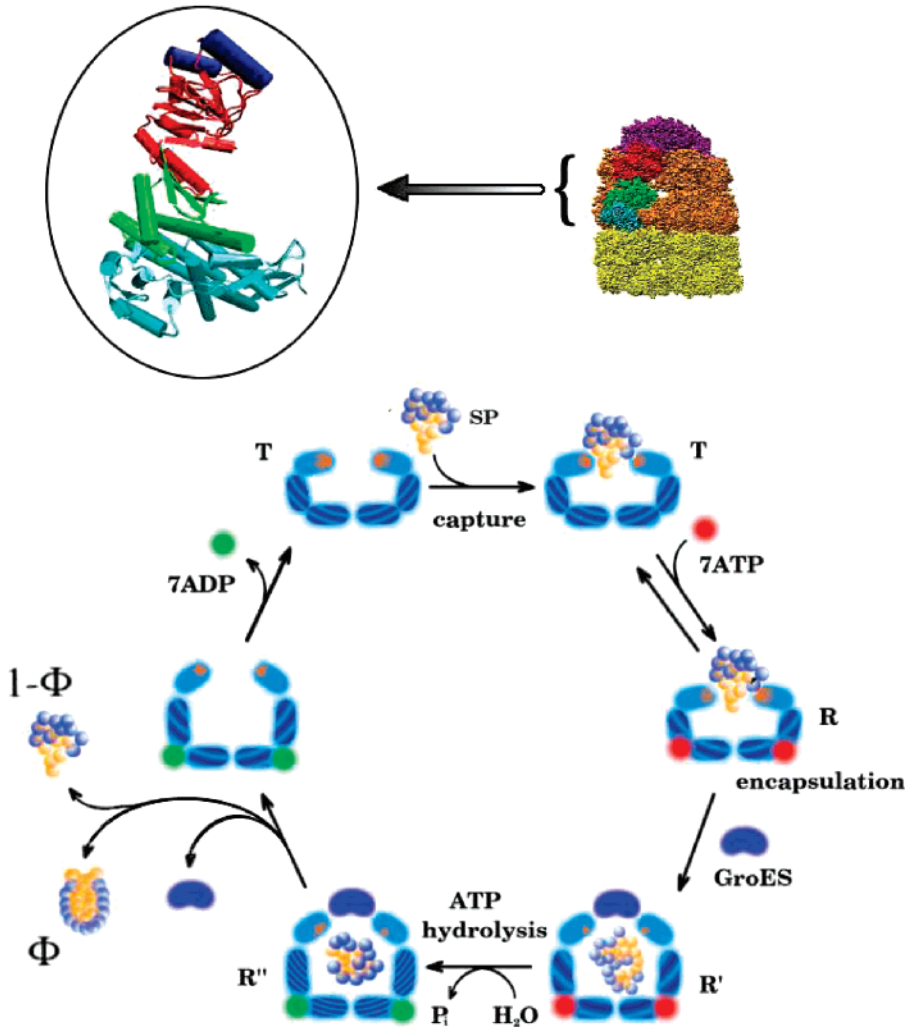

FIGURE 5: The top right shows the structure of the R″ (GroEL − $(ADP)_7$ − GroES) state. One of the GroEL subunits is shown in the circle on the top left. The apical domain is shown in red, the intermediate is in green, and the equatorial domain is in cyan. The hemicycle, which is completed in about 15 s at 37 °C, in the GroEL-assisted folding of proteins is shown in the bottom. For clarity only the steps in the ligand driven allosteric transitions in one of the rings (cis) is shown. In the initial step the substrate protein (SP) is captured by the GroEL in the T state. This step could induce minor conformational changes in GroEL. ATP binding triggers rigid body rotation of the intermediate domain toward the equatorial domain, leading GroEL to the R state. The R → R′ transition and GroES binding encapsulates the SP provided it is small enough to fit to the expanded cavity. After ATP hydrolysis the R′ → R″ transition takes place. The release of ADP, inorganic phosphate, and the SP (whether it is folded or not) is triggered by a signal from the trans ring (not shown). Only the T ↔ R is reversible. All other steps in the cycle are driven.

binding of ATP to the equatorial domain of GroEL, is entirely concerted. Simple geometric considerations show that the T ↔ R state is accompanied by the movement of the SP binding sites with nonadjacent ones moving farther than adjacent sites. The SP binding stimulates ATPase activity with the k_{cat} per subunit being about five times greater in the T state than in R state. Thus, the SP resists the T ↔ R transition. This suggests that in the process of T-to-R transition force is exerted on the SP (*117*), which implies that the annealing action of GroEL results in unfolding of the SP (*118*, *119*).

(b) Encapsulation. Upon encapsulation the SP goes from being bound to GroEL to a state in which it is confined in the volume permitted by the cavity (Figure 5). In the bound state the microenvironment experienced by the SP is largely hydrophobic, whereas in the sequestered state the SP is conformationally unrestrained because of weaker interactions with the GroEL cavity. Encapsulation is accompanied by a series of allosteric transitions in the GroEL which constitute the fundamental power stroke in the chaperonin cycle. The binding of MgATP triggers the domain movements that are exaggerated in the presence of GroES. The encapsulation process increases the inner volume of the cavity to about 185 000 $Å^3$. The polarity of the surface of the central cavity undergoes a dramatic change from being hydrophobic in the T state to hydrophilic in the R states. It remains so in this state until reverse domain movements return GroEL to the T state. The switching from the hydrophobic to hydrophilic surfaces that occurs in each *hemicycle results in the unfolding of the SP*. This event puts the SP in a higher point in the free energy landscape from which it can partition either to the folded state or be trapped in another misfolded conformation.

(c) ATP Hydrolysis. As a result of encapsulation the ATP molecules, which are locked in the active site, are committed to hydrolysis. At in vivo concentrations of ATP all seven ATP are hydrolyzed in a "quantized" manner in the presence of GroES (*120*). The hydrolysis of ATP serves as a timing

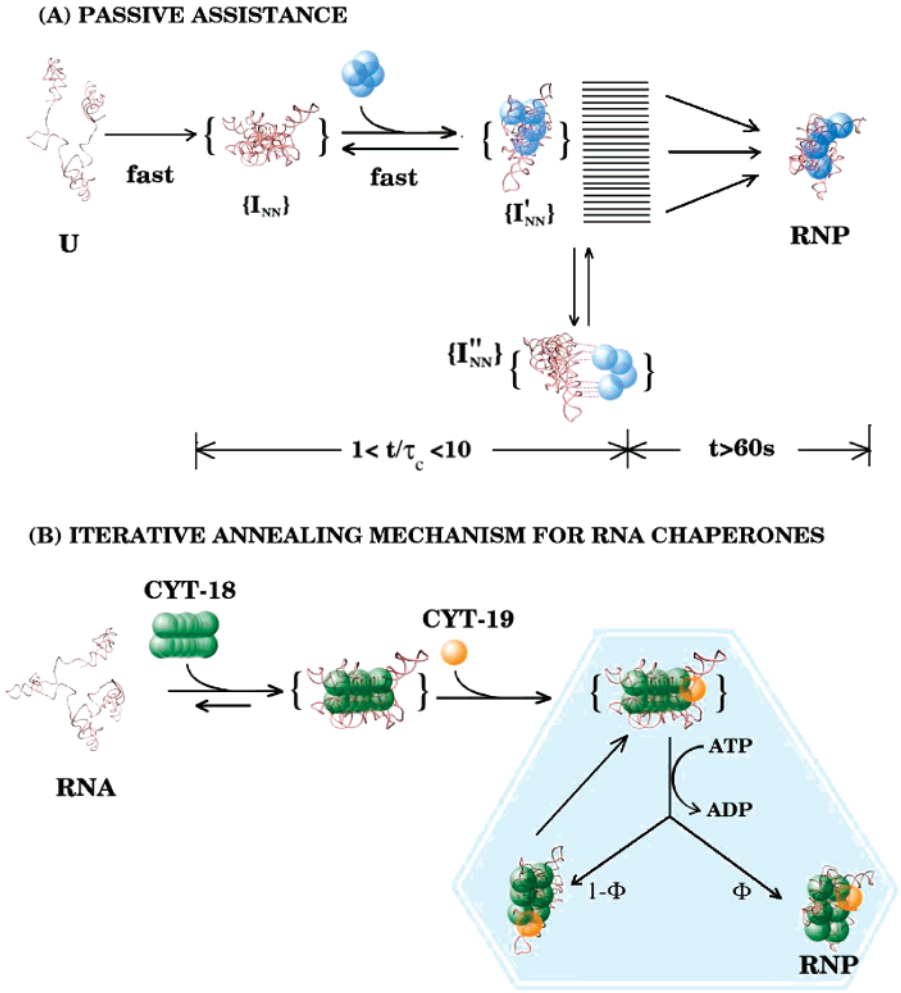

FIGURE 6: (A) A generalized tertiary capture model for protein-assisted (CBP2) RNA folding. This scheme closely resembles the one proposed by Garcia and Weeks (*136*). In this "passive" model the unfolded RNA rapidly collapses upon addition of multivalent counterions (valence greater than one). Protein (in blue) that binds by a diffusion-limited reaction leads to an ensemble of fluctuating protein−RNA complex. Protein binding restricts fluctuations in RNA. The structures in the protein−RNA can transiently unbind (shown as $\{I''_{NN}\}$), which allows the RNA to fold in a restricted volume. This process of binding and transient release (the annealing action) leads to the assembly competent RNP on long time scale. (B) Iterative annealing mechanisms for RNA chaperones. The binding of CYT-18 to misfolded RNA structures produce an ensemble of CYT-18-RNA complex. This step is similar to the events leading to $\{I'_{NN}\}$ in panel A. Binding of CYT-19 (shown in the second reaction step) to the mobile CYT-18−RNA structures results in a change in the structure of RNA. The free energy of ATP hydrolysis can place the RNA in a higher region of a free energy landscape from which it can kinetically partition to the RNP (probability in Φ) or can reach another CYT-18−RNA−CYT-19 complex (probability 1 − Φ). The cycle is iterated (blue region) until sufficient yield is obtained.

devise that controls the lifetime of the encapsulated state. Just as in other molecular machines (molecular motors or F1-ATPase) the power stroke of the nanomachine is associated with binding of the cofactors. Mutations in GroEL and addition of electrolytes can presumably alter the lifetime of the R″ state which in turn can have a profound effect on the annealing function of GroEL. In the extreme case, the single ring mutant (SR1) can artificially lock GroEL into the R″ state which in effect enables the SP to fold under confinement. Note that even in the SR1 mutant, with infinite lifetime for the R″ state (Figure 5), the SPs microenvironment changes once which is sufficient to enable near complete folding of even stringent SPs (RuBisCO for example) (*121*).

(d) Ligand Release and Domain Relaxation. The binding of another SP molecule to the distal trans ring takes place only after ATP hydrolysis in the cis ring. The release of all three ligands GroES, SP (folded or not), and ADP is triggered by ATP binding to the trans state. The rate of release of the ligands is a complex function of the concentration of ATP and whether a SP binds to the trans ring. Upon release of the ligands from the cis ring GroEL relaxes to the starting T state. The hemicycle then starts in the trans ring. In one round or iteration of binding, encapsulation, and release the SP reaches the native with an yield Φ. If Φ is small then the process is iterated several times until a sufficient amount of native state is generated. The number of iterations needed to obtain sufficient yield of the native state depends on a variety of factors related to the coupling of the dynamics of allosteric transitions and the rates of SP folding.

Scenarios for Protein-Assisted RNA Folding. The principal role of chaperones is to assist in the resolution of the multitude of alternative misfolded structures so that sufficient yield of the native material is realized in biologically viable time. A number of studies have implicated proteins in RNA folding. (1) Since the pioneering studies by Karpel and Fresco (*108*) it has been realized that proteins can facilitate in renaturating RNA. (2) Self-splicing of *Tetrahymena* pre-

rRNA occurs nearly 50 times faster in *Tetrahymena* cells than under in vitro conditions (*122*, *123*). (3) Interestingly, splicing of the intron in *E. coli* is also as fast as in *Tetrahymena* cells (*124*, *125*), which suggests that species specific mechanism is not operative in cellular RNA folding. It is likely, as model in vitro RNA assembly shows, that faster cellular splicing must be assisted by proteins or other cofactors. Unlike the well-identified chaperonin system for proteins there does not appear to be a "one-size fits all" protein that acts as RNA chaperones for a diverse class of RNA molecules.

For purposes of discussion let us classify RNA molecules that require proteins (or other RNAs) to fold into two classes. Although this is an oversimplification it allows us to consider assisted folding in two distinct limits. The class I RNAs may be those whose in vitro folding is greatly aided by protein cofactors but their assistance is not strictly required. An example is mitochondrial bI5 group I intron which can fold in 50 mM or greater Mg^{2+} at room temperature but interactions with cytochrome *b* pre-mRNA processing protein 2 (CBP2) allows folding at physiological levels ($\approx$7 mM) of Mg^{2+} (*126*, *127*). The well-studied *Tetrahymena* ribozyme and other group I introns may belong to stringent class II substrates by which we mean spontaneous folding, even at high counterion concentration is too slow, to be biologically viable. For this class of RNA molecules "active" interaction with protein cofactor (*128*, *129*), including perhaps ATP hydrolysis (*128*), is required to produce splicing competent structures. Similarly, proteins that interact with RNA to assist folding may also be classified into two classes. Proteins such as CBP2 appear to interact specifically with only a subgroup of group I introns whereas *Neurospora Crassa* mitochondrial tyrosyl tRNA synthetase (CYT-18) is more promiscuous (*130*, *131*). These observations suggest that the details of ribonucleoprotein assembly can vary greatly depending upon the RNA sequence and on the nature of interactions with the protein.

Because of the variations in the interactions of the proteins with the RNA there are possibly several scenarios of protein-assisted RNA folding. At a global level, it appears that the IAM (or a variant) for chaperonin-mediated folding can also be used to explain the role of RNA chaperones. Following the observations in an insightful note by Herschlag (*109*), who was the first to draw analogies to protein chaperones, we envision two distinct scenarios for protein-assisted RNA assembly.

Passive Assistance. A key advance in understanding protein-assisted ribonuleoprotein assembly was made in the studies initiated by Weeks and Cech (*126*) who probed splicing reactions in bI5 group I intron from yeast mediated by CBP2. The initial studies (*126*, *127*) and subsequent reports by Weeks and co-workers (*132*, *133*) showed that efficient folding of the RNA occurs by a tertiary structure capture mechanism (*134*, *135*). According to this model, CBP2 binds to a partially folded nativelike RNAs. The two-step assembly of the RNP complex (*126*), namely, the formation of the catalytic core by association of the conserved P5−P4−P6 and P7−P3−P8 domains of group I intron (Figure 3), and subsequent assembly of 5′ domain of bI5 core, can also be achieved by the noncognate *Neurospora crassa* mitochondrial tyrosyl tRNA synthetase (CYT-18 proteins) (*130*). In contrast to CBP2 a stable CYT-18-bI5 complex can form early in the RNA folding process. Thus, the promotion of splicing competent structures by CYT-18 and CBP2 may occur by a similar process but could differ in the assembly pathways (*130*).

A generalization of the tertiary structure capture model, which is very similar to the mechanism proposed by Garcia and Weeks (*136*), is necessary based on the following considerations. (1) Theory (*81*) and experiments (*103*) suggest that group I introns (L-21 Sca I for example) form compact structures on 0.1 s time scale. (2) Buchmuller and Weeks (*133*) showed that CBP2 binds to bI5 intron even at low (<3 mM) Mg^{2+} concentrations, which is sufficient to produce an ensemble of collapsed RNA structures. The revised model is based on the premise that the rapid collapse of RNA produces a manifold of compact structures *all of which are targets for binding to CBP2*. Thus, on time scales less than about 2 min, the estimated time for CBP2-assisted nativelike structure to form at 7 mM Mg^{2+} (over a wide temperature range) (*130*), an ensemble of fluctuating CBP2-bI5 complex can form (Figure 6A). The stabilities of these structures can vary and some of them can rapidly interconvert among each other. The structures probably differ only in the interface between CBP2 and the 5′ domain of the intron intron.

If a manifold of entropically stabilized CBP2-bI5 complex forms rapidly how does CBP2, at long times, facilitate folding at low Mg^{2+} concentration? If the association between CBP2 and the collapsed bI5 structures is too weak then the large bI5 conformational fluctuations can produce long-lived metastable kinetic traps. In this situation CBP2 or CYT-18 would have little effect on RNA folding. In the opposite limit, transient unfolding in the bI5 conformations, which may be needed for the transition from misfolded structures to the NBA, would be prohibited. Thus, an optimal stability of the CBP2-bI5 complex is needed to produce efficient RNA folding. In this picture, CBP2 essentially reduces the entropic barrier to RNA folding by dynamically "confining" it to a smaller folding volume. In other words, the protein cofactor "crowds" the RNA that results in an increase in stability and folding rates. An immediate consequence of this model is that enhancing the interaction between protein and RNA (say, by altering the nature of counterions or suitable mutations in CBP2 or RNA) would alter folding efficiency.

The entropy barrier version of the tertiary capture mechanism can be tested using single molecule FRET experiments that can monitor the RNA states starting from collapse time scales $\tau_c \approx 0.1$ s. If this picture is correct, we would expect large fluctuation in the time dependence of the FRET efficiency. A distribution of FRET values, characteristic of collapsed conformations, would be evidence of manifold of compact states that can interact with the protein. Although we have described CBP2-assisted bI5-folding in detail, we expect a similar short time description to hold good for CYT-18 induced folding of group I introns.

"*Active*" *Role for Proteins in RNA Folding.* Recent experiments show that proteins play an active role in rescuing kinetically trapped RNAs. The *E. coli* protein StpA assists folding of thymidylate synthase (td) group I intron by partially unfolding a compact RNA structure (*129*). In an important set of experiments Lambowitz and co-workers (*128*) showed CYT-18 and the DEAD-box protein CYT-19

(belonging to the ATPase family) work in a coordinated manner by utilizing ATP hydrolysis to promote splicing of *Neurospora crassa* group I intron. While the CYT-18 alone can rescue *N. crassa* mt group I intron at 37 °C, CYT-19 is required for efficient splicing at the normal growth temperature (25 °C). The "unfoldase" activity of StpA and temperature sensitive ATP-dependent nonspecific binding of CYT-19 is reminiscent of the way the GroEL machinery operates in resolving misfolded states of proteins (*137*).

The iterative annealing mechanism can be generalized to include the participation of protein-assisted assembly of group I intron (Figure 6B). We envision CYT-18 binding to the RNA to produce a manifold of conformations (Figure 6B) just as in Figure 6A. Subsequently, CYT-19 can bind to high free energy regions of the structure in such a way as to unfold incorrectly formed structures. Protein binding to single-stranded (high free energy regions) RNA mechanism was first proposed by Karpel and Fresco (*108*) who showed that helix-destabilizing proteins can promote RNA folding. Upon ATP binding and hydrolysis the RNA is placed in a high free energy state from which it can either partition, with probability Φ, either to the splicing competent state or populate a non-native conformation with $(1 - \Phi)$ probability (Figure 6B). If Φ is small then this process is iterated until sufficient yield of the RNA complex is obtained. The IAM for RNA chaperone activity shows that the yield of the splicing competent complex Ψ_N after n iteration is (*39*).

$$\Psi_N = 1 - (1 - \Phi)^n \qquad (2)$$

Two predictions of the model are the following: (a) The stoichiometry of ATP binding is not known. If we assume that in each iteration one ATP is consumed then for the experimental conditions in Mohr et al. (see Figure 5A of ref *128*). For $\Phi \approx 0.1$ $n \sim 7$ for $\Psi_N = 0.5$. Alternatively, if the number of ATP consumed in each iteration is known then the turnover time (events in aqua blue regime in Figure 6B) can be computed by measuring the time dependent accumulation of the excised intron. (b) If Φ is altered by point mutations then the number of iteration needed for equivalent yield should decrease. For the U273A mutant in group I intron for which $\Phi \approx 0.8$ two iterations are sufficient to drive the folding reaction to near completion.

In modeling protein-assisted RNA folding the conformational changes in proteins upon interactions with RNA are usually not discussed. The probability of diffusion-limited binding of protein by RNA is greatest if the capture cross-section is large. This argues in favor of binding-induced partial unfolding of both proteins and RNA prior to complex formation. In the context of the model in Figure 6 it is likely that the binding-induced conformational fluctuations in both protein and RNA persist at all times. Fluorescent labels on the protein also would be required to reveal the extent of its RNA-induced conformational change.

CONCLUSIONS

Even at the sequence level a polynucleotide is different from a polypeptide. RNA is made from only four nucleotides whereas the number of amino acids from which proteins are synthesized is twenty. Beyond these numerical differences there are profound variations in the chemical character of the building blocks. The backbone of RNA is hydrophilic just as in proteins. The chemical character of the "side chain" of RNA, namely, the bases are predominantly hydrophobic although they can form interbase hydrogen-bonds. In contrast, protein side chains have greater chemical diversity (hydrophobic, polar, and charged). In this sense RNA is closer to a homopolymer than proteins. The "homopolymer" nature results in RNA sequence to adopt favorable alternative structures (*138*). This contributes to the intrinsic ruggedness of the energy landscape compared to proteins. The homopolymer character of RNA makes their design and structure prediction even more difficult than the already complex protein folding problems. At a first glance it might seem that the constraint of Watson−Crick pairing and the inherent stability of RNA secondary structures (*9*) might make RNA folding manageable. However, database analysis shows that roughly a half of the base pairs are involved in non-WC structures (bulges, loops, etc.) (*139*). In addition, the role of valence and shape of cations is modulating RNA secondary structures and possibly altering them in the course of tertiary folding are difficult to anticipate (*57*, *140*). The homopolymer nature can lead to a rugged energy landscape, even at the secondary structure level (*141*), even for modest size polynucleotide sequences. Despite these difficulties it is remarkable that, at some global level, the principles of protein self-organization are also applicable to RNA folding. We anticipate that the cross-fertilization of ideas will lead to conceptual framework for understanding not only folding but also RNP assembly and protein−protein interactions.

ACKNOWLEDGMENT

One of us (D.T.) is indebted to D. K. Klimov, G. H. Lorimer, and S. A. Woodson for collaboration on protein and RNA folding. We acknowledge K. M. Weeks for useful discussions on RNP assembly. The remarks of W. A. Eaton and the reviewers have been particularly helpful.

REFERENCES

1. Fersht, A. R. (1999) *Structure and Mechanism in Protein Science*, Freeman, New York.
2. Doudna, J., and Cech, T. (2002) The chemical repertoire of natural ribozymes, *Nature 418*, 222−228.
3. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) Self-splicing RNA − Auto-excision and auto-cyclization of the ribosomal-RNA intervening sequence of *Tetrahymena*, *Cell 31*, 147−157.
4. Cech, T. R., Zaug, A. J., and Grabowski, P. J. (1981) *In vitro* splicing of the ribosomal-RNA precursor of *Tetrahymena*-involvement of a quianosine nucleotide in the excision of the intervening sequence, *Cell 27*, 487−496.
5. Guerrier-Takada, C., and Altman, S. (1984) Catalytic activity of an RNA molecule prepared by transcription in vitro, *Science 223*, 285−286.
6. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of Ribonuclease-P is the catalytic subunit of the enzyme, *Cell 35*, 849−857.
7. Prusiner, S. B. (1998) Prions, *Proc. Natl. Acad. Sci. U.S.A. 95*, 13363−13383.
8. Selkoe, D. J. (2003) Folding proteins in fatal ways, *Nature 426*, 900−904.
9. Tinoco, I., Jr., and Bustamante, C. (1999) How RNA folds, *J. Mol. Biol. 293*, 271−281.
10. Brion, P., and Westhof, E. (1997) Hierarchy and dynamics of RNA folding, *Annu. Rev. Biophys. Biomol. Struct. 26*, 113−137.
11. Eaton, W. A., Muñoz, V., Hagen, S. J., Jas, G. S., Lapidus, L. J., Henry, E. R., and Hofrichter, J. (2000) Fast Kinetics and Mechanisms in Protein Folding, *Annu. Rev. Biophys. Biomol. Struct. 29*, 327−359.

12. Eaton, W. A., Muñoz, V., Thompson, P. A., Henry, E. R., and Hofrichter, J. (1998) Kinetics and dynamics of loops, alpha-helices, beta-hairpins, and fast-folding proteins, *Acc. Chem. Res. 31*, 745−753.
13. Fersht, A. R., and Daggett, V. (2002) Protein Folding and Unfolding at Atomic Resolution, *Cell 108*, 573−582.
14. Rhoades, E., Gussakovsky, E., and Haran, G. (2003) Watching Proteins Fold One Molecule at a Time, *Proc. Natl. Acad. Sci. U.S.A. 100*, 3197−3202.
15. Deniz, A. A., Laurence, T. A., Beligere, G. S., Dahan, M., Martin, A. B., Chemla, D. S., Dawson, P. E., Schultz, P. G., and Weiss, S. (2000) Single-molecule protein folding: Diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5179−5184.
16. Lim, M., Hamm, P., and Hochstrasser, R. M. (1998) Protein fluctuations are sensed by stimulated infrared echoes of the vibrations of carbon monoxide and azide probes, *Proc. Natl. Acad. Sci. U.S.A. 95*, 15315−15320.
17. Zhuang, X., Bartley, L., Babcock, A., Russell, R., Ha, T., Hershlag, D., and Chu, S. (2000) A single-molecule study of RNA catalysis and folding, *Science 288*, 2048−2051.
18. Zhuang, X., and Reif, M. (2003) Single-molecule folding, *Curr. Opin. Struct. Biol. 13*, 86−97.
19. Yang, H., Luo, G., Karnchanaphanurach, P., Louie, T., Rech, I., Cova, S., Xun, L., and Xie, X. S. (2003) Protein conformational dynamics probed by single-molecule electron transfer, *Science 302*, 262−266.
20. Onuchic, J. N., and Wolynes, P. G. (2004) Theory of protein folding, *Curr. Opin. Struct. Biol. 14*, 70−75.
21. Onuchic, J., Luthey-Schulten, Z., and Wolynes, P. G. (1997) Theory of Protein Folding: The Energy Landscape Perspective, *Annu. Rev. Phys. Chem. 48*, 539−600.
22. Dill, K. A., and Chan, H. S. (1997) From Levinthal to Pathways to Funnels, *Nat. Struct. Biol. 4*, 10−19.
23. Thirumalai, D., and Klimov, D. K. (1999) Deciphering the Time Scales and Mechanisms of Protein Folding Using Minimal Off-Lattice Models, *Curr. Opin. Struct. Biol. 9*, 197−207.
24. Mirny, L., and Shakhnovich, E. (2001) Protein Folding Theory: From Lattice to All-Atom Models, *Annu. Rev. Biophys. Biomol. Struct. 30*, 397−420.
25. Thirumalai, D., and Woodson, S. A. (1996) Kinetics of Folding of Proteins and RNA, *Acc. Chem. Res. 29*, 433−439.
26. Thirumalai, D., Klimov, D. K., and Woodson, S. A. (1997) Kinetic partitioning mechanism as a unifying theme in the folding of biomolecules, *Theor. Chem. Acc. 96*, 14−22.
27. Thirumalai, D. (1995) From Minimal Models to Real Proteins: Time Scales for Protein Folding Kinetics, *J. Phys. I (Fr.) 5*, 1457−1467.
28. Pan, J., Thirumalai, D., and Woodson, S. A. (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.
29. Draper, D. E. (1996) Parallel worlds, *Nat. Struct. Biol. 3*, 397−400.
30. Rook, M. S., Treiber, D. K., and Williamson, J. R. (1998) Fast folding mutants of the *Tetrahymena* group I ribozyme reveal a rugged folding energy landscape, *J. Mol. Biol. 281*, 609−620.
31. Treiber, D., and Williamson, J. (2001) Beyond kinetic traps in RNA folding, *Curr. Opin. Struct. Biol. 11*, 309−314.
32. Ferré-D'Amaré, A. R., and Doudna, J. A. (1999) RNA FOLDS: Insights from Recent Crystal Structures, *Annu. Rev. Biophys. Biomol. Struct. 28*, 57−73.
33. Branden, C., and Tooze, J. (1991) *Introduction to Protein Structure*, Garland Publishing, New York.
34. Koculi, E., Lee, N., Thirumalai, D., and Woodson, S. A. (2004) Folding of the *Tetrahymena* Ribozyme by Polyamines: Importance of Counterion Valence and Size, *J. Mol. Biol. 341*, 27−36.
35. Yang, W. Y., Pitera, J. W., Swope, W. C., and Gruebele, M. (2004) Heterogeneous Folding of the trpzip Hairpin: Full Atom Simulation and Experiment, *J. Mol. Biol. 336*, 241−251.
36. Snow, C. D., Nguyen, H., Pande, V. S., and Gruebele, M. (2002) Absolute comparison of simulated and experimental protein-folding dynamics, *Nature 420*, 102−106.
37. Mayor, M., Guydosh, N. R., Johnson, C. M., Grossmann, J. G., Sato, S., Jas, G. S., Freund, S. M. V., Alonso, D. O. V., Daggett, V., and Fersht, A. (2003) The complete folding pathway of a protein from nanoseconds to microseconds, *Nature 421*, 863−867.
38. Pande, V. S., Grosberg, A. Y., Tanaka, T., and Rokhsar, D. S. (1998) Pathways for protein folding: is a new view needed? *Curr. Opin. Struct. Biol. 8*, 68−79.
39. Todd, M. J., Lorimer, G. H., and Thirumalai, D. (1996) Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism, *Proc. Natl. Acad. Sci. U.S.A. 93*, 4030−4035.
40. Sosnick, T., and Pan, T. (2003) RNA folding: models and perspectives, *Curr. Opin. Struct. Biol. 13*, 309−316.
41. Dill, K. A. (1990) Dominant Forces in Protein Folding, *Biochemistry 29*, 7133−7155.
42. Socci, N. D., and Onuchic, J. N. (1995) Kinetic and Thermodynamic Analysis of Protein Like Heteropolymers: Monte Carlo Histogram Technique, *J. Chem. Phys. 103*, 4732−4744.
43. Bryngelson, J. D., and Wolynes, P. G. (1990) A Simple Statistical Field Theory of Heteropolymer Collapse with Application to Protein Folding, *Biopolymers 30*, 177−188.
44. Camacho, C., and Thirumalai, D. (1993) Kinetics and Thermodynamics of Folding in Model Proteins, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6369−6372.
45. Plaxco, K. W., Millett, I. S., Segel, D. J., Doniach, S., and Baker, D. (1999) Chain collapse can occur concomitantly with the rate-limiting step in protein folding, *Nat. Struct. Biol. 6*, 554−556.
46. Bryngelson, J., Onuchic, J., Socci, N., and Wolynes, P. (1995) Funnels, pathways, and the energy landscape of protein-folding − A synthesis, *Proteins: Struct., Funct. Genet. 21*, 167−195.
47. Rich, A., and Schimmel, P. R. (1977) Introduction to Transfer RNA, *Acc. Chem. Res. 10*, 385−387.
48. Cole, P. E., Yang, S. K., and Crothers, D. M. (1972) Conformational Changes of Transfer Ribonucleic Acid. Equilibrium Phase Diagrams, *Biochemistry 11*, 4358−4368.
49. Stein, A., and Crothers, D. M. (1976) Equilibrium Binding of Magnesium(II) by *Escherichia coli* tRNA, *Biochemistry 15*, 157−160.
50. Stein, A., and Crothers, D. M. (1976) Conformational Changes of Transfer RNA. The Role of Magnesium(II), *Biochemistry 15*, 160−168.
51. Thirumalai, D., and Woodson, S. A. (2000) Maximizing RNA folding rates: a balancing act, *RNA 6*, 790−794.
52. Bryngelson, J. D., and Wolynes, P. G. (1987) Spin glasses and the statistical mechanics of protein folding, *Proc. Natl. Acad. Sci. U.S.A. 84*, 7524−7528.
53. Mézard, M., Parisi, G., and Virasoro, M. (1988) *Spin Glass Theory and Beyond*, World Scientific, River Edge, NJ.
54. Thirumalai, D., Ashwin, V., and Bhattacharjee, J. K. (1996) Dynamics of Random Hydrophobic−Hydrophilic Copolymers with Implications for Protein Folding, *Phys. Rev. Lett. 77*, 5385−5388.
55. Guo, Z., and Thirumalai, D. (1995) Kinetics of Protein Folding: Nucleation Mechanism, Time Scales, and Pathways, *Biopolymers 36*, 83−102.
56. Chavez, L. L., Onuchic, J. N., and Clementi, C. (2004) Quantifying the Roughness on the Free Energy Landscape: Entropic Bottlenecks and Protein Folding Rates, *J. Am. Chem. Soc. 126*, 8426−8432.
57. Wu, M., and Tinoco, I., Jr. (1998) RNA folding causes secondary structure rearrangement, *Proc. Natl. Acad. Sci. U.S.A. 95*, 11555−11560.
58. Thirumalai, D. (1998) Native secondary structure formation in RNA may be a slave to tertiary folding, *Proc. Natl. Acad. Sci. U.S.A. 95*, 11506−11508.
59. Jackson, S. E. (1998) How do proteins fold? *Folding Des. 3*, R81−R91.
60. Garcia-Mira, M. M., Sadqi, M., Fischer, N., Sanchez-Ruiz, J. M., and Muñoz, V. (2002) Experimental Identification of Downhill Protein Folding, *Science 298*, 2191−2195.
61. Sabelko, J., Ervin, J., and Gruebele, M. (1999) Observation of strange kinetics in protein folding, *Proc. Natl. Acad. Sci. U.S.A. 96*, 6031−6036.
62. Teilum, K., Maki, K., Kragelund, B. B., Poulsen, F. M., and Roder, H. (2002) Early kinetic intermediate in the folding of acyl-CoA binding protein detected by fluorescence labeling and ultrarapid mixing, *Proc. Natl. Acad. Sci. U.S.A. 99*, 9807−9812.
63. Akiyama, S., Takahashi, S., Kimura, T., Ishimori, K., Morishima, I., Nishikawa, Y., and Fujisawa, T. (2002) Conformational landscape of cytochrome *c* folding studied by microsecond-resolved small-angle X-ray scattering, *Proc. Natl. Acad. Sci. U.S.A. 99*, 1329−1334.

64. Wolynes, P. G. (1997) Folding funnels and energy landscapes of larger proteins within the capillarity approximation, *Proc. Natl. Acad. Sci. U.S.A. 94*, 6170−6175.
65. Fersht, A. R. (1997) Nucleation mechanisms in protein folding, *Curr. Opin. Struct. Biol. 7*, 3−9.
66. Li, L., Mirny, L. A., and Shakhnovich, E. I. (2000) Kinetics, thermodynamics and evolution of non-native interactions in a protein folding nucleus, *Nat. Struct. Biol. 7*, 336−342.
67. Qi, P. X., Sosnick, T. R., and Englander, S. W. (1998) The burst phase in ribonuclease A folding and solvent dependence of the unfolded state, *Nat. Struct. Biol. 5*, 882−884.
68. Guo, Z., and Thirumalai, D. (1997) The Nucleation-Collapse Mechanism in Protein Folding: Evidence for the Non-Uniqueness of the Folding Nucleus, *Folding Des. 2*, 423−442.
69. Thirumalai, D., and Klimov, D. K. (1998) Fishing for folding nuclei in lattice models and proteins, *Folding Des. 3*, 481−496.
70. Shakhnovich, E. I. (1998) Folding nucleus: Specific or multiple? Insight from lattice models and experiments, *Folding Des. 3*, R108−R111.
71. Onuchic, J. N., Socci, N. D., Luthey-Schulten, Z., and Wolynes, P. G. (1996) Protein folding funnels: The nature of the transition state ensemble, *Folding Des. 1*, 441−450.
72. Lindberg, M., Tngrot, J., and Oliveberg, M. (2002) Complete change of the protein folding transition state upon circular permutation, *Nature Struct. Biol. 9*, 818−822.
73. Muñoz, V. (2002) Folding plasticity, *Nat. Struct. Biol. 9*, 792−794.
74. Zhuang, X., Kim, H., Pereira, M., Babcock, H., Walter, N., and Chu, S. (2002) Correlating Structural Dynamics and Function in Single Ribozyme Molecules, *Science 296*, 1473−1476.
75. Koculi, E., Thirumalai, D., and Woodson, S. A. Counterion charge density determines the position and plasticity of RNA folding transition states, submitted to *Proc. Natl. Acad. Sci. U.S.A.*
76. Zarrinkar, P. P., and Williamson, J. R. (1994) Kinetic Intermediates in RNA Folding, *Science 265*, 918−924.
77. Pan, J., and Woodson, S. A. (1998) Folding Intermediates of a Self-splicing RNA: Mispairing of the Catalytic Core, *J. Mol. Biol. 280*, 597−609.
78. Russell, R., and Herschlag, D. (1999) New pathways in folding of the *Tetrahymena* group I RNA enzyme, *J. Mol. Biol. 291*, 1155−1167.
79. Pan, T., and Sosnick, T. R. (1997) Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity, *Nat. Struct. Biol. 4*, 931−938.
80. Pan, J., Thirumalai, D., and Woodson, S. A. (1997) Folding of RNA involves parallel pathways, *J. Mol. Biol. 273*, 7−13.
81. Thirumalai, D., Lee, N., Woodson, S. A., and Klimov, D. K. (2001) Early Events in RNA Folding, *Annu. Rev. Phys. Chem. 52*, 751−762.
82. Pan, J., Deras, M. L., and Woodson, S. A. (2000) Fast Folding of a Ribozyme by Stabilization Core Interactions: Evidence for Multiple Folding Pathways in RNA, *J. Mol. Biol. 296*, 133−144.
83. Hyeon, C., Zhuang, X., and Thirumalai, D. Exploring the folding routes of *Tetrahymena thermophila* ribozyme, unpublished.
84. Baker, D. (2000) A surprising simplicity to protein folding, *Nature 405*, 39−42.
85. Plaxco, K. W., Simons, K. T., and Baker, D. (1998) Contact order, transition state placement and the refolding rates of single domain proteins, *J. Mol. Biol. 277*, 985−994.
86. Gutin, A. M., Abkevich, V. I., and Shakhnovich, E. I. (1996) Chain Length Scaling of Protein Folding Time, *Phys. Rev. Lett. 77*, 5433−5436.
87. Finkelstein, A. V., and Badretdinov, A. Y. (1997) Rate of protein folding near the point of thermodynamic equilibrium between the coil and the most stable chain fold, *Folding Des. 2*, 115−121.
88. Ivankov, D. N., Garbuzynskiy, S. O., Alm, E., Plaxco, K. W., Baker, D., and Finkelstein, A. V. (2003) Contact order revisited: influence of protein size on the folding rate, *Protein Sci. 12*, 2057−2062.
89. Li, M. S., Klimov, D. K., and Thirumalai, D. (2004) Thermal denaturation and folding rates of single domain proteins: size matters, *Polymer 45*, 573−579.
90. Yang, W. Y., and Gruebele, M. (2003) Folding at the speed limit, *Nature 423*, 193−197.
91. Galzitskaya, O. V., Garbuzynskiy, S. O., Ivankov, D. N., and Finkelstein, A. V. (2003) Chain length is the main determinant of the folding rate for proteins with three-state folding kinetics, *Proteins: Struct., Funct., Genet. 51*, 162−166.
92. Ivankov, D. N., and Finkelstein, A. V. (2004) Prediction of protein folding rates from the amino acid sequence-predicted secondary structure, *Proc. Natl. Acad. Sci. U.S.A. 101*, 8942−8944.
93. Perl, D., Holtermann, G., and Schmid, F. X. (2001) Role of the Chain Termini for the Folding Transition State of the Cold Shock Protein, *Biochemistry 40*, 15501−15511.
94. Porschke, D., and Eigen, M. (1971) Cooperative nonenzymic base recognition III. Kinetics of the helix-coil transition of the oligoribouridylic•oligoriboadenylic acid system and of oligoriboadenylic acid alone at acid pH, *J. Mol. Biol. 62*, 361−381.
95. Porschke, D., Uhlenbec, O. C., and Martin, F. H. (1973) Thermodynamics and Kinetics of Helix-Coil Transition of Oligomers Containing GC Base Pairs, *Biopolymers 12*, 1313−1335.
96. Heilman-Miller, S. L., Pan, J., Thirumalai, D., and Woodson, S. A. (2001) Role of Counterion Condensation in Folding of *Tetrahymena* Ribozyme I. Equilibrium stabilization by cations, *J. Mol. Biol. 306*, 1157−1166.
97. Heilman-Miller, S. L., Pan, J., Thirumalai, D., and Woodson, S. A. (2001) Role of Counterion Condensation in Folding of *Tetrahymena* Ribozyme II. Counterion-dependence of Folding Kinetics, *J. Mol. Biol. 309*, 57−68.
98. Misra, V. K., and Draper, D. E. (1998) On the Role of Magnesium Ions in RNA Stability, *Biopolymers 48*, 113−135.
99. Qiu, L., Zachariah, C., and Hagen, S. J. (2003) Fast Chain Contraction during Protein Folding: "Foldability" and Collapse Dynamics, *Phys. Rev. Lett. 90*, 168103.
100. Perez-Salas, U. A., Rangan, P., Krueger, S., Briber, R. M., Thirumalai, D., and Woodson, S. A. (2004) Compaction of a Bacterial Group I Riboyme Coincides with the Assembly of Core Helices, *Biochemistry 43*, 1746−1753.
101. Akiyama, S., Takahashi, S., Kimura, T., Ishimori, K., Morishima, I., Nishikawa, Y., and Fujisawa, T. (2002) Conformational landscape of cytochrome *c* folding studied by microsecond-resolved small-angle X-ray scattering, *Proc. Natl. Acad. Sci. U.S.A. 99*, 1329−1334.
102. Cárdenas, A. E., and Elber, R. (2003) Kinetics of cytochrome C folding: Atomically detailed simulations, *Proteins: Struct., Funct., Genet. 51*, 245−257.
103. Das, R., Kwok, L., Millett, I., Bai, Y., Mills, T., Jacob, J., Maskel, G., Seifert, S., Mochrie, S., Thiyagarajan, P., Doniach, S., Pollack, L., and Herschlag, D. (2003) The Fastest Global Events in RNA Folding: Electrostatic Relaxation and Tertiary Collapse of the *Tetrahymena* Ribozyme, *J. Mol. Biol. 332*, 311−319.
104. Takamoto, K., Das, R., He, Q., Doniach, S., Brenowitz, M., Herschlag, D., and Chance, M. R. (2004) Principles of RNA Compaction: Insights from the Equilibrium Folding Pathway of the P4−P6 RNA Domain in Monovalent Cations, *J. Mol. Biol. 343*, 1195−1206.
105. Lee, N., and Thirumalai, D. (2001) Dynamics of collapse of flexible polyelectrolytes in poor solvents, *Macromolecules 32*, 3446−3457.
106. Lee, N., and Thirumalai, D. (2000) Dynamics of collapse of flexible polyampholyte, *J. Chem. Phys. 113*, 5126−5129.
107. Fenton, W. A., and Horwich, A. L. (1997) GroEL-mediated protein folding, *Protein Sci. 6*, 743−760.
108. Karpel, R. L., Miller, N. S., and Fresco, J. R. (1982) Mechanistic Studies of Ribonucleic Acid Renaturation by a Helix-Destabilizing Protein, *Biochemistry 21*, 2102−2108.
109. Herschlag, D. (1995) RNA Chaperones and the RNA Folding Problem, *J. Biol. Chem. 270*, 20871−20874.
110. Schroeder, R., Grossberger, R., Pichler, A., and Waldsich, C. (2002) RNA folding in vivo, *Curr. Opin. Struct. Biol. 12*, 296−300.
111. Lambowitz, A. M., Caprara, M. G., Zimmerly, S., and Perlman, P. S. (1999) Group I and group II ribozymes as RNPs: clues to the past and guides to the future in *The RNA World*, Cold Spring Harbor, New York.
112. Viitanen, P. V., Gatenby, A. A., and Lorimer, G. H. (1992) Purified chaperonin 60 (groEL) interacts with the nonnative states of a multitude of *Escherichia coli* proteins, *Protein Sci. 1*, 363−369.
113. Buckle, A. M., Zahn, R., and Fersht, A. R. (1997) A structural model for GroEL-polypeptide recognition, *Proc. Natl. Acad. Sci. U.S.A. 94*, 3571−3575.
114. Stan, G., Brooks, B. R., and Thirumalai, D. Probing the "annealing" mechanism of GroEL minichaperone using molecular dynamics simulation, unpublished.
115. Wang, J. D., Michelitsch, M. D., and Weissman, J. S. (1998) GroEL-GroES-mediated protein folding requires an intact central cavity, *Proc. Natl. Acad. Sci. U.S.A. 95*, 12163−12168.

116. Thirumalai, D. (1994) *Statistical Mechanics, Protein Structure, and Protein − Substrate Interactions*. (Doniach, S., Ed.) Theoretical Perspectives on In Vitro and In Vivo Protein Folding, pp 115−134, New York, Plenum.
117. Shtilerman, M., Lorimer, G. H., and Englander, S. W. (1999) Chaperonin function: folding by forced unfolding, *Science 284*, 822−825.
118. Betancourt, M. R., and Thirumalai, D. (1999) Exploring the kinetic requirements for enhancement of protein folding rates in the GroEL cavity, *J. Mol. Biol. 287*, 627−644.
119. Thirumalai, D., and Lorimer, G. H. (2001) Chaperonin-mediated protein folding, *Annu. Rev. Biophys. Biomol. Struct. 30*, 245−269.
120. Todd, M. J., Viitanen, P. V., and Lorimer, G. H. (1994) Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding, *Science 256*, 659−666.
121. Brinker, A., Pfeifer, G., Kerner, M. J., Naylor, D. J., Hartl, F. U., and Hayer-Hartl, M. (2001) Dual Function of Protein Confinement in Chaperonin-Assisted Protein Folding, *Cell 107*, 223−233.
122. Brechm, S. L., and Cech, T. R. (1983) Fate of an Intervening Sequence Ribonucleic Acid: Excision and Cyclization of the *Tetrahymena* Ribosomal Ribonucleic Acid Intervening Sequence in Vivo, *Biochemistry 22*, 2390−2397.
123. Bass, B. L., and Cech, T. R. (1984) Specific interaction between the self-splicing RNA of *Tetrahymena* and its guanosine substrate: implications for biological catalysis by RNA, *Nature 308*, 820−826.
124. Zhang, F., Ramsay, E. S., and Woodson, S. A. (1995) In vivo facilitation of *Tetrahymena* group I intron splicing in *Escherichia coli* pre-ribosomal RNA, *RNA 1*, 284−292.
125. Woodson, S. A. (2000) Recent insights on RNA folding mechanisms from catalytic RNA, *Cell. Mol. Life. Sci. 57*, 796−808.
126. Weeks, K. M., and Cech, T. R. (1995) Efficient Protein-Facilitated Splicing of the Yeast Mitochondrial bI5 Intron, *Biochemistry 34*, 7728−7738.
127. Weeks, K. M., and Cech, T. R. (1995) Protein facilitation of group-I intron splicing by assembly of the catalytic core and the 5′-splice-site domain, *Cell 82*, 345−348.
128. Mohr, S., Stryker, J. M., and Lambowitz, A. M. (2002) A DEAD-Box Protein Functions as an ATP-Dependent RNA Chaperone in Group I Intron Splicing, *Cell 109*, 769−779.
129. Waldsich, C., Grossberger, R., and Schroeder, R. (2002) RNA chaperone StpA loosens interactions of the tertiary structure in the td group I intron in vivo, *Genes Dev. 16*, 2300−2312.
130. Webb, A. E., Rose, M. A., Westhof, E., and Weeks, K. M. (2001) Protein-dependent transition states for ribonucleoprotein assembly, *J. Mol. Biol. 309*, 1087−1100.
131. Chen, X., Gutell, R. R., and Lambowitz, A. M. (2000) Function of tyrosyl-tRNA synthetase in splicing group I introns: an induced-fit model for binding to the P4−P6 domain based on analysis of mutations at the junction of the P4−P6 stacked helices, *J. Mol. Biol. 301*, 265−283.
132. Buchmueller, K. L., Webb, A. E., Richardson, D. A., and Weeks, K. M. (2000) A collapsed non-native RNA folding state, *Nat. Struct. Biol. 7*, 362−366.
133. Buchmueller, K. L., and Weeks, K. M. (2003) Near Native Structure in an RNA Collapsed State, *Biochemistry 42*, 13869−13878.
134. Weeks, K. M., and Cech, T. R. (1996) Assembly of a ribonucleoprotein catalyst by tertiary structure capture, *Science 271*, 345−348.
135. Weeks, K. M. (1997) Protein-facilitated RNA folding, *Curr. Opin. Struct. Biol. 7*, 336−342.
136. Garcia, I., and Weeks, K. M. (2004) Structural Basis for the Self-Chaperoning Function of an RNA Collapsed State, *Biochemistry 43*, 15179−15186.
137. Lorsch, J. R. (2002) RNA Chaperones Exist and DEAD Box Proteins Get a Life, *Cell 109*, 797−800.
138. Schultes, E. A., and Bartel, D. P. (2000) One Sequence, Two Ribozymes: Implications for the Emergence of New Ribozyme Folds, *Science 289*, 448−452.
139. Dima, R. I., Hyeon, C., and Thirumalai, D. (2005) Extracting stacking interaction parameters for RNA from the data set of native structures, *J. Mol. Biol.*, in press.
140. Wilkinson, K. A., Merino, E. J., and Weeks, K. M. (2005) RNA SHAPE chemistry reveals nonhierarchical interactions dominate equilibrium structural transitions in $tRNA^{Asp}$ transcripts, *J. Am. Chem. Soc. 127*, in press.
141. Fresco, J. R., Adams, A., Ascione, R., Henley, D., and Lindahl, T. (1966) Tertiary Structure in Transfer Ribonucleic Acids, *Cold Spring Harbor Symp. Quant. Biol. 31*, 527−537.
142. Kiefhaber, T. (1995) Kinetic traps in lysozyme folding, *Proc. Natl. Acad. Sci. U.S.A. 92*, 9029−9033.
143. Matagne, A., Radford, S. E., and Dobson, C. M. (1997) Fast and slow tracks in lysozyme folding: insight into the role of domains in the folding process, *J. Mol. Biol. 267*, 1068−1074.
144. Millet, I. S., Townsley, L. E., Chiti, F., Doniach, S., and Plaxco, K. W. (2002) Equilibrium Collapse and the Kinetic "Foldability" of Proteins, *Biochemistry 41*, 321−325.
145. Riesner, D., and Römer, R. (1973) *Physico-Chemical Properties of Nucleic Acids*, Vol. 2, Thermodynamics and Kinetics of Conformational Transitions in Oligonucleotides and tRNA, pp 237−318, Academic Press, London.
146. Deras, M. L., Brenowitz, M., Ralston, C. Y., Chance, M. R., and Woodson, S. A. (2000) Folding mechanism of the *Tetrahymena* ribozyme p4-p6 domain, *Biochemistry 39*, 10975−10985.
147. Rangan, P., Masquida, B., Westhof, E., and Woodson, S. A. (2003) Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme, *Proc. Natl. Acad. Sci. U.S.A. 100*, 1574−1579.
148. Fang, X., Pan, T., and Sosnick, T. R. (1999) Mg-dependent folding of a large ribozyme without kinetic traps, *Nat. Struct. Biol. 12*, 1091−1095.
149. Xiao, M., Leibowitz, M. J., and Zhang, Y. (2003) Concerted folding of a Candida ribozyme into the catalytically active structure posterior to a rapid RNA compaction, *Nucleic Acids Res. 31*, 3901−3908.
150. Zarrinkar, P. P., Wang, J., and Williamson, J. R. (1996) Slow folding kinetics of RNase P RNA, *RNA 2*, 564−573.
151. Sclavi, B., Sullivan, M., Chance, M. R., Brenowitz, M., and Woodson, S. A. (1998) RNA Folding at Millisecond Intervals by Synchrotron Hydroxyl Radical Footprinting, *Science 279*, 1940−1943.
152. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal Structure of a Group I Ribozyme Domain: Principles of RNA Packing, *Science 273*, 1678−1685.
153. Lehnert, V., Jaeger, L., Michel, F., and 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.

BI047314+