Theor Chem Acc (1997) 96:14–22

Theoretical
Chemistry Accounts

Overview

Kinetic partitioning mechanism as a unifying theme in the folding of biomolecules

D. Thirumalai, D.K. Klimov, S.A. Woodson

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

Received: 14 January 1997 / Accepted: 23 January 1997

Abstract. We describe a unified approach to describe the kinetics of protein and RNA folding. The underlying conceptual basis for this framework relies on the notion that biomolecules are topologically frustrated due to their polymeric nature and due to the presence of conflicting energies. As a result, the free energy surface (FES) has, in addition to the native basin of attraction (NBA), several competing basins of attraction. A rough FES results in direct and indirect pathways to the NBA, i.e., a kinetic partitioning mechanism (KPM). The KPM leads to a foldability principle according to which fast folding sequences are characterized by the folding transition temperature T_F being close to the collapse transition temperature T_θ, at which a transition from the random coil to the compact structure takes place. Biomolecules with $T_\theta \approx T_F$, such as small proteins and tRNAs, are expected to fold rapidly with two-state kinetics. Estimates for the multiple time scales in KPM are also given. We show that experiments on proteins and RNA can be understood semi-quantitatively in terms of the KPM.

Key words: Kinetic partitioning mechanism – Nucleation collapse – RNA folding

1 Introduction

The pioneering experiments of Anfinsen and co-workers [1] in the early 1960s showed that the folding of proteins into a unique structure with a well-defined three-dimensional topology is a self-assembly process, that is, the information needed to reach the native conformation is encoded in the primary sequence itself. Although these experiments were instrumental in demonstrating the possibility that the native conformation of proteins corresponds to the global free energy minimum, they did not address the question of how the native state is reached in a biologically relevant time scale, t_B. The question of kinetic accessibility of the native conformation came into focus when Levinthal argued [2] that the time for random search of all available conformations far exceeds t_B. This seemed to imply that in order for proteins to reach the native conformation on the time scale, t_B, there have to be preferred pathways that will essentially limit the search of the conformational space. The Levinthal paradox, simplistic as it is, has served as an intellectual impetus to understand how proteins find their native conformations in a relatively short time.

In the past several years, through a combination of sophisticated experiments [3–10] and a study of minimal models of proteins [11–20], a new framework for understanding folding kinetics has emerged. Several reviews summarizing various aspects of the theoretical framework have appeared [21–23]. In this article we describe a complementary, but different, perspective on how biomolecules (both proteins and RNA) reach their native conformations under folding conditions. The basic idea behind all theoretical studies is that the underlying topography of the free energy landscape of biomolecules is rugged [21, 23], consisting of many minima separated by barriers of varying heights. It should be stressed that the notion of a complicated free energy surface (FES) has been invoked in a variety of contexts. The rugged nature of potential energy surfaces was introduced to understand slow dynamics in structural glasses nearly 30 years ago [24]. More recently, extensive investigations have revealed that the hallmark of several classes of disordered systems [25] is that the energy surface is complex, which in turns leads to activated scaling and non-Arrhenius temperature dependence of transport coefficients. Natural proteins are unique in the sense that despite their complicated FES there is a dominant basin of attraction that is accessible on the time scale t_B. Therefore, the challenge is to understand how a polypeptide chain explores the complex FES in order to reach the global free energy minimum. A natural, but tautological, answer is that biomolecular sequences have evolved to fold rapidly. The major contribution of theoretical studies is to show how the rapid assembly of proteins and RNA takes place by examining

Correspondence to: D. Thirumalai

the kinetic processes that are involved in the exploration of the complicated FES. In addition, spurred in part by the theoretical developments new experimental techniques have been elaborated to provide details (almost at the atomic level) of the events in the folding process ranging from the tens of nanoseconds to submillisecond time scales. The joint efforts of theoreticians and the experimental community are leading to rapid progress in our ability to describe, in some instances quantitatively, the kinetics of biomolecular self-assembly.

In a recent article [26] we have suggested that concepts developed in the context of protein folding may be used to understand the kinetics of RNA folding. The purpose of this review is to describe a unified approach that uncovers the global features that are expected in the folding kinetics of proteins and RNA. We also provide comparisons between theoretical predictions and experimental measurements. Our goal is to point out that one can describe in some detail the general kinetic principles of folding of biomolecules from the statistical mechanics perspective. Furthermore, we will show that the expressions for the various time constants can be estimated (at least, to within an order of magnitude) in terms of experimental parameters leading to validation of the theoretical concepts.

2 Topological frustration and the kinetic partitioning mechanism

Most of the qualitative features of the folding kinetics of biomolecules can be understood by introducing the notion of topological frustration. A crucial feature of proteins is that the primary sequence contains a certain fraction of monomers that are hydrophobic. The fraction of hydrophobic residues in globular proteins is slightly in excess of 0.5 [27]. The linear density of the hydrophobic residues is roughly uniform along the contour of the polypeptide chain, which means that the hydrophobic residues are dispersed throughout the chain. If this were not the case, proteins would tend to aggregate. A consequence of the uniform density of hydrophobic residues is that on any length scale l that is not equal to the size of the chain there is a propensity for the hydrophobic residues to form tertiary contacts under folding conditions. The resulting structures, formed by having contacts between proximal hydrophobic residues, would in all likelihood be incompatible with the unique global fold. The incompatibility of the structures on local length scales with the native conformation leads to topological frustration.

It is important to appreciate that topological frustration is an inherent consequence of the polymeric nature of proteins (connectivity of residues) as well as the presence of competing interactions (hydrophobic species, which prefer to form compact structures, and hydrophilic residues, which are better accommodated by extended conformations). Thus, all proteins are topologically frustrated, with long chains being more so than smaller ones. A direct consequence of topological frustration in proteins is that the underlying topography of the FES is complex, consisting of many minima that are separated by a distribution of barriers. We have recently shown [26] that similar considerations apply to RNA as well.

The nature of the low-lying minima is easy to describe qualitatively. On any length scale l there are numerous ways of constructing structures that are in conflict with global fold. Many, perhaps most, of these structures would have high free energies and be consequently unstable to thermal fluctuations. However, certain of these structures are truly stable low-energy minima that represent the conformations that can have many structural features in common with the native state. The energetic difference between these low-energy misfolded structures and the native state can be easily compensated by the entropy associated with the misfolded structures. Thus, such structures potentially act as kinetic traps in a typical folding experiment and slow down the rate of protein folding.

The fundamentals of the kinetic partitioning mechanism (KPM) can be deciphered from the concept that there are low-energy minima (in which the proteins are misfolded) separated by free energy barriers from the native state. Imagine an ensemble of denatured molecules under folding conditions (achieved by diluting the concentration of the denaturant, for example) in search of the deepest basin of attraction in the rugged free energy landscape. A fraction of the molecules Φ would map onto the native basin of attraction (NBA) directly, and would reach it rapidly without being trapped in other states. The remaining fraction would inevitably be stuck in one or several of the low-energy misfolded structures, and only on a longer time scale reach the native state by suitable activation processes. Thus, because of the topological frustration that gives rise to a rugged free energy landscape, the pool of denatured molecules partitions into fast folders and slow folders that reach the native state by indirect slow off-pathway processes.

A schematic sketch of the KPM is shown in Fig. 1. The outline of the kinetic scheme that emerges from Fig. 1 can be conveniently written as

$$\begin{array}{ccccc} \mathbf{U} & & \xrightarrow[k_1]{\Phi} & & \mathbf{F} \\ & \mathbf{k_2} \searrow & & \nearrow \mathbf{k_3} & \\ & & \{\mathbf{MS}\} & & \end{array}$$

where **U** refers to the unfolded states, **F** is the folded state, and $\{\mathbf{MS}\}$ denotes the collection of the low-energy misfolded states. For simplicity we have not indicated the rates for the backward processes in the above kinetic scheme.

The yield of the fast process is given by Φ, the partition factor. Since the shape and structure of the underlying FES determine Φ, it becomes apparent that Φ depends not only on the factors intrinsic to sequences but also on external conditions such as pH, temperature and ionic strength. Furthermore, Φ can be easily altered by mutations so that a wild-type protein with $\Phi \approx 1$ can be made into a slow or moderate folder. In the subsequent sections we will discuss in some detail the dominant time scales that arise in the KPM. We will also show that various experiments can be at least qualitatively understood in terms of the KPM.

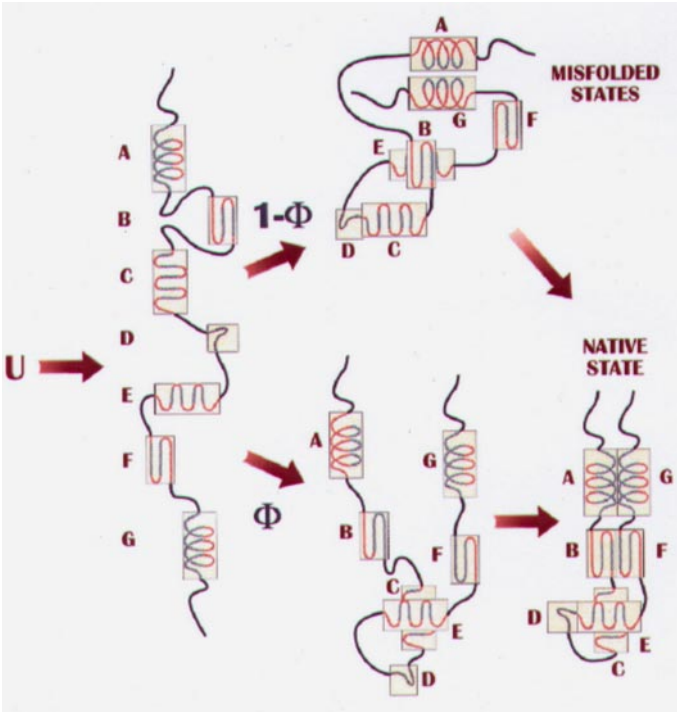

Fig. 1. A pictorial representation of the kinetic partition mechanism. The unfolded structures collapse rapidly (in, perhaps, microsecond time scales). These structures contain almost all of the secondary structures. These structural elements are shown as blocks labeled as *A*–*G*. Some of these blocks show helices, which are expected to form in submicrosecond time scales, while other blocks show beta-strands, which form in about 10 μs. The subsequent packing of these secondary structural elements results in a fraction of the population Φ going directly to the native state via the native conformation nucleation-collapse mechanism. An example of a transition state obtained along this pathway is displayed as an expanded version of the native conformation. This structure contains all native-like contacts, and the lack of native contacts between blocks *A* and *G* and *B* and *F* makes this structure somewhat larger than the native state. The remaining fraction of the molecules, 1–Φ, gets trapped in misfolded structures, an example of which is shown in the upper-right corner. In this case helices *A* and *G* have incorrect orientation and non-native contact between *B* and *E* has been formed. The activated transitions from the misfolded structures to the native state involve partial unraveling of the polypeptide chain to break the incorrect contacts and establish native contacts. In this highly simplified representation hydrophobic portions of sequence are shown in *blue* and the hydrophilic are given in *red*

3 Experimental evidence for KPM

Several recent experimental studies support the basic ideas of the KPM [7, 26]. These experiments suggest that foldable proteins can be divided into two classes. First are the fast folders which reach the native conformation in a two-state kinetic process without being trapped in any intermediates. Typically, fast-folding proteins are relatively small. From our theoretical perspective one would conclude that the σ values (see Eq. 1) for these are small and the underlying energy landscape is dominated by a single basin of attraction corresponding to the native conformation. The time constant for reaching the native conformation is of the order of a few milliseconds, which is consistent with the theoretical prediction given in Eq. (2). These proteins have the partition factor, Φ, that is quite close to unity.

The other class of proteins is the moderate folders which exhibit multiple kinetic phases predicted by the KPM. The most detailed verification of KPM has come from the study of the refolding of lysozyme. In important early experiments, Radford et al. [7, 28] observed that the protection kinetics in hydrogen exchange labeling experiments is well described by biphasic kinetics. If we follow the interpretation suggested by Thirumalai and Guo [29] that the fast phase in these experiments describes the mechanism of refolding to the native state by the nucleation-collapse process, then the corresponding amplitude gives an estimate of the partition factor. With this interpretation the experiments of Radford et al. [7, 28] suggest that $\Phi \approx 0.25$ at $T = 20^{\circ}$C, pH 5.2 and at the concentration of guanidinium chloride (GdmCl) of 0.54 M.

More recently, Kiefhaber [30] has performed an ingenious experiment that very clearly shows the full range of kinetic behavior predicted by KPM. He used interrupted refolding experiments on hen egg lysozyme to directly measure the value of Φ. The refolding experiments are initiated by diluting completely unfolded lysozyme into final folding conditions (0.6 M GdmCl, pH 5.2, $T = 20^{\circ}$C). At various times the folding is interrupted by transferring the solution to 5.3 M GdmCl and pH 1.8. Apparently, under these conditions the native lysozyme unfolds completely in about 20 s, while any partially formed misfolded structures unravel in a few milliseconds. Thus, the amplitude of the slow unfolding process gives the amount of the native structure when the folding process is interrupted. By varying the time of interruption a history of the kinetics of formation of the native state can be constructed. Kiefhaber [30] shows, by analyzing the folding kinetics at 0.6 M GdmCl, that the partition factor Φ is 0.14, which implies that 14% of the initial population of denatured lysozyme reach the native state in 50 ms. The majority of the population gets to the native conformation by indirect processes that involve transitions out of the kinetic traps. As shown below the time constant for this process using Eq. (5) is about 100 ms, which is in rough agreement with experimental estimate of 420 ms [30].

4 Dominant time scales in KPM

The basic ideas leading to the KPM which are described above have been substantiated using computer simulations of simplified models of proteins [26, 31–33]. The dependence of the various parameters characterizing KPM, namely, Φ, k_1, k_2, k_3, etc., on the properties intrinsic to the sequence have been identified [33]. In particular, it has been shown in a series of papers that for given external conditions the kinetic parameters of KPM are largely determined by equilibrium temperatures that are intrinsic to the protein sequence [32, 33]. It is now known that for foldable sequences there are at least two equilibrium temperatures that determine the "phases" of proteins [34, 35]. One is the high temperature, T_θ, at which the chain undergoes a transition from

a random coil conformation to a collapsed state; T_θ is very similar to the collapse transition temperature introduced by Flory to describe the so-called θ-point in homopolymers. Since there are several distinct energy scales in proteins that discriminate between the exponentially large number of compact conformations, the chain undergoes a folding transition to the native state at a temperature $T_F \lesssim T_\theta$. The transition at T_F is usually first order [35], while the one at T_θ can be either first or second order depending upon a number of factors [36], such as the relative strength of effective two- and three-body interactions in the polypeptide chain. Both T_θ and T_F are experimentally measurable. The folding transition temperature, T_F, is usually associated with the midpoint of the temperature dependence of denaturation plots. The collapse temperature is somewhat harder to measure experimentally. Recently T_θ has been experimentally determined for a small number of proteins [37, 38].

Using lattice models, Monte Carlo simulations, off-lattice models and Langevin simulations we have established that folding times, τ_F, correlate remarkably well with a single parameter that is intrinsic to the sequence [32, 33, 35, 39], namely

$$\sigma = (T_\theta - T_F)/T_\theta. \tag{1}$$

Therefore, τ_F can be varied by altering T_θ and T_F, both of which depend not only on the sequence but also on the external conditions. It is clear that T_F depends precisely on the sequence and hence can be altered by mutations. Surprisingly, T_θ also depends on the sequence [33] (less sensitively than T_F). The reason that T_θ depends on the sequence and not just on the sequence composition is that in addition to the hydrophobic interactions, due to the finite size of proteins the interfacial interactions between the surface residues and the solvent make a large contribution in determining the precise topology. The combination of hydrophobic and interfacial interactions determines T_θ, resulting in the collapse transition temperature being sequence dependent.

The correlation between folding times and σ suggests that the rates of folding can be altered by changing σ while leaving external conditions fixed. Thus the wild-type protein and a mutated one can have very different folding rates depending upon σ. This has also been observed in RNA, in which the folding behavior can be drastically altered by a single point mutation [40, 41]. Using the concepts of polymer theory one of us [42] has shown that the time scales characteristic of KPM can be established in terms of σ and other experimentally controlled parameters. Remarkably, the dominant time scales in the folding process are once again controlled by σ.

5 Fast processes and native conformation nucleation-collapse mechanism

The fast processes, by which a certain fraction Φ of the initial ensemble of denatured molecules reaches the native state, have been shown to occur via a native conformation nucleation-collapse (NCNC) mechanism [29, 43, 44]. According to the NCNC folding is initiated by the formation of native tertiary contacts. Once a critical number of residues form tertiary native contacts, establishing an overall near-native topology in the transition state, the polypeptide chain rapidly reaches the native state. In this mechanism the processes of collapse and the acquisition of the native state are almost synchronous [42], and hence would be nearly indistinguishable. The time scale for the NCNC has been argued to be [42]

$$\tau_{\mathrm{NCNC}} \approx \frac{\eta a}{\gamma} f(\sigma) N^{\omega}, \tag{2}$$

where η is the solvent viscosity, a is roughly the persistent length of the protein, γ is the surface tension, which tends to minimize the exposed surface area of the hydrophobic species, and N is the number of amino acid residues in the protein. The exponent ω lies in the range $3.8 \leq \omega \leq 4.2$. The function $f(\sigma)$ was originally shown to be algebraic in σ. Numerical studies indicate that $f(\sigma) \simeq \exp(\sigma/\sigma_0)$ [32, 39] provides a better fit to the folding times.

There are a number of remarks concerning the NCNC and τ_{NCNC} that are worth making:

1. If σ is small, which implies that collapse and folding are almost synchronous, then $\Phi \approx 1$ and the folding time coincides with τ_{NCNC}. Typically this is only expected for small proteins under optimal folding conditions. Under these circumstances folding kinetics is expected to display two-state behavior. Several experiments suggest that small proteins exhibit the predicted two-state behavior [47, 48].
2. From a theoretical perspective sequences with small σ are extremely well optimized so that the simultaneous requirements of thermodynamic stability and the kinetic accessibility of the native state can be achieved over a relatively large temperature range. Numerical estimates [45] suggest that for these sequences $\tau_F \approx \tau_{\mathrm{NCNC}}$ indeed scales algebraically with N where $\omega \approx 4$, confirming the theoretical predictions [42].
3. In a recent experiment Schönbrunner et al. [46] have suggested that for the 74-residue all-β-sheet-forming protein tendamistat, collapse and folding are essentially indistinguishable. Using the experimental parameters for η and an estimate for γ and a the calculated folding time according to Eq. (2) is about 7 ms, which is remarkably close to the measured value of 10 ms [46]. This estimate also suggests that experiments in the submillisecond regime are required to directly observe the nucleation-dominated processes.

6 Three-stage multipathway mechanism

In the case of moderate and slow folders it is likely that a large fraction of initially denatured molecules do not reach the native state by the NCNC mechanism described above. Such is the case in the refolding of hen egg lysozyme, for example [30]. The partition factor Φ under these circumstances is small. We now describe the approach to the native state of the pool of molecules that follow the indirect off-pathway processes. Extensive numerical studies of lattice and off-lattice models show that the formation of the native structure by indirect

pathways can be conveniently described by three-stage multipathway kinetics [28, 35]. A brief description of each of the stages along with estimates of the associated time scale is given below.

6.1 Non-specific collapse

After the folding process is initiated the polypeptide chain collapses into a relatively compact phase due to the hydrophobic driving force. The kinetics in this stage is quite complex and can perhaps be described by a distribution of time scales leading to a stretched exponential behavior [35]. It should be stressed that in contrast to homopolymer collapse the initiation of collapse in proteins is not completely random. The possible structures that are seen in this phase could depend on loop formation probability, dihedral angle transitions, etc. It is likely that certain secondary structure elements such as helices, which form rapidly, are already present at this stage. By a small generalization of the arguments presented by de Gennes [49] the time scale for non-specific collapse can be written as

$$\tau_c \simeq \frac{\eta a}{\gamma}\left(\frac{T_\theta - T_F}{T_\theta}\right)^3 N^2. \tag{3}$$

An estimate for τ_c can be made by taking $\eta \approx 0.01$ P, $a \approx (5\text{–}10)$ Å and $\gamma \approx (40\text{–}60)\,\mathrm{cal/(Å^2\,mol)}$. With $T_F \simeq T_\theta/2$, τ_c is found to be between (0.02–3)μs for $N = 100$. This time scale roughly coincides with the time for forming a small number of contacts (see Eq. 9) between residues that are far apart in the sequence space.

6.2 Kinetic ordering

In this stage the folding chain discriminates between the exponentially large number of compact conformations to form as many native contacts as possible which would result in a lower free energy. In this phase free energy biases inherent in foldable sequences become operative and the chain navigates to the competing basins of attraction (CBAs) by a very cooperative motion. At the end of this stage the chain reaches one of the low-energy misfolded structures which have many elements in common with the native conformation. The search among these large number of compact structures leading to the low-energy misfolded conformations has been argued to proceed by a reptation-like process with the time constant [42]

$$\tau_{ko} \approx \tau_D N^\zeta, \tag{4}$$

with $\zeta \approx 3$. The time constant τ_D corresponds to a local dihedral angle transition and is approximately 10^{-8} s. Thus, $\tau_{ko} \approx 10$ ms for $N = 100$.

6.3 All-or-none

The last stage in the off-pathway processes involves activated transition from one of the many minimum energy structures to the native state. This process necessarily involves unraveling of the chain (at least partially) in order to break the incorrect contacts and subsequently form the native contacts. The partial unraveling of the chain in the process of transition to the native state has been observed in numerical simulations of minimal models of proteins [15, 35]. It has been argued that the average free energy barrier separating the misfolded structures and the native state scales as $\sqrt{N}k_BT_F$ [42] under certain optimal folding conditions, so that the folding time for the slow process is

$$\tau_F \approx \tau_0 e^{\sqrt{N}} \tag{5}$$

at $T \simeq T_F$. Numerical simulations and more recently experiments [50] suggest that $\tau_0 \simeq 10^{-6}$s so that τ_F for $N = 100$ is 0.1s.

Since the barrier height scales only sublinearly with N it is clear that foldable sequences do not encounter the Levinthal paradox under folding conditions, even if they fold by indirect pathways. There are multiple pathways leading to the second stage, whereas only relatively few pathways connect the misfolded structures and the native state. This is because, as suggested elsewhere, the number of low-energy compact structures only scales as $\ln N$ [51]. It is also clear that if the molecules follow the three-stage kinetic approach to the native state then the transition states occur closer to the native conformation.

7 Foldability principle

There are now numerous examples of proteins that reach the native state in few tens of milliseconds under optimal folding conditions [46–48]. The theoretical reasoning given above indicates that under these conditions the value of σ for these proteins is relatively small. These observations suggest a foldability principle which can be stated as follows: *A sequence is rapidly foldable if* $T_\theta \approx T_F$. By foldability we mean that both the kinetic accessibility and thermodynamic stability are simultaneously satisfied. The foldability principle naturally applies to small single-domain proteins whose sequences can be optimized relatively easily.

It is, in fact, tempting to suggest that the foldability principle, which expresses the kinetic accessibility criterion in terms of properties intrinsic to the sequence, is a quantitative realization of the consistency principle [52] and the principle of minimal frustration [53]. Go [52] realized that spontaneous folding of proteins requires that short-range interactions, which are responsible for secondary structure formation, be compatible with long-range interactions, which confer global topology. Here "short" and "long" refer to distances along the sequence. Go also realized that it is not possible to produce an "ideal protein" in nature, in which there is a complete harmony between long- and short-range interactions. More recently, Bryngelson and Wolynes [53] suggested, using the random energy model [54] as a paradigm for protein folding, that the conflicts between various energy scales should be minimized.

If we take these principles into account we can argue that the minimization of σ should be a natural criterion

for achieving a nearly ideal protein sequence. A heuristic argument leading to this conclusion goes as follows: The collapse transition temperature T_θ is primarily determined by a combination of the driving force that tends to bury the hydrophobic residues in the core of the protein and the forces that place the hydrophilic residues at the surface. The free energy scale D determining T_θ is obtained by a balance between the hydrophobic interactions and the interfacial energies that tend to place hydrophilic residues at the surface to create a nearly compact structure. Thus, $k_B T_\theta \approx D$. We have recently argued [39] that T_F can be approximated as

$$T_F \approx \frac{\Delta_s}{S_{NN}}, \tag{6}$$

where Δ_s is roughly the stability gap [21] and S_{NN} is the entropy associated with the low-energy non-native states. It is reasonable to assume that S_{NN} also depends on D. If the driving force is very large, then $D/k_B T >> 1$, and the polypeptide chain will collapse into one of an exponentially large number of conformations. Since Δ_s is only weakly dependent on N [42], it follows that for large D, $T_F \approx 0$, which is the homopolymer limit. In the opposite limit $D/k_B T << 1$ there is not enough driving force for collapse and S_{NN} once again grows with N (excluding logarithmic corrections), leading to small T_F. A sufficiently high value of T_F is obtained only when S_{NN} is small or the number of low-energy non-native structures is not too big. Thus an optimum value of D is required so that S_{NN} remains small. The existence of optimal D is also consistent with the observation that in natural proteins the fraction of hydrophobic residues is in slight excess of 0.5 [27, 51]. If we use the bound that $T_F \lesssim T_\theta$ then we see that the optimal value for D results for $T_F \approx T_\theta$. An optimal value of D implies a proper balance between long-range and short-range interactions so that the hydrophobic interactions are in harmony with interfacial forces. Thus, at least heuristically we can conclude that the consistency principle [52] and the principle of minimal frustration [53] suggest that σ should be small for optimally designed proteins. Since minimizing σ seems probable only for small proteins it is tempting to suggest that in nature bigger proteins are not optimized.

8 Early events in protein folding

The time scale estimate for τ_{NCNC} for proteins that reach the native state by the nucleation-collapse folding process with $\Phi \approx 1$ is roughly between 0.1ms to few tens of milliseconds, depending on the length of the protein and external conditions (see Eq. 2). In addition, the time scale (Eq. 4) for reaching the low-energy misfolded structures by off-pathway processes is also about a few milliseconds. It is, therefore, of interest to ask about processes that take place on a submillisecond time scale. Following the pioneering work of Eaton and co-workers [10, 55], there has been an explosion of experimental papers [56–59] probing protein folding events on short time scales using a variety of techniques. In the original experiments Jones et al. [10] used optical triggering to refold cytochrome c in a chemical denaturant. More recently, temperature jump [57], electron transfer [56], and other novel mixing techniques [59] have been used to induce and observe protein folding on a submillisecond time scale. The major conclusion of almost all these experiments is that significant self-assembly of proteins begins on time scales as small as a microsecond.

From a general perspective a question of some importance is whether there is an upper limit for the rate of protein folding. In fact, in a recent interesting article explaining the so called "new view" of protein folding Dill and Chan [23] put down a "wish list" for experimental studies of folding kinetics. One of the questions in the wish list was "What is the fastest speed a protein can fold?". Hagen et al. [50] have recently attempted to provide an imaginative answer to this question by using the following reasoning. They conducted an experiment to probe the time it takes for two residues that are far apart in sequence space to form a contact. Such a transient contact may either be native (i.e., the contact is present in the final folded conformation) or non-native. Using optical triggering to refold cytochrome c Hagen et al. [50] estimated that the diffusion controlled rate for two sites separated by $\sim 50-60$ residues to make a contact would be about $(100\,\mu\mathrm{s})^{-1}$. Using the loop formation probability for stiff chains [60] they argued that the more probable contacts between sites separated by 10 or 20 residues can occur in about $1\,\mu$s. Since the formation of a single tertiary contact is the most elementary folding process (besides, say, the formation of secondary structure like a helix) in the route towards the global fold Hagen et al. [50] argued that the upper limit for folding rate of a protein should be $1(\mu\mathrm{s})^{-1}$.

These experimental estimates that 1μs should be an important time scale in the initiation of certain events in the folding process is consistent with theoretical estimates of the time scales $\tau(l)$ for diffusion-limited contact formation between two sites that are separated by l residues. Guo and Thirumalai [61] showed using scaling arguments that $\tau(l)$ can be estimated as

$$\tau(l) \approx \frac{\langle R_l^2 \rangle}{P(l) D_0}, \tag{7}$$

where $\langle R_l^2 \rangle$ is the spatial distance separating the two sites, $P(l)$ is the probability of loop formation [60], and D_0 is the effective monomer diffusion constant. The distribution function $P(l)$ can be calculated by assuming that the backbone is stiff on the scale of the order of persistent length and is given by [60]

$$P(l) \approx \Omega(N) \begin{cases} 0, & l < l_{\min} \\ \frac{1-\exp(-\frac{l_a}{l_p})}{l^{\theta_3}}, & l \geq l_{\min} \end{cases}, \tag{8}$$

where $l_{\min}$ is length of the shortest loop possible, θ_3 is an universal exponent whose value in three dimensions is 2.2, l_a is the length of a residue, and l_p is an effective persistent length which measures the stiffness of the polypeptide chain backbone. The normalization factor $\Omega(N)$ depends on the total number of residues. If we use the Flory results for $\langle R_l^2 \rangle \approx l_p^2 l^{2\nu}$, then the time constant τ_l for $l_a > l_p$ becomes

$$\tau_l \approx \Omega(N) \frac{l_p^2 \, l^{2\nu+\theta_3}}{D_0}. \quad (9)$$

If we take the experimental estimates for $D_0 \simeq 10^{-6}\text{cm}^2/\text{s}$, $l_p \approx 5$ Å, $\nu \simeq 0.6$ then τ_l for $l = 10$ according to Eq. (9) turns out to be about 10 μs, where $\Omega(N)$ is computed using $l_{min} \simeq 7$ and $N \to \infty$. This theoretical estimate is consistent with the experimental measurements given the inherent uncertainties in D_0 and l_p.

In retrospect it is not surprising that 1 μs or so turns out to be an important time scale in the early processes of protein folding. This was already realized based on theoretical arguments given in the context of the refolding of bovine pancreatic trypsin inhibitor [60, 62]. However, an understanding of how the formation of these contacts leads to further self-assembly of proteins is still lacking. This requires experiments that can directly observe correlated events. It is only through such experiments that a molecular basis for the nucleation-collapse mechanism can be provided.

Experimental [50] and theoretical [61] studies clearly suggest that transient long-range contacts occur on some microsecond time scale. If these contacts are non-native and stable then the polypeptide chain will subsequently collapse into a misfolded structure on a short (much less than about milliseconds) time scale. In proteins with small σ the initial stable contacts are expected to be native, and once a sufficient number of such contacts forms then a rapid transition to the native state takes place, presumably via the nucleation-collapse mechanism. As suggested above, experiments that can observe correlation between multiple contacts will be required to further elucidate the nature of the nucleation-collapse mechanism.

9 Folding of RNA

The folding pathways of large RNA are now beginning to be probed experimentally. In the past year the existence of possible connections between the folding of RNA and proteins has been pointed out [26, 63]. From the perspective of the rough FES it is natural to expect some common qualitative elements for folding kinetics of proteins and RNA [26]. The usual arguments like the incompatibility of times for random search of all conformations and biological folding times apply equally well to nucleotide sequences. For RNA it is necessary to form correct secondary structures, namely Watson-Crick pairs between complementary sequences. The correctly formed secondary structures assemble to achieve the correct three-dimensional organization of the structural elements. Although we expect certain common trends in the folding of RNA and proteins there are also fundamental differences [26, 63] that have not been explored. One major difference is that collapse of RNA typically requires binding of divalent ions [64, 65].

As for proteins, it is found that certain RNA sequences fold rapidly without being trapped in misfolded states. These include tRNAs [66] as well as a certain small group self-splicing introns [48] for which Φ under in vitro conditions appears to be close to unity. The folding time for tRNA was estimated to be about 0.1–1 s [66] suggesting that perhaps folding occurs via a nucleation-collapse mechanism.

In a recent article [26] we have begun to analyze in quantitative terms the folding kinetics of *Tetrahymena* ribozyme in terms of the KPM. The availability of considerable structural information makes the *Tetrahymena* ribozyme an attractive model for studying the folding of large RNAs [67]. The arguments given for topological frustration suggest that the low-energy misfolded structures become more prominent for larger length chains, resulting in a smaller value of Φ [26].

These expectations are borne out in the quantitative analysis of the experiments on the refolding of precursor RNA containing the *Tetrahymena* ribozyme. The experiments of Emerick and Woodson [68, 69] showed, using self-splicing kinetics and gel electrophoresis, that a population containing a mixture of active and inactive conformers is in slow exchange at $T = 30°\text{C}$. The majority of the population ($> 70\%$) of the wild-type *Tetrahymena* precursor RNA appears to be misfolded after transcription at $T = 30°\text{C}$. If the RNA is heated to $T = 75°\text{C}$ and annealed to $T = 30°\text{C}$ the percentage of inactive molecules decreases to about 20–30%. Thus the inactive conformations (presumably misfolded) can be made to reach the native state by an annealing process.

The experimental findings of Emerick and Woodson [68, 69] confirm the basic picture of folding predicted by KPM. According to KPM for larger RNAs one expects the chain to be trapped in one of the low-energy structures. The relatively slow folding is a consequence of escape from these traps by an activated process. The partition factor at $T = 30°\text{C}$ is small ($\Phi \approx 0.2$), which implies that most of the molecules reach the native conformation by off-pathway processes.

Another theoretical prediction made in the context of proteins that appears to be consistent with experiments on RNA is the activation energy separating the misfolded states and the active folded conformation. Based on the temperature dependence of the conversion of the inactive to active form the barrier height was estimated to be (10–15) kcal/mol [68, 69]. According to the theoretical arguments such barriers are expected to scale as $\sqrt{N} k_B T$ [42] which for *Tetrahymena* precursor RNA ($N = 650$) turns out to be 15 kcal/mol. The good agreement between theoretical estimate and experiments suggests that typical free energy barriers in biomolecules are small.

10 Conclusions

It is gratifying that certain general principles of folding kinetics of proteins and RNA can be deciphered from simple considerations [26]. Due to their polymeric nature and the presence of conflicting energy scales, biomolecules are intrinsically topologically frustrated. As a consequence the FES is complex and contains not only the NBA, but also CBAs as well. The basic features of the KPM naturally emerge from this idea. The concepts

outlined in this article should be viewed as a tentative unified proposal to conveniently classify the possible scenarios that can arise in the complex self-assembly of proteins and RNA. It is possible to extend these concepts to make testable predictions for the folding of specific proteins [60] and RNA, assuming more detailed models that account for solvent conditions and other aspects that are left out in the simplified description. It is nonetheless clear that our understanding of the folding kinetics will continue to grow rapidly through an interplay between theoretical ideas and experimental advances.

A unified description of the folding process of proteins and RNA is expected to advance the study of RNA folding. For example, it is logical to suggest that fast processes in RNA could well determine the extent to which misfolded structures are going to slow down the folding process. If the similarities between the nature of folding of proteins and RNA are further pursued, then it would imply that the organization of the folded structure of RNA also involves parallel pathways [70]. The recent semi-quantitative analysis [26] of the experiments of Emerick and Woodson [68] strongly suggests that folding of large RNAs does take place by multiple parallel pathways. Additional experiments on faster time scales, are needed to further elucidate the nature of these pathways.

Finally, the KPM also points to the need for chaperonin-assisted folding of proteins and RNAs [26]. The arguments based on KPM would suggest that only when the partition factor is small (less than 10%) does one require the chaperonin machinery [26, 71]. Typically, as suggested here and elsewhere [26], this happens only for large proteins. For these, N is sufficiently large so that the folding time given by Eq. (5) not only exceeds t_B but starts to become comparable with the time scale for aggregation processes. Under these circumstances the chaperones are predicted to rescue the misfolded structures by a process referred to as iterative annealing mechanism [71, 72]. A similar reasoning would suggest that for large RNA as well, there must exist RNA chaperones [26] which presumably function in a manner similar to GroEL and GroES. The RNA chaperones have not yet identified, although certain non-binding RNA proteins have been shown to enhance the rate of RNA-catalyzed reactions in vitro [73, 74].

Acknowledgements. We would like to thank Oksana Klimova for preparing the figure. This work was supported in part by grants from the National Science Foundation, the Air Force Office of Scientific Research, the National Institutes of Health, the American Cancer Society, the Pew Charitable Trust, and the Camille and Henry Dreyfus Foundation.

References

1. Anfinsen CA (1973) Science 181: 223
2. Levinthal C (1969) In: Debrunner P, Tsibris JCM, Münck E (eds) Mössbauer spectroscopy in biological systems. University of Illinois Press, Urbana
3. Englander SW, Poulsen A (1969) Biopolymers 7: 329
4. Bai Y, Englander SW (1996) Proteins Struct Funct Genet 24: 145
5. Udgaonkar JB, Baldwin RL (1988) Nature 335: 694
6. Roder H, Elöve GA, Englander SW (1988) Nature 335: 700
7. Radford SE, Dobson CM (1995) Phil Trans R Soc Lond B 348: 17
8. Matouschek A, Kellis JT Jr, Serrano L, Bycroft M, Fersht AR (1990) Nature 346: 440
9. Fersht AR (1995) Curr Opin Struct Biol 5:79
10. Jones CM, Henry ER, Hu Y, Chan CK, Luck SD, Bhuyan A, Roder H, Hofrichter J, Eaton WA (1993) Proc Natl Acad Sci USA 90: 11860
11. Bryngelson JD, Wolynes PG (1989) J Phys Chem 93: 6902
12. Honeycutt JD, Thirumalai D (1990) Proc Natl Acad Sci USA 87: 3526
13. Shakhnovich E, Farztdinov G, Gutin AM, Karplus M (1991) Phys Rev Lett 67: 1665
14. Leopold PE, Montal M, Onuchic JN (1992) Proc Natl Acad Sci USA 89: 8721
15. Chan HS, Dill KA (1994) J Chem Phys 100: 9238
16. Zwanzig R, Szabo A, Bagchi B (1992) Proc Natl Acad Sci USA 89: 20
17. Miller R, Danko CA, Fasolka MJ, Balazs AC, Chan HS, Dill KA (1992) J Chem Phys 96: 768
18. Skolnick J, Kolinski A (1990) Science 250: 1121
19. Skolnick J, Kolinski A (1991) J Mol Biol 221: 499
20. Zwanzig R (1995) Proc Natl Acad Sci USA 92: 9801
21. Bryngelson, JD, Onuchic JN, Socci ND, Wolynes PG (1995) Proteins Struct Funct Genet 21: 167
22. Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS (1995) Protein Sci 4: 561
23. Dill KA, Chan HS (1997) Nature Struct Biol 4: 10
24. Goldstein M (1969) J Chem Phys 51: 3328
25. Mezard M, Parisi G, Virasaro MA (1987) Spin glass theory and beyond. World Scientific, Singapore
26. Thirumalai D, Woodson SA (1996) Acc Chem Res 29: 433
27. White SH, Jacobs RE (1993) J Mol Evol 36: 79
28. Radford SE, Dobson CM, Evan PA (1992) Nature 358: 302
29. Thirumalai D, Guo Z (1995) Biopolymers 35: 137
30. Kiefhaber T (1995) Proc Natl Acad Sci USA 92: 9029
31. Mirny LA, Abkevich V, Shakhnovich EI (1996) Folding Design 1: 103
32. Klimov DK, Thirumalai D (1996) Proteins Struct Funct Genet 26: 411
33. Veitshans T, Klimov DK, Thirumalai D (1997) Folding Design 2: 1
34. Bryngelson JD, Wolynes PG (1990) Biopolymers 30: 177
35. Camacho CJ, Thirumalai D (1993) Proc Natl Acad Sci USA 90: 6369
36. Camacho CJ, Schanke T (1997) Europhys Lett (in press)
37. Alexander P, Fahnestock S, Lee T, Orban J, Bryan P (1992) Biochemistry 31: 3597
38. Catanzano F, Giancola C, Barone G (1996) Biochemistry 35: 13378
39. Camacho CJ, Thirumalai D (1996) Europhys Lett 35: 627
40. Jaeger L, Westhoff E, Michel F (1993) J Mol Biol 234: 331
41. Emerick VL, Woodson SA (1993) Biochemistry 32: 14062
42. Thirumalai D (1995) J Phys (France) 5: 1457
43. Abkevich VI, Gutin AM, Shakhnovich E (1994) Biochemistry 33: 10026
44. Fersht AR (1995) Proc Natl Acad Sci USA 92: 10869; Sosnick TR, Englander SW (1996) Proteins Struct Funct Genet 24: 413
45. Abkevich VI, Gutin AM, Shakhnovich EI (1996) Phys Rev Lett 77: 5433
46. Schönbrunner N, Koller KP, Kiefhaber T (1997) J Mol Biol (in press)
47. Schindler T, Herrler M, Marahiel MA, Schmid FX (1995) Nat Struct Biol 2: 663
48. Kragelund BB, Robinson CV, Kundsen J, Dobson CM, Poulsen FM (1995) Biochemistry 34: 7217
49. de Gennes PG (1985) J Phys Lett 46: L639
50. Hagen SJ, Hofrichter J, Szabo A, Eaton WA (1996) Proc Natl Acad Sci USA 93: 11615

51. Camacho CJ, Thirumalai D (1993) Phys Rev Lett 71: 2505
52. Go N (1983) Annu Rev Biophys Bioeng 12: 183
53. Bryngelson JD, Wolynes PG (1987) Proc Natl Acad Sci USA 84: 7524
54. Derrida B (1980) Phys Rev Lett 45: 79
55. Eaton WA, Thompson PA, Chan CK, Hagen SJ, Hofrichter J (1996) J Struct 4: 1133
56. Pascher T, Chesick JP, Winkler JR, Gray HB (1996) Science 271: 1558
57. Ballew RM, Sabelko J, Gruebele M (1996) Proc Natl Acad Sci USA 93: 5759
58. Williams K, Cansgrove TP, Gilmanshin R, Fang KS, Callander RH, Woodroff WH, Dyer RB (1996) Biochemistry 35: 691
59. Yeh SR, Takahashi S, Fan B, Roussean DL (1997) Nat Struct Biol 4: 51
60. Camacho CJ, Thirumalai D (1995) Proc Natl Acad Sci USA 92: 1277
61. Guo Z, Thirumalai D (1995) Biopolymers 36: 83
62. Thirumalai D (1993) Lectures in statistical mechanics. Protein structure, NATO Workshop, Corsica
63. Zarrinkar PP, Williamson JR (1996) Nat Struct Biol 3: 432
64. Lindahl T, Adams A, Fresco JR (1966) Proc Natl Acad Sci USA 35: 1941; Grosshans CA, Cech TR (1984) Biochemistry 28: 6888
65. Pyle AM (1993) Science 261: 709
66. Crothers DM, Cole PE, Hilbers CW, Schulman RG (1974) J Mol Biol 87: 63
67. Cech TR (1993) In: Gesteland RF, Atkins JF (eds) The RNA world. Cold Spring Harbor Press, New York
68. Emerick VL, Woodson SA (1994) Proc Natl Acad Sci USA 91: 9675
69. Emerick VL, Pan J, Woodson SA (1996) Biochemistry 35: 13469
70. Pan J, Woodson SA, Thirumalai D (1997)
71. Todd MJ, Lorimer GH, Thirumalai D (1996) Proc Natl Acad Sci USA 93: 4030
72. Chan HS, Dill KA (1996) Proteins Struct Funct Genet 24: 345
73. Tsuchihashi Z, Khosla M, Herschlag D (1993) Science 262: 99
74. Coetzee T, Herschlag D, Belfort M (1994) Genes Dev 8: 1575