Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 4030–4035, April 1996
Biochemistry

Chaperonin-facilitated protein folding: Optimization of rate and yield by an iterative annealing mechanism

MATTHEW J. TODD*, GEORGE H. LORIMER*, AND D. THIRUMALAI†

*Central Research and Development Department, DuPont Company Experimental Station, Wilmington, DE 19880-0402; and †Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742

Communicated by I. C. Gunsalus, Environmental Research Laboratory, Gulf Breeze, FL, December 15, 1995 (received for review June 1, 1995)

ABSTRACT **We develop a heuristic model for chaperonin-facilitated protein folding, the iterative annealing mechanism, based on theoretical descriptions of "rugged" conformational free energy landscapes for protein folding, and on experimental evidence that (*i*) folding proceeds by a nucleation mechanism whereby correct and incorrect nucleation lead to fast and slow folding kinetics, respectively, and (*ii*) chaperonins optimize the rate and yield of protein folding by an active ATP-dependent process. The chaperonins GroEL and GroES catalyze the folding of ribulose bisphosphate carboxylase at a rate proportional to the GroEL concentration. Kinetically trapped folding-incompetent conformers of ribulose bisphosphate carboxylase are converted to the native state in a reaction involving multiple rounds of quantized ATP hydrolysis by GroEL. We propose that chaperonins optimize protein folding by an iterative annealing mechanism; they repeatedly bind kinetically trapped conformers, randomly disrupt their structure, and release them in less folded states, allowing substrate proteins multiple opportunities to find pathways leading to the most thermodynamically stable state. By this mechanism, chaperonins greatly expand the range of environmental conditions in which folding to the native state is possible. We suggest that the development of this device for optimizing protein folding was an early and significant evolutionary event.**

The native structure of proteins is thought to represent the global minimum free energy state (1). All of the information needed for folding of the protein is contained in the various solvent-induced interactions between the amino acids of the protein. From this perspective, protein folding reduces to a statistical mechanical description of finite-sized branched heteropolymers.

Despite advances in physical and genetic techniques, no unified theory for protein folding has emerged, perhaps because of the bewildering diversity of folded structures. However, two novel aspects of protein folding have emerged. (*i*) One aspect, grounded in statistical mechanics, provides insight into the kinetics and thermodynamics of protein folding (2–4). (*ii*) The other aspect introduces what seems like an unwelcome complication, the chaperonin proteins. Chaperonins permit the folding of some other proteins, most significantly under *in vitro* conditions that are not permissive for spontaneous folding (5). We show here that these two seemingly disparate approaches to protein folding can be unified by the theory of iterative annealing to account for chaperonin-facilitated protein folding. Theoretical studies have also begun to consider the role of chaperonins in protein folding (6–9).

Two concepts are needed to obtain a coherent picture of the kinetic mechanism of the chaperonins. (*i*) The conformational free energy landscape for protein folding can be rough. There may be kinetic traps in which nonnative structures can accumulate; if deep enough, these traps can retard the appearance of the native state. Especially for large proteins, spontaneous folding of all of the molecules to the native state may not occur *in vitro* on a biologically meaningful time scale.‡ (*ii*) With energy from the coordinated hydrolysis of ATP, the chaperonins provide a mechanism to traverse the landscape, circumventing the barriers separating the misfolded structures and the native state, as suggested on theoretical (6) and on experimental (11, 12) grounds. The interaction of the chaperonin with the substrate protein is stochastic in nature; the appearance of the native state is the consequence of multiple rounds of binding and release (iterations) (11, 12).

Rugged Energy Landscapes and the Kinetic Partitioning Mechanism

It is clear that the underlying conformational free energy landscape of proteins can be rugged, implying that several minima exist separated by barriers of various heights (2–4). This leads to features in the folding process that are not encountered in simpler chemical reactions. A sketch of two free energy landscapes is shown in Fig. 1. Note (*i*) multiple pathways lead from the unfolded random coil to the native state (13), and (*ii*) there can be several energy minima (conformations) that are separated by small barriers of a few kT, where T is the temperature and k is the Boltzmann constant. These barriers may be easily overcome by thermal fluctuations and their presence may not interfere with the formation of the native state (Fig. 1*A*). On the other hand, there can be an ensemble of low energy minima that require large (many kT) activation energies to escape from their trapped states (Fig. 1*B*). The structures in these low energy basins may have many aspects in common with the native state; however, they remain nonnative or misfolded by virtue of incorrect tertiary interactions. Since there can be many such misfolded structures, the collective entropy of this ensemble can easily make up for their higher energy with respect to the native state. Thus, after an initial collapse of the unfolded protein, many of the molecules may become trapped for arbitrarily long times as one or another of these misfolded structures (14). The portion of molecules collapsing to the native state may depend upon the stochastic distribution that undergoes correct versus incorrect nucleation (15, 16).

The qualitative kinetic behavior of the folding protein can be obtained from the rough free energy landscape. In general there are two distinct mechanisms (16, 17) involved in the acquisition of the native conformation (Fig. 2*A*). In the first, described by k_1 in Fig. 2*A*, the chain follows a direct path in which the native structure is acquired rapidly after the collapse transition. We will refer to molecules that follow this path as fast folders and to the pathways involved in this process as

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IAM, iterative annealing mechanism; Rubisco, ribulose bisphosphate carboxylase.

‡Whether or not the *in vitro* rate of spontaneous or chaperonin-facilitated folding of a given protein is sufficient to sustain the growth of the organism is a question that is rarely, if ever, asked (10).

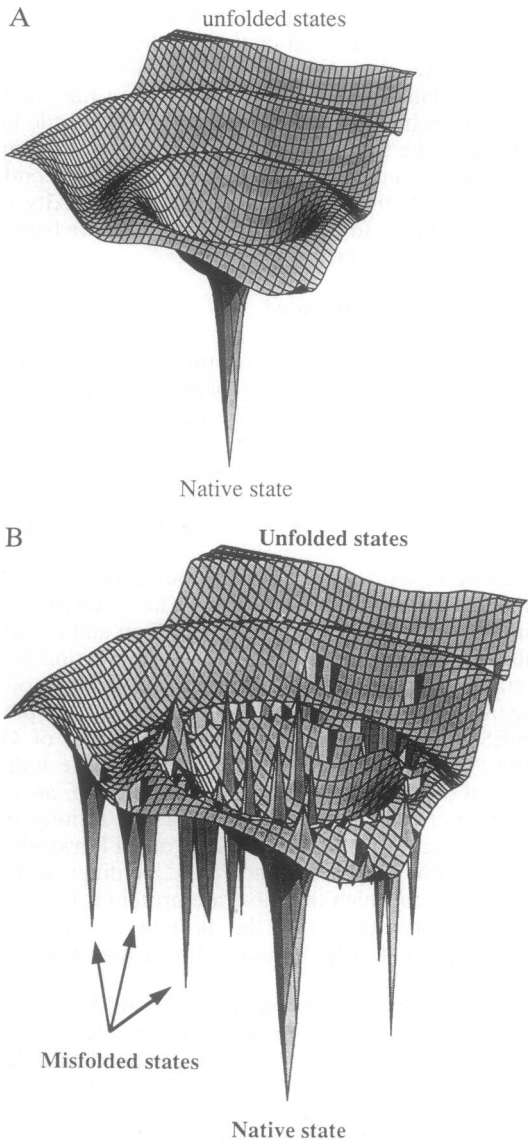

FIG. 1. Three-dimensional energy landscapes depicting multiple (hypothetical) reaction profiles for protein folding. Each folding trajectory starts at the upper rim of the free energy surface and proceeds through a series of conformations that are defined by having a unique subset of native contacts and nonnative contacts. These trajectories converge as they approach the native state where only native contacts remain (center of funnel = native state). A given energy landscape may have no significant energy barriers or only small barriers that can be surmounted by thermal motion (*A*). On such landscapes, the protein folds spontaneously and rapidly to the native state, perhaps without detectable intermediates. For other substrates/conditions, barriers in the energy landscape may exist that preclude progression to the native state on a biologically meaningful time scale (*B*), thus creating kinetically trapped species. Each of these kinetically trapped conformers is defined by a subset of native and nonnative hydrophobic contacts (6). It is proposed that the chaperonin reaction cycle, which is driven by the energy of ATP hydrolysis, lifts these trapped species from their deep free energy valleys by disrupting some of the contacts (in a random fashion) thus increasing the free energy of the nonnative protein. The released conformer may subsequently progress directly to the native state or fall into another (or the same) free energy valley, again resulting in a kinetically trapped species. In the latter case, a further cycle of binding and release by the chaperonin is required. In this view, chaperonin facilitated folding is a process of iterative annealing, with each cycle requiring the hydrolysis of ATP and the release of the substrate protein, still in a nonnative state. Whether the released protein proceeds to the native state or rebinds to the chaperonin depends upon whether it achieves the native state or falls into the energy valley of a misfolded structure.

direct. These pathways involve specific collapse, implying that the topology of the native structure appears simultaneously with collapse, i.e., correct nucleation (Fig. 3). There are several examples of fast folding proteins that become native in less than 10 ms without the formation of detectable intermediates (18–25). In the other mechanism (k_2 in Fig. 2*A*), incorrect nucleation occurs and the chain becomes trapped in one of the low energy minima after collapse (Fig. 3). Since this collapse is nonspecific, the trapped structures represent an ensemble of misfolded species. We will refer to molecules that follow this path as misfolders and to the pathway they take as indirect. These misfolded structures, which may collectively constitute the molten globule state, have to overcome activation barriers before the native conformation is reached. The average activation barrier separating the misfolded structures and the native state is predicted to scale as $N^{1/2}$, where N is the number of amino acids in the protein (26).

We further arbitrarily classify the misfolders into two subclasses, slow folders and no folders, depending on the magnitude of the activation barrier(s) separating the misfolded conformations and the native state. The relevant time scale for slow folders ranges from milliseconds to several minutes. For the no folders, the relevant time scale is in excess of that which is biologically relevant. We stress that the distinction between slow folding and no folding for a given protein is arbitrary and depends upon the external factors such as ionic strength, pH,

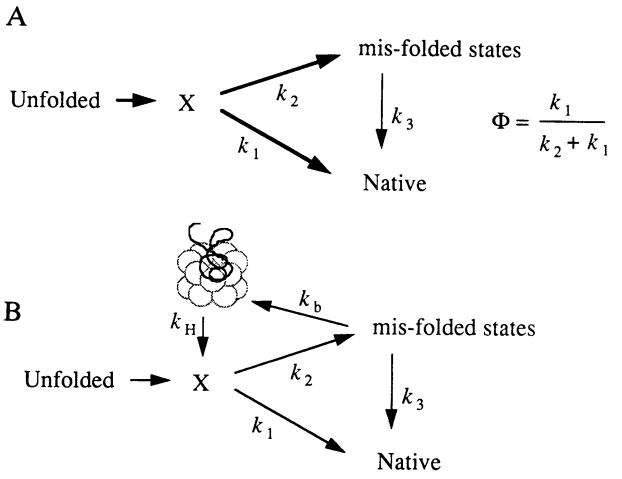

FIG. 2. Kinetic scheme(s) for the protein folding reaction. (*A*) Upon removal of denaturing conditions, unfolded proteins will be in a high-energy state and will rapidly form secondary structures with lower energy (X). These nonnative structures will progress to even lower-energy conformations, partitioning between native and misfolded states (with average rates k_1 and k_2, respectively). A partition factor, Φ, can be defined that represents the fraction of molecules that collapse directly to the native state. Spontaneous folding will be characterized by the sum of the molecules that fold during the initial rapid collapse and those that achieve the native state after activated transitions from the misfolded states (with average rate k_3). It is likely that this process involves partial unfolding of the protein. The conformation of such an expanded protein may correspond to one of the ensemble of structures that constitutes X. The rate for this combined process, which could be stepwise or concerted, is k_3. If the free energy barriers between the misfolded states and the native conformation are large, then k_3 becomes vanishingly small. Therefore, for some proteins, spontaneous folding yields may be quite small. (*B*) Chaperonin-assisted folding introduces an alternative kinetic pathway to escape from the deep free-energy valleys characteristic of misfolded structures. Rapid multivalent binding of misfolded proteins to chaperonins (k_b), the quantized hydrolysis of ATP (k_H) (11), and the release of unfolded proteins (11, 12) may allow sufficient disruption of misfolded structures that repartitioning between native and misfolded states will once again ensue. The iterative nature of this process will ensure that proteins that fold with difficulty (i.e., those with numerous deep kinetic traps) will have multiple opportunities to reach the native state.

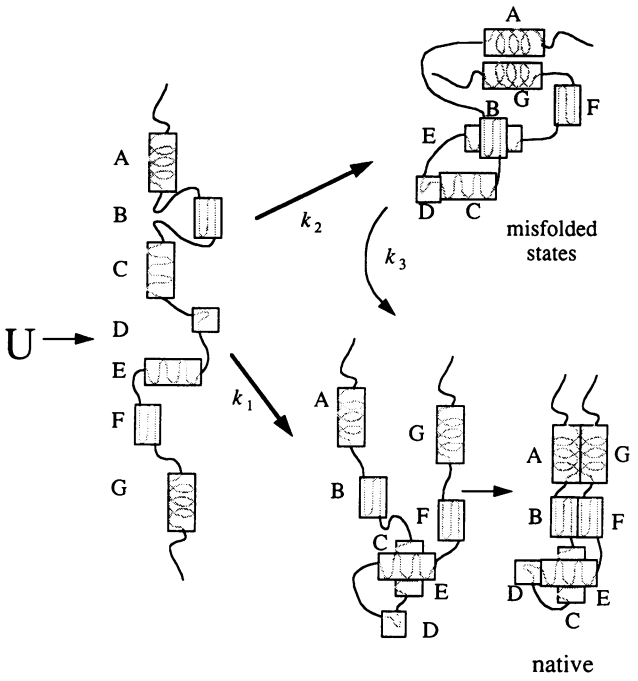

FIG. 3. Unfolded protein (U) rapidly collapses to a species that contains nearly quantitative amounts of secondary structure (represented by diagrams A–G). Subsequent progression along the protein folding reaction coordinate results in a fraction of the population rapidly reaching the native state via correct nucleation (k_1). Folding intermediates that form incorrect tertiary associations reach a collection of misfolded states (k_2). One such (hypothetical) misfolded structure is shown (*Upper Right*), where elements of secondary structure A and G have associated in the wrong direction and/or structures B and E have associated rather than the correct interactions of B with F and C with E (*Lower Right*).

and temperature, etc. (3, 5). For example, ribulose bisphosphate carboxylase (Rubisco) behaves as a no folder at low ionic strength, but as a slow folder at higher ionic strength (5). Likewise, malate dehydrogenase and asparatate transaminase (5, 27, 28) behave as slow folders at 10–20°C, but essentially as no folders at 37–40°C, the physiological temperature. Each of these no folders can be rescued by the complete chaperonin system. An alternative fate of both slow folders and no folders is irreversible aggregation (29).

An important consequence of the kinetic scheme shown in Fig. 2*A* for *in vitro* folding is that the fraction of molecules that reach the native state via the direct or correct nucleation pathway is governed by a partition factor, Φ, that depends on intrinsic factors (amino acid sequence) and extrinsic factors (temperature, pH, ionic strength, etc.). Those arriving indirectly must overcome the energy barrier associated with k_3, which is subject to similar intrinsic and extrinsic factors. As a result, one can obtain a variety of scenarios for protein folding that will depend upon an interplay of k_1, k_2, and k_3 and other processes that occur on a shorter time scale. In general it is difficult to predict Φ theoretically or to measure it experimentally. The *in vitro* folding of a number of small proteins (18, 21, 24, 25) is consistent with the existence of fast and slow folding pathways. A rough free energy landscape thus results in direct and indirect pathways to the native conformation, i.e., a kinetic partitioning mechanism (Fig. 2).

Protein Folding *in Vitro* and *in Vivo*: The Need for Chaperonins

Many proteins fold spontaneously *in vitro* (30). If the search is sufficiently exhaustive, a conformational free energy landscape can usually be found that is smooth enough to permit at least some of the unfolded protein molecules to progress to the native state on an observable time scale. Yet even under supposedly optimal conditions, the yield of active native protein often falls short of 100%. In addition the conditions employed in such experiments are sometimes of little biological relevance: the temperature is often reduced to nonphysiological levels to maximize yield; and the *in vitro* conditions seldom approach the concentration and complexity of the intracellular mileau. In contrast, most proteins are believed to fold *in vivo* with high efficiency.

Properties of the Chaperonins

Chaperonins permit the folding of substrate proteins to occur under nonpermissive conditions, where spontaneous folding is scarcely detectable (5, 27, 28). The chaperonins react with a vast number of nonnative proteins that are structurally unrelated in their native states (31), predominantly via hydrophobic interactions (32, 33). Analyses of proteins complexed to GroEL indicate the absence of higher-order structures (34–36).

The native structure of GroEL is a double toroid, with seven subunits in each ring (32). Each subunit has been proposed to exist in one of two conformations, with high and low affinity for substrate proteins (37–40). The interconversion between these states is driven by a K^+-dependent hydrolysis of ATP, turning over approximately every 10–12 s (38). In the presence of GroES, the hydrolysis of ATP by each ring of GroEL becomes "quantized" (11). This ensures that the individual subunits of a ring change their conformation as a unit, alternating between states of high and low affinity for the substrate protein (11). Multiple rounds of ATP hydrolysis are associated with chaperonin-facilitated folding; with each round, the ring to which the substrate protein is bound cycles to a low-affinity state, and the protein is released in a nonnative state (11, 12). Relaxation to a high-affinity state completes the cycle.

The promiscuity with which chaperonins bind substrate proteins and the release of proteins in nonnative states point to an active albeit nonspecific process rather than a directed process. It has been argued (11) that chaperonins utilize the energy of quantized ATP hydrolysis to facilitate the escape of those protein molecules that are located in deep kinetic traps. This basic idea, when combined with recent theoretical (2–4, 6–9, 16, 17) and experimental work on *in vitro* protein folding (11, 12, 18–25), offers a novel picture of the kinetic action of chaperonins, which we refer to as iterative annealing. The purpose of this paper is to use the theoretical ideas of the rough energy landscape to analyze the consequences of the iterative annealing mechanism (IAM) on protein folding *in vivo*.

The IAM of Chaperonin-Facilitated Protein Folding

The IAM is based on the premise that in the presence of the chaperonins the partially folded protein is offered multiple chances to reach the native state. It was conjectured based on the rough energy landscape perspective of protein folding that the chaperonins rescue those protein molecules that are trapped in deep energy minima (6). The IAM is a biological example of the annealing protocols used in computer simulations of a variety of condensed matter systems (41). The hallmark of these systems is their tendency to become trapped in unfavorable minima. Transitions out of these minima are unlikely on the observational time scale. By raising the temperature of the system, energy can be provided to overcome the barriers and by repeating this procedure one can iteratively approach the global minimum of the system.

We propose that exactly the same thing is accomplished in the folding of proteins in the presence of the complete chaperonin system. By repeatedly allowing the protein to

escape from deep traps, a high yield of folded protein is achieved under otherwise nonpermissive conditions (11, 12). According to the kinetic partition mechanism described above, a fraction of initial molecules, Φ, reach the native state spontaneously and very rapidly ($\approx 10^{-3}$ s) by a specific collapse process, involving the formation of a critical nucleus at a very early stage in the folding process. The remainder of the molecules quickly become trapped in misfolded structures that are stabilized by nonnative tertiary contacts or by nonnative solvent-mediated interactions (Figs. 2 and 3). These misfolded structures rearrange on longer time scales if at all, by overcoming the free energy barriers separating them from other low energy states. Rearrangement must involve at least partial unfolding of the protein molecule; hence, it is a slower process. In some cases, it may not occur on a biologically relevant time scale. However, the misfolded structures can bind to the chaperonin long before these activated transitions are reached. The chaperonin–protein complex undergoes a cycle of ATP hydrolysis and releases the protein in a less folded state; i.e., the free energy of the substrate protein immediately after release from the chaperonin exceeds that of the substrate protein immediately before association with the chaperonin. An important caveat needs to be mentioned here: depending on the affinity of the substrate protein for the low-affinity state, it is possible that some of the protein may not be released from the chaperonin during the period (10–12 s maximum) when the chaperonin ring is in the low-affinity state. However, while the substrate protein is unrestrained, repartitioning of the molecules occurs. A portion (Φ) nucleate correctly and rapidly reach the native state, while the remainder nucleate incorrectly and condense to an ensemble of misfolded structures that are subsequently rebound by the chaperonin. The time scale for folding to the native state or rebinding to the chaperonin can be estimated to be 1–10 ms from a consideration of the chaperonin concentration and the high rates of association (10, 42). In the presence of chaperonins, the rate of folding will be determined by k_H instead of k_3. The cycle of ATP hydrolysis, release, partitioning, and rebinding (Fig. 2*B*) continues, ensuring all of the starting material is converted to the native state.

The IAM accounts for the random nonspecific nature of the interaction of chaperonins with many substrate proteins. By stochastically disrupting misfolded structures, no specificity is required. By releasing these disrupted structures, giving them an opportunity to partition to the native state in free solution, the essential element of Anfinsen's postulate is upheld. The kinetic partition mechanism thus serves as a natural unifying theme that is applicable to *in vitro* and *in vivo* protein folding. The important element that the IAM adds to this scenario is the multiple opportunities accorded to the protein to realize the folding potential inherent in its primary sequence. Two features of this theory are axiomatic. (*i*) Partitioning to the native state should occur on a time scale faster than the rebinding to the chaperonin. (*ii*) It is imperative that the native state of a protein has a much lower affinity for the chaperonin.

It is useful to contrast the IAM to the kinetic proofreading model for chaperonin action (7). In the latter, the chaperonins are assumed to provide a bias in the folding to the native state that is argued to be responsible for the rate enhancement. A bias could theoretically exist in substrate recognition or in the disruption of misfolded structures. Like the kinetic proofreading model, the IAM allows that the chaperones may have lower affinity for species more closely resembling the native state. However, according to IAM, there is a stochastic disruption of the protein bound to GroEL—we can envision no method to selectively disrupt nonnative contacts while leaving native contacts intact. In addition, the kinetic proofreading model was developed by assuming the unfolded protein remained associated with the chaperonin (yet unbound and able to fold) (7), which is inconsistent with experimental evidence (11, 12). The iterative nature of the present model reflects the opportunity of an unfolded protein to explore the entire available conformational space.

An Experimental Test of Two Aspects of Iterative Annealing

Two consequences of the IAM are experimentally testable. (*i*) The chaperonins should be able to rescue misfolded conformers located in the deep kinetic traps that prevent spontaneous folding to the native state; i.e., no folders can be converted to the native state by the complete chaperonin system. (*ii*) The chaperonins should display the properties of a catalyst, enhancing the rate of folding in a manner proportional to the concentration of the chaperonins. This has been difficult to accomplish because chaperonin substrates, nonnative proteins, are particularly susceptible to aggregation. Consequently, molar equivalents of GroEL have generally been necessary to prevent aggregation. At substoichiometric concentrations of chaperonins, efficient folding requires a substrate that neither folds spontaneously nor aggregates. In the present report, we describe conditions for creating kinetically trapped conformers of Rubisco that are long-lived yet can be restored to the native state by the chaperonins. Thus multiple turnovers with substoichiometric amounts of chaperonins could be monitored.

Neutralization of acid-denatured Rubisco could produce stable nonnative conformers that did not fold to the native state and only slowly lost their ability to be folded (Fig. 4*A*). Only upon addition of the complete chaperonin system was folding observed. The maximum yield of folded Rubisco decreased with time, suggesting that these conformers were only transiently stable. However, the rate constant for folding was independent of the delay, consistent with rapid availability of all conformers able to serve as substrates for GroEL. Thus a pool of misfolded Rubisco was generated that was not itself folding competent but could be converted to a folding-competent state by the chaperonins. Similar observations have been made with malate dehydrogenase at 37°C (27, 43).

Likewise, introducing permissive spontaneous folding conditions after a delay allowed a fraction of the Rubisco to fold spontaneously (Fig. 4*B*). If permissive folding conditions were present initially, delayed addition of chaperonins increased the yield substantially. These results can be easily understood in terms of the IAM. For Rubisco, high ionic strength may lower the energy barriers between some of the traps and their respective activation energies. Since the net spontaneous yield remains lower than the chaperonin-assisted yield (40% vs. 80%), at the higher ionic strength, not all molecules escape from their trap. Trapped conformers that cannot be rescued by increased ionic strength, however, can be rescued by the complete chaperonin system. The presence of multiple folding pathways that may be differentially affected by the reaction conditions (while increasing the complexity of the folding reaction) may explain the substoichiometric yield often observed during folding reactions. A portion of these traps may be too deep to exit under a given spontaneous folding condition.

With this transiently stable pool of nonnative protein substrate, the effect of substoichiometric quantities of GroEL on the rate of folding was examined (Fig. 5*A*). At low GroEL concentrations (e.g., 1–2 nM 14-mer) at least 10 turnovers were observed. The reaction kinetics were adequately fitted by a single exponential that results from solving the rate equations describing the IAM (unpublished results). The rate was initially directly proportional to the GroEL concentration (Fig. 5*B*), increasing to a limiting rate at a $GroEL_{14}$/Rubisco molar ratio of unity. From this experimental data the fraction of molecules that has reached the native state as a result of a single iteration (i.e., the partition factor Φ) and the number of

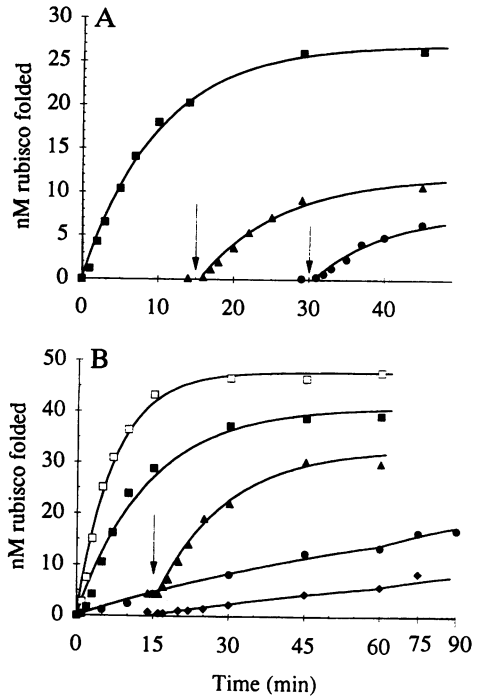

FIG. 4. Stability of acid-denatured unfolded Rubisco after neutralization. (*A*) Acid-denatured Rubisco was rapidly neutralized in folding buffer containing ATP and $GroES_7$; $GroEL_{14}$ was added to initiate folding at 0 (■), 15 (▲), or 30 (●) min. Unfolded Rubisco slowly lost the ability to be recognized by GroEL, as evidenced by the reduced folding yields; however, the Rubisco that was still able to fold did so at a rate independent of the delay before GroEL addition. Twenty microliters of 2.5 μM Rubisco in 10 mM HCl was rapidly added to 1.0 ml of folding buffer (100 mM Hepes/10 mM dithiothreitol/10 mM KOAc/5 mM $Mg(OAc)_2$/1.0 mM MgATP/200 nM $GroES_7$/0.01% Tween) giving a stable suspension of ≈50 nM unfolded Rubisco. $GroEL_{14}$ (125 nM) was either present initially, added at 15 min, or added at 30 min to initiate folding. Aliquots of 100 μl were quenched with glucose/hexokinase [to 5 mM glucose/hexokinase (0.2 mg/ml; 270 units/mg, Sigma)]. Rubisco assays were as described (11). Data was fit to a first-order exponential decay by using KALEIDAGRAPH. First-order decay constants were 0.101 ± 0.006 min^{-1} (GroEL at time zero), 0.092 ± 0.014 min^{-1} (GroEL at 15 min), and 0.103 ± 0.021 min^{-1} (GroEL at 30 min). Rubisco folding in the absence of chaperonins was undetectable. (*B*) Spontaneous Rubisco folding in the presence of 0.25 M NaCl: Delayed GroEL addition (arrow) allows recovery of folding incompetent conformers. Rubisco folding with GroEL present initially [with (■) or without (□) 0.25 M NaCl] or added after 15 min (▲). Spontaneous folding was measured with 0.25 M NaCl present initially (●) or added after 15 min (◆).

iterations can be easily computed. For simplicity, we consider only the limiting case, when $GroEL_{14}$ is at least stoichiometric with the protein substrate. Under these conditions, if a given molecule does not fold in some specific iterative cycle, it rebinds to another $GroEL_{14}$, and losses due to aggregation can be ignored. It is straightforward to show that the fraction of molecules in the native state Ψ_N after n iterations is given by

$$\Psi_N = [1 - (1 - \Phi)^n]. \quad [1]$$

An estimate for the value of n can be made on the basis of the rate at which the complete chaperonin system consumes ATP. In the presence of GroES, each GroEL toroid turns over as a unit at a rate of 2–3 min^{-1}. With this value, the data in Fig. 5*A* can be used to obtain a value for Φ between 0.045 and 0.067. In deriving the above relation, we assume that Φ does not depend on the iteration. This may not always be valid.

Biological Ramifications of the IAM

The IAM, which follows from the analysis of the free energy landscape of proteins, hinges on the assumption that quantized

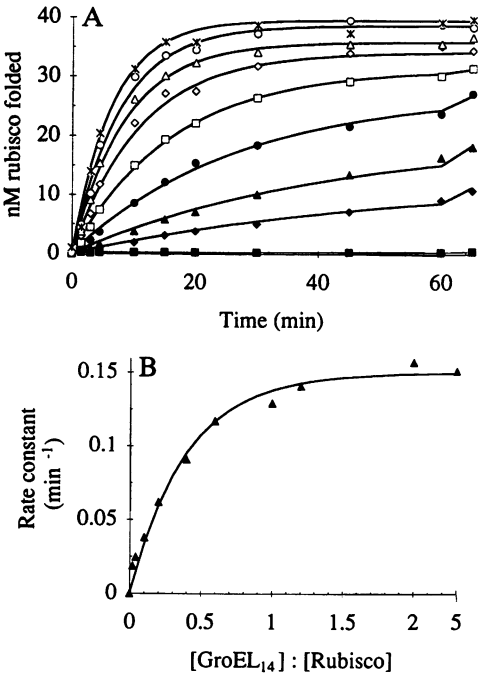

FIG. 5. Rate of Rubisco folding at 25°C is proportional to chaperonin concentration. (*A*) Multiple turnovers occur when aggregation of substrate protein is inhibited. Rubisco (50 nM acid-denatured) was neutralized in folding buffer containing 1 mM ATP, 200 nM $GroES_7$, and 0 (■), 1 (◆), 2 (▲), 5 (●), 10 (□), 20 (◇), 30 (△), 50 (○), or 100 (*) nM $GroEL_{14}$. (*B*) Pseudo first-order rate constant for Rubisco folding as a function of GroEL concentration (k_{app}). Folding assays were done as described in Fig. 4. Results were fit to the equation folded Rubisco(t) = Rubisco (t = 0) + V_i (1 − $e^{-k_{app}*t}$). Background Rubisco activity [$Rubisco_{(t=0)}$] was ≤0.1 nM during the course of these assays.

ATP hydrolysis by the GroEL toroids provides the energy needed to accord misfolded structures additional opportunities to reach the native conformation. It is remarkable that the kinetic partition mechanism, originally introduced to describe *in vitro* folding, can be suitably modified to interpret the kinetics of *in vivo* folding. Our mechanism also serves to emphasize the fact that despite the seeming simplicity of IAM, the possible scenarios for *in vitro* and *in vivo* folding can be quite diverse. This is because the parameters involved in IAM, namely, Φ and various rate constants, depend on both intrinsic and extrinsic factors. It is tempting to suggest that whenever Φ is small and the subsequent traps are deep, chaperonins would be required for proteins to fold with a yield and a rate sufficient to sustain the survival of the organism. For example, when cells are exposed to a heat shock (a higher temperature than normal), nonnative kinetically trapped conformers of various proteins may become populated. The enhanced synthesis of chaperonins, which is part of the heat shock response, provides a mechanism for returning these proteins to their native states.

We have shown that the complete chaperonin system serves to enhance the rate of Rubisco folding, because $\Phi k_H > k_3$ (Fig. 2). However, our kinetic scheme also suggests that under certain conditions, the chaperonins could lead to a retardation of folding rates, i.e., when $k_3 > \Phi k_H$. Such behavior has also been experimentally observed (42, 44).

The indispensability of the chaperonins attests to their biological importance. The IAM of chaperonin-facilitated protein folding confers selective advantage on the organism. The complete chaperonin system optimizes protein folding by permitting progression to the native state under a much wider range of conditions than the spontaneous process allows. But optimization comes at a price—additional ATP must be hydrolyzed. However, a simple calculation shows that the

benefits of the IAM far outweigh the costs. Consider the synthesis of Rubisco, a protein with 450 amino acid residues. By assuming folding conditions are nonpermissive within the cell, one can calculate from the data in Fig. 5*A* that about 20 iterations are required to fold 90% of the Rubisco molecules. Since each quantized iteration requires the expenditure of 7 ATPs, the IAM for Rubisco consumes about 140 ATPs, or 3–10% of the energy expended to form the random coil.

From an evolutionary standpoint, the chaperonins could ameliorate the effects of mutations. Point mutations are thought to alter the energy landscape and folding dynamics (3, 4) because the partition factor Φ depends on the amino acid sequence. Point mutations may increase the fraction of molecules that become kinetically trapped. The ability of chaperonins to rescue these conformers and allow them multiple chances to advance to the most stable state provides an explanation for the observation that plasmids overexpressing chaperonins are able to suppress a wide variety of temperature-sensitive mutations (45, 46). On the other hand, as evolution searches for new biological activities by introducing mutations into copies of existing genes, chaperonins ensure that a mutant protein is able to explore the entire free energy landscape without becoming trapped in local minima. Evolution may occur when enough mutations accumulate that a particular kinetic trap becomes a new energy minimum with altered function. Thus, we conclude, the development of the chaperonins and of the IAM for optimizing protein folding must surely have been an early and significant evolutionary event.

This work was supported in part by a grant to G.H.L from the New Energy and Industrial Technology Development Organization/ Research Institute of Innovative Technology for the Earth (Japan).

1. Anfinsen, C. B. (1973) *Science* **181,** 223–230.
2. Wolynes, P. G., Onuchic, J. N. & Thirumalai, D. (1995) *Science* **267,** 1619–1620.
3. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995) *Proteins Struct. Funct. Genet.* **21,** 167–196.
4. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995) *Protein Sci.* **4,** 561–602.
5. Schmidt, M., Buchner, J., Todd, M. J., Lorimer, G. H. & Viitanen, P. V. (1994) *J. Biol. Chem.* **269,** 10304–10311.
6. Thirumalai, D. (1994) in *Statistical Mechanics, Protein Structure, and Protein–Substrate Interactions*, ed. Doniach, S. (Plenum, New York), pp. 115–134.
7. Gulukota, K. & Wolynes, P. G. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 9292–9296.
8. Sfatos, C. D., Gutin, A. M., Abkevich, V. I. & Shakhnovich, E. I. (1996) *Biochemistry* **35,** 334–339.
9. Chan, H. S. & Dill, K. A. (1996) *Proteins Struct. Funct. Genet.* **24,** 345–351.
10. Lorimer, G. H. (1996) *FASEB J.* **10,** 5–9.
11. Todd, M. J., Viitanen, P. V. & Lorimer, G. H. (1994) *Science* **265,** 659–666.
12. Weissmann, J. S., Kashi, Y., Fenton, W. A. & Horwich, A. L. (1994) *Cell* **78,** 693–702.
13. Harrison, S. C. & Durbin, R. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 4028–4030.
14. Go, N. (1983) *Annu. Rev. Biophys. Bioeng.* **12,** 183–210.
15. Wetlaufer, D. B. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 697–701.
16. Thirumalai, D. & Guo, Z. (1994) *Biopolymers* **35,** 137–140.
17. Guo, Z. & Thirumalai, D. (1995) *Biopolymers* **36,** 83–102.
18. Radford, S. E., Dobson, C. M. & Evans, P. A. (1992) *Nature (London)* **358,** 302–307.
19. Kuszewski, J., Clore, M. G. & Gronenborn, A. M. (1994) *Protein Sci.* **3,** 1945–1952.
20. Elöve, G., Bhuyan, A. K. & Roder, H. (1994) *Biochemistry* **33,** 6925–6935.
21. Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994) *Nat. Struct. Biol.* **1,** 149–156.
22. Jackson, E. & Fersht, A. (1991) *Biochemistry* **30,** 10428–10435.
23. Schnidler, T., Herrler, M., Marahiel, M. A. & Schmid, F. X. (1995) *Nat. Struct. Biol.* **2,** 663–673.
24. Kiefhaber, T. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 9029–9033.
25. Udkaonkar, J. B. & Baldwin, R. L. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 8197–8201.
26. Thirumalai, D. (1995) *J. Phys. (Paris) I* **5,** 1457–1467.
27. Peralta, D., Hartman, D. J., Hoogenraad, N. J. & Høj, P. B. (1994) *FEBS Lett.* **339,** 45–49.
28. Mattingly, J. A., Iriate, A. & Martinez-Carrion, M. (1995) *J. Biol. Chem.* **270,** 1138–1148.
29. Kiefhaber, T., Rudolph, R., Kohler, H. H. & Buchner, J. (1991) *Bio/Technology* **9,** 825–829.
30. Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol.* **49,** 117–237.
31. Viitanen, P. V., Gatenby, A. A. & Lorimer, G. H. (1992) *Protein Sci.* **1,** 363–369.
32. Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Jochimiak, A., Horwich, A. L. & Sigler, P. B. (1994) *Nature (London)* **371,** 578–586.
33. Lin, Z., Schwarz, F. P. & Eisenstein, E. (1995) *J. Biol. Chem.* **270,** 1011–1014.
34. Zahn, R., Spitzfaden, C., Ottiger, M., Wüthrich, K. & Plückthun, A. (1994) *Nature (London)* **368,** 261–265.
35. Robinson, C. V., Gross, M., Eyles, S. J., Ewbank, J. J., Mayhew, M., Hartl, F. U., Dobson, C. M. & Radford, S. E. (1994) *Nature (London)* **372,** 646–651.
36. Okazaki, A., Ikura, T., Nikaido, K. & Kuwajima, K. (1994) *Nat. Struct. Biol.* **1,** 439–446.
37. Jackson, G. S., Staniforth, R. A., Halsall, D. J., Holbrook, J. J., Clarke, A. R. & Burston, S. G. (1993) *Biochemistry* **32,** 2554–2563.
38. Todd, M. J., Viitanen, P. V. & Lorimer, G. H. (1993) *Biochemistry* **32,** 8560–8567.
39. Bochkareva, E. S. & Girshovich, A. S. (1994) *J. Biol. Chem.* **269,** 23869–23871.
40. Kovalenko, O., Yifrach, O. & Horovitz, A. (1994) *Biochemistry* **33,** 14974–14978.
41. Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983) *Science* **220,** 671–680.
42. Grey, T. E. & Fersht, A. R. (1993) *J. Mol. Biol.* **232,** 1197–1207.
43. Ranson, N. A., Dunster, N. J., Burston, S. G. & Clarke, A. R. (1995) *J. Mol. Biol.* **249,** 581–586.
44. Viitanen, P. V., Donaldson, G. K., Lorimer, G. H., Lubben, T. H. & Gatenby, A. A. (1991) *Biochemistry* **30,** 9716–9723.
45. van Dyk, T., Gatenby, A. A. & La Rossa, R. (1989) *Nature (London)* **342,** 451–453.
46. Gordon, C. L., Sather, S. K., Casjens, S. & King, J. (1994) *J. Biol. Chem.* **269,** 27941–27951.