Annu. Rev. Biochem. 2001. 70:603–47

FOLDING OF NEWLY TRANSLATED PROTEINS IN VIVO: The Role of Molecular Chaperones

Judith Frydman
Department of Biological Sciences, Stanford University, Stanford, California 94305-5020; e-mail: jfrydman@stanford.edu

Key Words cotranslational folding, chaperonin, Hsp70, GroEL, TRiC/CCT

■ **Abstract** Recent years have witnessed dramatic advances in our understanding of how newly translated proteins fold in the cell and the contribution of molecular chaperones to this process. Folding in the cell must be achieved in a highly crowded macromolecular environment, in which release of nonnative polypeptides into the cytosolic solution might lead to formation of potentially toxic aggregates. Here I review the cellular mechanisms that ensure efficient folding of newly translated proteins in vivo. De novo protein folding appears to occur in a protected environment created by a highly processive chaperone machinery that is directly coupled to translation. Genetic and biochemical analysis shows that several distinct chaperone systems, including Hsp70 and the cylindrical chaperonins, assist the folding of proteins upon translation in the cytosol of both prokaryotic and eukaryotic cells. The cellular chaperone machinery is specifically recruited to bind to ribosomes and protects nascent chains and folding intermediates from nonproductive interactions. In addition, initiation of folding during translation appears to be important for efficient folding of multidomain proteins.

CONTENTS

One of the greatest challenges facing modern biology is to understand the complex interrelationship between the amino acid sequence of a protein and its three-dimensional shape. The protein folding problem intersects with a number of other areas of contemporary cell and molecular biology, such as translocation into organelles, regulation of protein conformation and function, intracellular protein degradation, and the molecular basis of amyloid diseases. Indeed, because protein folding translates the linear genetic code into three dimensions, it has been called the second, and more intractable, half of the genetic code. Despite considerable effort expended on understanding the relationship between amino acid sequence and native structure, a comprehensive solution to this problem has remained elusive.

Over the past 15 years, the protein folding problem has taken on an additional level of complexity, as it has become clear that folding in vivo differs considerably from folding in a test tube. The difference likely arises from the profound dissimilarity in the conditions under which proteins fold in vivo and in vitro. To ensure that proteins reach the native state with high efficiency, cells have evolved a complex machinery that assists the folding of newly synthesized polypeptides and rescues existing proteins from partial stress-induced denaturation. This machinery includes several conserved protein families, which are collectively termed molecular chaperones because they guide proteins to their "proper" fate but do not remain associated with the final product.

Given the voluminous information accumulated about protein folding, a comprehensive review on this topic exceeds the space limitations of this chapter. This review focuses on the processes that lead to folding of newly translated proteins in the cytosol of prokaryotic and eukaryotic cells, and provides an overview of the components involved and functional principles that govern de novo folding in vivo. The folding of secretory proteins, which occurs following translocation into the endoplasmic reticulum (ER) is not covered (for reviews see 1–3). Although folding of ER proteins is generally coupled to glycosylation and to disulfide bond formation, and thus certain ER chaperones are specialized to assist these processes, there are otherwise remarkable similarities between folding in the ER and the cytosol, which are highlighted throughout this review.

PROTEIN FOLDING IN VITRO AND IN VIVO

Folding experiments in vitro, using chemically denatured proteins, have demonstrated that the information for the three-dimensional structure of a protein is ultimately determined by its amino acid sequence (4). Consequently, the native

state is the thermodynamically most stable conformation of a polypeptide. To reach the native state, the unfolded polypeptide does not sample randomly all possible conformations, but rather proceeds via one or more pathway(s) in which rapid formation (on the order of milliseconds) of compact folding intermediates restricts the conformational space available to the polypeptide (5–8). It appears that removal of hydrophobic side chains away from the aqueous solution to form the hydrophobic core of the folded protein is a major driving force for formation of folding intermediates.

Most in vitro studies of protein folding have used small, single-domain proteins that undergo cooperative and reversible folding reactions. In these experiments, a pure protein is unfolded in denaturant and rapidly diluted into aqueous solution, where it folds spontaneously. From the perspective of a physical chemist, such models are manageable systems, as the folding reaction is not complicated by off-pathway side reactions, such as aggregation, irrelevant to the central question addressed. Consequently, such studies are usually performed under conditions that minimize off-pathway reactions, such as very high dilution and low temperature. These conditions make such studies experimentally tractable but certainly do not represent the conditions prevailing in the cytoplasm or within cellular organelles. Even under ideal conditions, spontaneous folding is observed only for small proteins. The large majority of cellular proteins form kinetically trapped, slow-folding or nonproductive, compact intermediates, which contain a considerable amount of exposed hydrophobic surface and are thus prone to aggregation. As aggregation is driven at least partly by hydrophobic interactions, it is even more pronounced when folding is attempted under the physiological conditions occurring inside the cell.

The interior of a cell is a very crowded and dynamic environment; the effective protein concentration inside a typical cell has been estimated to be as high as 300 mg/ml (9) (for an artist's rendition of bacterial cell cytosol, visit http://www.scripps.edu/pub/goodsell/gallery/cell.html). Because of its high concentration of proteins and other macromolecules, the cytoplasm no longer behaves as an ideal fluid. This nonideal behavior can influence a number of biological processes. Crowding in the cytoplasm gives rise to excluded volume effects, which can result in significant increases (over several orders of magnitude) in the affinities and rate constants of many reactions, particularly association reactions such as aggregation (reviewed in 10, 11). Macromolecular crowding appears to influence several aspects of folding, including substrate binding to chaperones (12) and the properties of folding intermediates (13). For instance, the effect of crowding on the behavior of folding polypeptides has recently been examined using unfolded reduced lysozyme, a small protein that can fold spontaneously in an aqueous buffer (14). Strikingly, when refolding was performed under crowding conditions, lysozyme aggregated. Interestingly, lysozyme folding was restored by inclusion of the chaperone protein disulfide isomerase, consistent with the view that one role of chaperones may be to prevent aggregation under the conditions that exist in the cytosol (15). Even though the exact properties of the cellular milieu are difficult

to determine experimentally, more studies are needed to elucidate how crowding influences folding in vivo.

Another major difference between refolding of a chemically denatured protein and folding as it occurs in the cell is that in vivo, newly synthesized polypeptides enter the cytosol vectorially because the N terminus is synthesized before the C terminus. Although the N terminus of a translating polypeptide is available for folding before the rest of the polypeptide, stable folding of a folding unit or domain cannot occur until it is completely synthesized (reviewed in 16, 17). The rate of chain elongation in eukaryotes is approximately 4 amino acids per second (18), so a polypeptide of 1000 residues would require more than 4 minutes for its synthesis. In contrast, formation of collapsed folding intermediates occurs on a time scale of seconds. Thus, protein synthesis is considerably slower than formation of collapsed intermediates, which could bury hydrophobic aggregation-prone stretches of the nascent chains away from the solution. Although a nascent polypeptide cannot fold stably, neither can it remain extended, particularly in the cytosol. This situation poses a problem for nascent chains emerging from the ribosome. Nascent chains must avoid forming misfolded intermediates and avoid aggregation with other nascent chains. Aggregation appears to be a protein-specific phenomenon. Thus, identical chains synthesized on neighboring ribosomes in a polysome may be especially prone to aggregation (19). Hence, another potential function of chaperones may be to prevent aggregation that would result from nonnative interactions betweeen nascent chains.

Finally, another distinctive feature of the cellular milieu is its high degree of compartmentalization. This property was revealed by fluorescence recovery after photobleaching (FRAP) experiments, where a section of a mammalian cell loaded with fluorescently labeled protein or a small solute is photobleached and the rate of diffusion is measured as the fluorescent molecules redistribute back into the bleached section (20, 21). Such experiments indicated that whereas small solutes ($\leq$500 daltons) diffuse freely in the cytosol, diffusion of larger macromolecules, including proteins, is greatly retarded. These findings imply that most cellular processes, including chaperone-mediated folding, are highly compartmentalized and spatially organized.

Differences in folding conditions in vivo and in vitro lead to a number of interesting questions, which recent work has begun to address. Clearly, molecular chaperones have evolved to assist protein folding in vivo, both for newly translated proteins and for proteins that become denatured during thermal stress and cellular injuries. Because of their role in rescuing cellular proteins from denaturation, several classes of chaperone proteins are induced under conditions of stress, including high temperature (hence the generic name Hsp, for heat shock protein). However, the processes of folding newly made polypeptides and rescuing stress-denatured ones appear to differ both topologically and mechanistically (e.g. 22). This review focuses on processes that lead to folding of newly translated proteins in the cytosol of prokaryotic and eukaryotic cells.

CHAPERONE SYSTEMS IN EUKARYOTIC AND PROKARYOTIC CELLS

The idea that a class of proteins facilitates folding originally emerged from studies of protein folding inside chloroplasts and mitochondria (reviewed in 23). These experiments were followed by about 10 years of intense research into the identity and mechanisms of the cellular folding helpers and catalysts. Mounting evidence now indicates that molecular chaperones interact with, and stabilize, nascent and translocating polypeptides in vivo and prevent nonproductive reactions such as aggregation. Two major classes of ATP-dependent chaperones, the Hsp70s and the cylindrical chaperonin complexes, have been implicated in de novo protein folding in the cytosol of eukaryotic and prokaryotic cells, as well as in organelles of presumed endosymbiotic origin, such as mitochondria and chloroplasts (reviewed in 24, 26). Although substrate binding by both of these chaperone systems is regulated by ATP binding and hydrolysis, the Hsp70s and the chaperonins are structurally and functionally distinct and represent radically different principles of chaperone action. Because of their greater simplicity, prokaryotic chaperone systems are better understood than their eukaryotic counterparts, in terms of both mechanism and cellular function. The mechanistic aspects of these chaperone systems have been summarized in several excellent reviews (25, 27–31). In this section, a brief overview is provided of the basic mechanisms and substrate binding properties of the chaperone systems implicated in the folding of newly translated proteins in vivo.

The Hsp70 System

Hsp70s are a highly conserved family of proteins, distributed ubiquitously in all prokaryotes and in cellular compartments of eukaryotic organisms. Some compartments contain multiple Hsp70 homologs with distinct cellular functions (32) (Table 1). For instance, the cytosol of the yeast *Saccharomyces cerevisiae* contains four functionally redundant Hsp70 homologs, called Ssa1, Ssa2, Ssa3, Ssa4 (herein Ssa1–4) and three ribosome-associated Hsp70s, called Ssb1, Ssb2, and Pdr13 (or Ssz1) (33–35). Genetic studies indicate that only Ssa-type function is essential for viability and that Ssb activity cannot substitute for Ssa activity (32). Most Hsp70s have a molecular mass of approximately 70 kilodaltons (kDa) and consist of two functionally coupled domains, which have been crystallized separately (Figure 1*a*). The 44-kDa N-terminal domain mediates ATP binding (36), whereas the 18-kDa C-terminal domain binds the substrate polypeptide (37). Binding and release of the substrate rely on modulation of the intrinsic peptide affinity of Hsp70 by cycles of ATP binding and hydrolysis by the N-terminal domain (25, 26). In the ATP-bound state, Hsp70 binds and releases substrates rapidly, whereas the ADP-bound form binds and releases slowly. In *Escherichia coli*, cycling of the Hsp70 homolog, DnaK, between its different nucleotide-bound

TABLE 1 Members of the Hsp70 and chaperonin systems involved in de novo folding

Chaperone family	Bacterial	Yeast	Mammalian
Hsp70 70-kDa ATPases, bind extended polypeptides enriched in hydrophobic amino acids	DnaK: de novo folding and recovery from stress, binds nascent chains	Ssa1—4: de novo folding and recovery from stress Ssb1—Ssb2: binds ribosomes, nascent chains Pdr13p/Ssz: binds ribosomes	Hsc70: constitutive, binds nascent chains Hsp70: stress inducible — —
Hsp40 40-kDa cofactors that stabilize Hsp70-substrate interactions by stimulating ATP hydrolysis	DnaJ: has chaperone activity, interacts with DnaK	Ydj1: has chaperone activity, interacts with Ssa1 Sis1: ribosome-associated, has chaperone activity, interacts with Ssa1 Zuotin: ribosome-associated, interacts with Pdr13p/Ssz	Hdj2: binds newly translated proteins Hdj1, Hsp40: binds nascent polypeptides
Nucleotide exchange factors	GrpE: 23-kDa Hsp70-nucleotide exchange factor, interacts with prokaryotic Hsp70	—	—
	—	Sti1: 60-kDa protein, links Hsp70 and Hsp90, may promote nucleotide exchange	Hop, p60: Hsp70-organizing protein, Sti1 homolog
	—	Bag homologs: present in *Saccharomyces pombe*; not detected in *S. cerevisiae*	Bag1—Bag5: family that contains a Bag domain with Hsp70-nucleotide exchange activity

ADP-stabilizing factor	—	—	Hip: binds newly translated proteins, stabilizes Hsp70-ADP complex
Hsp70-like chaperones Stabilize extended polypeptides; prevent aggregation	Trigger factor: ribosome-associated, binds nascent chains, has prolyl isomerase activity	—	—
	—	GimC/prefoldin: mechanism unknown, possibly TRiC cofactor; 6 subunits Gim1-Gim6	GimC/prefoldin: binds to nascent chains
Chaperonin/Hsp60 Barrel-like structures that provide a protected folding environment for substrates. Promote ATP-dependent folding in the chamber of the double-ring complex	GroEL/Hsp60: 800-kDa, 14-mer complex, de novo folding and recovery from stress	TRiC/CCT: 1000 kDa, 16-mer hetero-oligomeric ring complex, has built-in lid; 8 subunits Cct1-Cct8	TRiC/CCT: same as yeast complex, assists de novo folding, binds nascent chains
	GroES/Hsp10: 7-mer ring complex, associates with GroEL, acts as a lid for the folding chamber	—	—

—: No homolog detected.

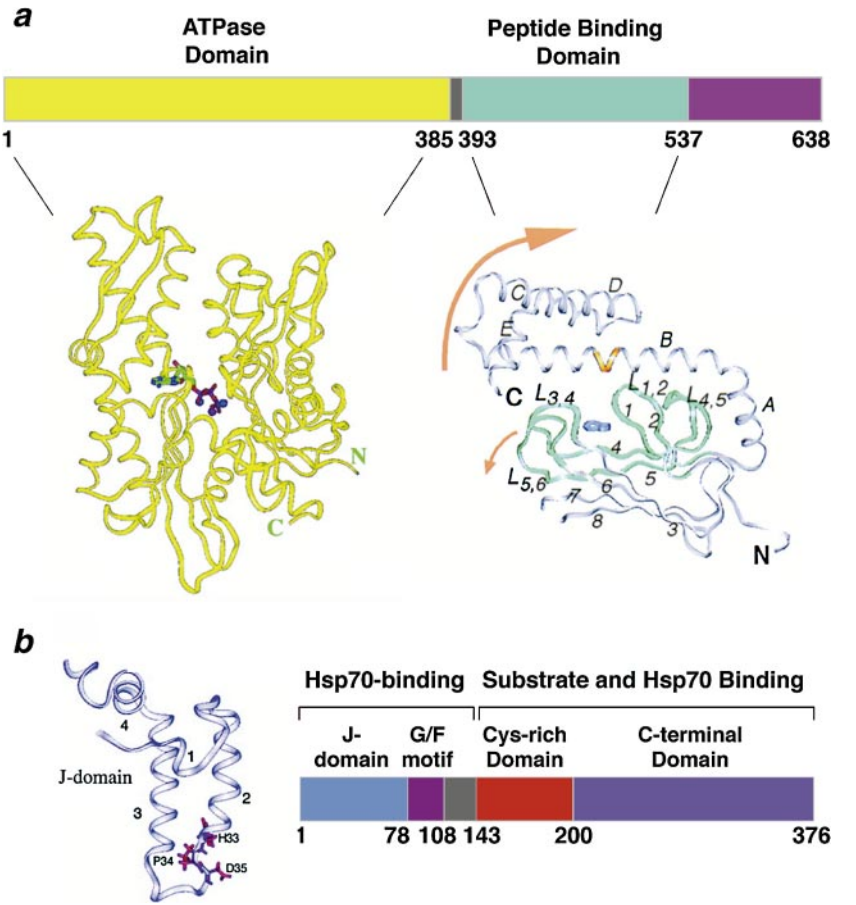


Figure 1 Domain structure of the Hsp70 and DnaJ chaperones. (*a*, *Top*) Domain structure of the bacterial Hsp70, DnaK. Residue numbers correspond to the approximate domain boundaries of DnaK. The structures of the ATPase domain and the peptide binding domain are highly conserved among all prokaryotic and eukaryotic Hsp70s. (*a*, *Lower left*) Ribbon diagram of the ATPase domain of bovine Hsc70 (heat-shock cognate protein), with bound ADP, Pi, Mg^{2+} (magenta), and two K^+ ions (blue). (*a*, *Lower right*) Ribbon diagram of the substrate binding domain of DnaK, highlighting the peptide substrate (small blue ovals) and the β-strands and loops lining the substrate binding pocket (blue green; the rest of the domain is light blue). The helical lid of the binding site is thought to open and close (orange arrows) in response to ATP-mediated conformational changes in the ATPase domain. (*b*, *Right*) Domain structure of bacterial DnaJ. G/F motif: Glycine-phenylalanine-rich region; Cys-rich domain: cysteine-rich, Zn-binding domain. (*b*, *Left*) The ribbon diagram of the J-domain of *E. coli* DnaJ, highlighting (in red) the side chains of the HPD tripeptide motif essential for Hsp70 binding. (Structures modified from Reference 25, with permission.)

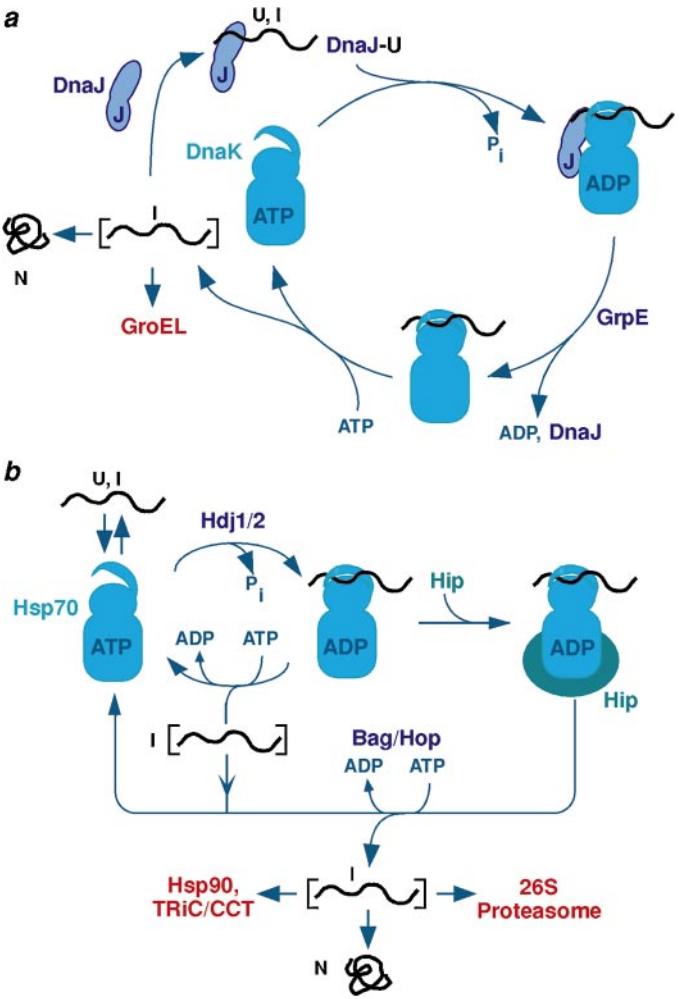
states is regulated by two cofactors, DnaJ and GrpE (Figure 2*a*) (25, 31, 38–40). The 41-kDa DnaJ protein is itself a chaperone, which can bind to unfolded polypeptides and prevent their aggregation (Figure 1*b*) (41, 42). DnaJ binds to DnaK and stimulates its ATPase activity, generating the ADP-bound state of DnaK, which interacts stably with the polypeptide substrate (Figure 2*a*) (43). The 23-kDa GrpE protein acts as a nucleotide exchange factor; it binds to the ATPase domain of DnaK and, by distorting the nucleotide binding pocket, induces release of bound ADP (Figure 2*a*) (44). Finally, rebinding of ATP triggers dissociation of the DnaK-substrate complex (Figure 2*a*).

How does Hsp70 recognize unfolded polypeptides? The peptide-binding site of Hsp70 contains a cleft that binds the peptide in an extended conformation (Figure 1*a*) (37). Analysis of the substrate specificity of Hsp70 using a number of different approaches, including phage display and synthetic peptide libraries, indicated that this chaperone recognizes linear polypeptide sequences enriched in hydrophobic amino acids (45–47). Because of its hydrophobic nature, this binding motif would typically be located in the interior of a correctly folded protein; consequently, surface exposure of such a sequence may be a distinctive feature of nonnative conformations. Such hydrophobic regions are probably present in all unfolded polypeptides, and it has been predicted that an Hsp70-binding site occurs, on average, every 40 residues (45). Association with Hsp70 results in the stabilization of the substrate protein in an extended conformation, thereby preventing its aggregation. In vitro, polypeptides can undergo multiple rounds of binding and release from Hsp70 (Figure 2*a*). This process is sufficient to promote folding of some model substrates, such as firefly luciferase (40). However, in many instances, the Hsp70-bound substrate must be transferred to another type of chaperone, such as a chaperonin complex, for productive folding (see below, Cooperation Between Different Chaperone Systems).

Elucidation of the DnaK reaction cycle has provided a paradigm for all Hsp70s. In fact, homologs of bacterial DnaJ, collectively called Hsp40 proteins, have been identified in all cellular compartments that contain an Hsp70 (Table 1) (48). All DnaJ homologs possess a so-called J-domain, a scaffold of four α-helices and a solvent-accessible loop region that exposes a conserved tripeptide (His-Pro-Asp) essential for interaction of the J-domain with Hsp70 (Figure 1*b*) (49, 50). J-domain-containing proteins can stimulate ATP hydrolysis by Hsp70 and generate the ADP-bound state. In eukaryotic cytosol, the DnaJ homologs, Hdj1 (also called Hsp40, or Sis1 in yeast) and Hdj2 (Ydj1 in yeast), regulate the activity of Hsp70 homologs (see below) (51–53). In addition to the N-terminal J-domain, Ydj1 (Hdj2) and Sis1 (Hdj1) contain C-terminal chaperone domains that bind unfolded polypeptides (52–54). In Ydj1, this substrate binding domain contains two essential cysteine-rich, Zn^{2+}-binding domains, also found in DnaJ. The C-terminal domain of Sis1 has recently been crystallized; its structure reveals several patches of hydrophobic side chains that are required for substrate binding (55). In addition to these bona fide DnaJ homologs, several eukaryotic proteins contain only a J-domain, which serves to recruit Hsp70 family members to specific cellular

sites (48). For instance, the ER membrane protein, Sec63, binds to the luminal Hsp70 (BiP/Kar2) via a J-domain (56). Similarly, auxilin, a mammalian protein associated with the clathrin coat of endocytic vesicles, interacts through its J-domain with cytosolic Hsp70, which is required for the ATP-dependent uncoating of the vesicles (57). Interestingly, a ribosome-associated, J-domain-containing protein called zuotin interacts with the cytosolic Hsp70 homolog Pdr13 (58, 59) (see below, Cooperation Between Different Chaperone Systems).

In striking contrast to DnaJ homologs, GrpE-related proteins in eukaryotes appear to be restricted to compartments of prokaryotic origin, i.e. mitochondria and



chloroplasts. Absence of a GrpE-like nucleotide exchange factor in eukaryotic cytosol can be explained by differences in the nucleotide-binding and hydrolysis properties of DnaK and Hsp70. ADP release is the rate-limiting step when DnaK is associated with DnaJ (38), and maximal stimulation of the ATPase activity of DnaK requires both DnaJ and the ADP-ATP exchange factor GrpE. In contrast, ATP hydrolysis appears to be the rate-limiting step in eukaryotic Hsp70s, and the steady-state ATPase activity of Hsp70 is significantly stimulated by Hsp40 (Hdj1) alone (52, 53) (Figure 2*b*). Interestingly, mammalian cells contain a 48-kDa Hsp70-interacting protein (termed Hip) that binds to the ATPase domain of Hsp70 and prevents ADP release from Hsp70 (60). It thus appears that the regulation of eukaryotic Hsp70 differs from DnaK in significant aspects and is considerably more complex (Figure 2*b*) (61). Similar to DnaJ in the bacterial reaction cycle, Hdj1 or Hdj2 (Hsp40) stimulate the Hsp70 ATPase and generate the ADP-bound form with a high affinity for substrate. However, whereas GrpE promotes nucleotide exchange on DnaK, the eukaryotic Hsp70 complex appears to be the target of multiple regulatory factors. Hip binding slows dissociation of ADP from Hsp70 (Figure 2*b*). Through stabilization of the ADP-bound conformation of Hsp70, Hip presumably stabilizes the Hsp70 substrate complex. The action of Hip appears to be antagonized by several proteins that stimulate nucleotide exchange by Hsp70 and thus serve as functional GrpE analogs in the eukaryotic cytosol. For

←

Figure 2 Model of the Hsp70 reaction cycle in the (*a*) bacterial and (*b*) mammalian cytosol. (*a*) In the bacterial system, the cycle starts with association of DnaJ with an unfolded polypeptide substrate (U) or a folding intermediate (I). DnaK enters the cycle in an ATP-bound conformation, characterized by rapid substrate association and dissociation kinetics. DnaJ binds to DnaK and stimulates conversion of the bound ATP to ADP, which stabilizes interaction of DnaK with the polypeptide substrate. GrpE protein promotes release of bound ADP from DnaK. Subsequent rebinding of ATP triggers dissociation of the polypeptide substrate in an intermediate conformation (I), which may fold to the native state (N), undergo another cycle of interaction with the DnaK chaperone system, or bind to a different chaperone system (e.g. GroEL). (*b*) Eukaryotic DnaJ homologs, Hdj1 or Hdj2, bind to unfolded polypeptides and promote the interaction of mammalian Hsp70 with the polypeptide substrate by driving Hsp70 into the ADP-bound conformation. In contrast to the bacterial cycle, ADP release is not a rate-limiting step when Hsp70 is in a complex with Hdj1 or Hdj2. Thus, ADP-to-ATP exchange and dissociation of the polypeptide substrate can occur in the absence of a GrpE homolog. This minimal reaction cycle may be sufficient for Hsp70 to fulfill many of its functions in the eukaryotic cytosol. However, the ADP-bound state of Hsc70 may be stabilized by the Hip protein, thereby increasing the half-life of the chaperone-substrate complex. Subsequent complex dissociation appears to be initiated by Bag homologs (and perhaps also by Hop), which stimulate ADP-to-ATP exchange, resulting in recycling of Hsp70. The released substrate may either fold, rebind to Hsp70, or bind to other cytosolic components (e.g. TRiC, Hsp90, or the 26S proteasome). (Modified from Reference 121 with permission.)

example, the Hsp70-interacting protein, Bag1 (62–64) appears to promote release of the bound ADP from Hsp70 and results in substrate release (Figure 2*b*). The Hsp70- and Hsp90-interacting protein, Hop (also called p60, or Sti1 in yeast) has also been proposed to stimulate nucleotide exchange by Hsp70 (65). Bag1 is a member of a large family of homologs that contain a conserved domain called the Bag-domain and variable N-terminal extensions (66). It has been proposed that, like GrpE, the Bag-domain stimulates nucleotide exchange by Hsp70 (64, 67, 68).

The modular domain structure of these Hsp70 regulators raises the interesting possibility that nucleotide exchange and subsequent substrate release may be coupled to downstream cellular processes (Figure 2*b*). For instance, Bag1 contains a ubiquitin-homology domain in its N-terminal extension and has been proposed to direct Hsp70-bound substrates to the 26S proteasome (69, 70). In the case of Hop, the substrate released from Hsp70 may be directed to bind to Hsp90 (see below, The Hsp90 System). Similarly, the N-terminal extensions of Bag homologs may play a role in directing the released polypeptide toward different folding or degradation pathways.

"Small" or "Holding" Chaperones: Complexes with an Hsp70-Like Function

Recent findings revealed that prokaryotic and eukaryotic cells contain chaperone complexes that overlap functionally with the Hsp70 system. These complexes have been collectively termed small chaperones to distinguish them from the ring-shaped chaperonin complexes (71). The small chaperones share with Hsp70 the capacity to "hold" or stabilize unfolded polypeptides and prevent their aggregation, but unlike Hsp70, their binding activity is not regulated by nucleotides. In *E. coli*, the activity of the DnaK system in folding newly translated proteins is functionally redundant with a prolyl-isomerase trigger factor (TF) (72, 73) (see below, Contribution of Hsp70 and Other Small Chaperones to de Novo Folding). TF was originally isolated as a protein that facilitated export of proteins into the periplasmic space (74). It can bind unfolded polypeptides and stabilize them against aggregation (75). Although TF displays prolyl-isomerase activity in vitro, its chaperone activity does not require the presence of proline residues in the substrate (75, 76). The binding specificity of TF has been examined using a cellulose-bound peptide array (76). Similar to DnaK, it displays a preference for linear determinants of approximately five residues enriched in hydrophobic side chains; but unlike DnaK, it appears to favor aromatic side chains (see above, The Hsp70 System).

It has been suggested that in eukaryotic cells and archaea, a 200-kDa hetero-oligomeric complex, termed the Gim complex (GimC) or prefoldin, acts like Hsp70 to stabilize nascent chains (77–79). The role of this complex in cellular folding is poorly understood, but *S. cerevisiae* lacking functional GimC are defective in actin and tubulin folding (77, 78). In support of an Hsp70-like function, biochemical and structural analysis using purified GimC from the archaebacterium

Methanobacterium thermoautotrophicum indicated that GimC/prefoldin is a baseball glove–shaped oligomeric complex that can bind unfolded polypeptides, prevent their aggregation, and transfer them to a chaperonin complex (80, 80a) (see below, Cooperation Between Different Chaperone Systems).

Ring-Shaped Chaperones: The Chaperonins

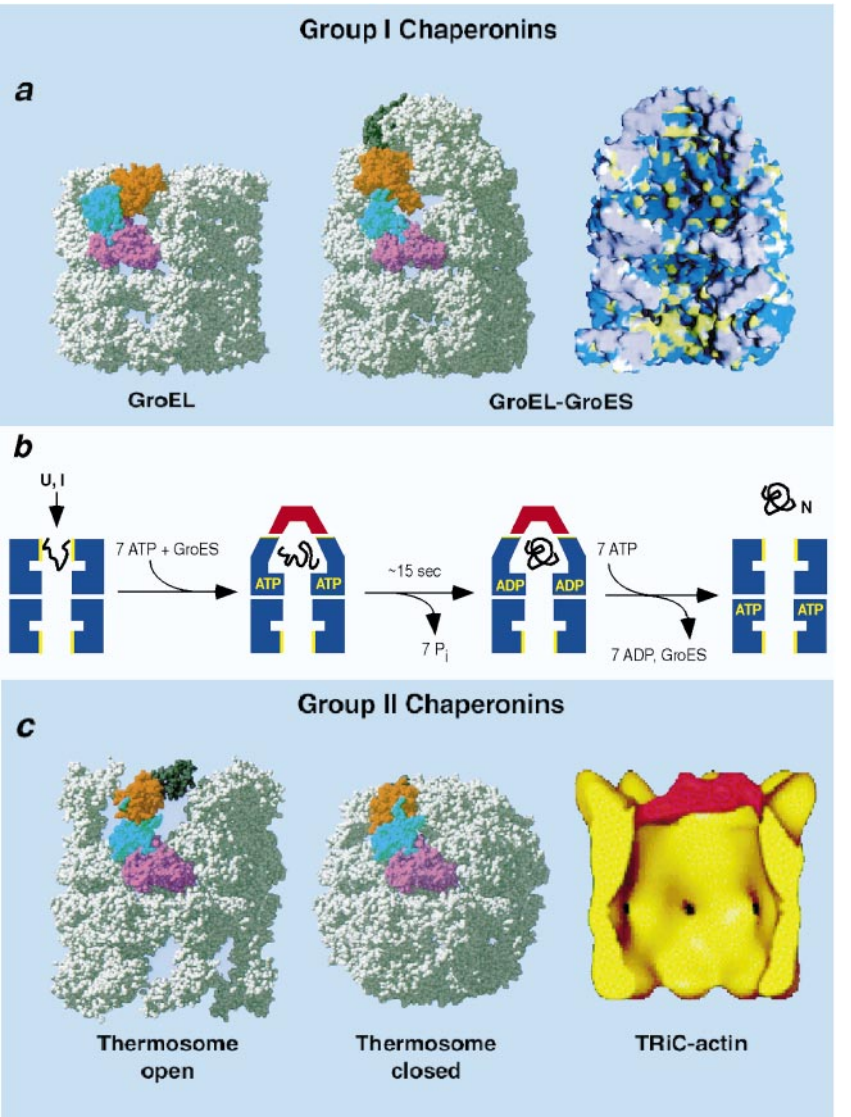
The chaperonins are large cylindrical protein complexes consisting of two stacked rings (reviewed in 25, 26, 28, 81, 82). In some chaperones, the rings have seven subunits, in others eight or nine. Chaperonins differ substantially from Hsp70 in architecture, as well as in their mechanism. However, as in Hsp70, ATP binding and hydrolysis also produce conformational changes that drive cycles of substrate binding and release.

There are substantial differences between group I chaperonins, found in eubacterial cells, and the distantly related group II chaperonins found in Archaea and Eukarya (see Table 1). Group I chaperonins, such as GroEL of *E. coli* and Hsp60 in mitochondria and chloroplasts, function in conjunction with a ring-shaped cofactor, GroES or Hsp10, that forms the lid on a folding cage in which polypeptide substrates are enclosed during folding (83, 84). In contrast, such a cofactor has not been found for group II chaperonins.

Group I Chaperonins The group I chaperonins are perhaps the best characterized of all chaperones, though significant questions remain about their mechanism (reviewed in 25, 26, 81, 85). The Hsp60 chaperonins from chloroplasts and mitochondria were the first complexes implicated in oligomeric assembly (23) and protein folding (23a). Structurally, the *E. coli* chaperonin, GroEL, contains 14 identical subunits arranged in two stacked rings of seven subunits each (Figure 3*a*). The ring-shaped structure of GroEL is essential for its folding activity (86, 87), which allows it to promote the folding of substrates that the Hsp70 system is unable to fold. Each subunit of GroEL consists of two discrete domains, joined by a hinge-like intermediate domain (88). The equatorial domains contain the ATP-binding pocket, whereas the apical domains contain a patch of hydrophobic amino acids that face the interior of the cavity and bind the unfolded substrate polypeptide through hydrophobic contacts. Unlike Hsp70s, GroEL does not bind linear peptides, but interacts efficiently with nonnative proteins. Binding to GroEL appears to be multivalent, and bound folding intermediates presumably expose hydrophobic surfaces that allow them to interact with several GroEL subunits simultaneously (89–91).

The substrate-binding residues in the apical domain of GroEL are also responsible for interacting with the cofactor GroES, a ring-shaped complex composed of seven identical subunits, which is essential for GroEL-mediated folding (92). In the presence of ATP, GroES binds to GroEL and induces a conformational change in the apical domains that displaces the substrate from its binding sites. The substrate is thus released it into the central cavity, which is now lined with hydrophilic side

chains (Figure 3*a*, *right*) (93). GroES also promotes ATP hydrolysis in the proximal (*cis*) ring of GroEL. Enclosure of the substrate polypeptide within the chamber of this GroEL-GroES-ADP complex is essential for folding (Figure 3*b*). The substrate remains enclosed in this cavity for approximately 15 seconds. It has been proposed that the enclosed cavity functions as an Anfinsen cage, i.e. a protected



chamber, that isolates the polypeptide under conditions of infinite dilution and allows it to fold according to its thermodynamic potential (94). There is negative cooperativity between the two GroEL rings; the GroEL-ADP-GroES complex is dissociated by ATP binding to the *trans* ring (Figure 3*b*) (95). GroES release returns the apical domains to the conformation that exposes their hydrophobic binding sites toward the cavity, which permits a still folded polypeptide to rebind and undergo another cycle of folding. However, if the substrate has achieved a folded conformation, it will no longer expose sufficient hydrophobic surfaces to mediate binding and will be released. Measurements of the kinetics of folding in vitro and in vivo indicate that some substrates, such as rhodanese, do interact with GroEL through several nucleotide-driven cycles of GroES interaction before reaching the folded state (96, 97). Despite our detailed structural and mechanistic understanding of GroEL, a fundamental question remains: How does confinement of the substrate to a space that alternatively exposes hydrophobic and hydrophilic surfaces facilitate polypeptide folding? Clearly, GroEL prevents aggregation and provides a sequestered folding cage for newly synthesized chains; however, it is not clear whether in addition it acts to unfold trapped intermediates and thus speeds up the rate of folding of some polypeptides (98).

←

Figure 3 Structure and function of the group I and group II chaperonin systems. (*a*) Domain arrangement and conformational states of group I chaperonins, shown as space-filling models for GroEL and GroES. (*Left*) The nucleotide-free state of GroEL. The domain organization of the GroEL protomer is highlighted in color, showing the equatorial (magenta), hinge or intermediate (cyan), and apical (orange) domains. (*Center*) The ADP-bound state of GroEL-GroES. The dark green area (*upper left*) highlights one GroES protomer in the GroEL-GroES complex. (*Right*) The same ADP-bound structure in the center as an interior view; the GroEL-GroES-(ADP)7 complex is sliced vertically along the central axis and colored to reflect the relative hydrophobicity of the interior surface. Hydrophobic atoms of the side chains are light green; polar and charged atoms of the side chains are dark purple; solvent-excluded surfaces at the interfaces with the missing subunits are light blue; and exposed backbone atoms are white. Note the hydrophilic nature of the GroES-enclosed cavity. (*b*) Schematic model of the GroEL-GroES reaction cycle for folding. GroES (red) and GroEL (blue) are sliced vertically to highlight the domain structure of individual subunits. The hydrophobic sites of substrate binding in the apical domains are highlighted in yellow. (*c*) Domain arrangement and conformational states of group II chaperonins. (*Left, center*) Space-filling models for the thermosome from *T. acidophilum*. The domain organization of the α-protomer is highlighted as in (*a*), showing the equatorial (magenta), intermediate (cyan), and apical (orange) domain. The helical protrusion is shown in dark green. (*Left*) A model for the nucleotide-free, open state of the thermosome based on electron tomographic studies on the α-only thermosome. (*Center*) The closed conformation observed in the X-ray structure of the thermosome (which presumably reflects the nucleotide-bound state). (*Right*) An interior view of the complex of TRiC (yellow) with its substrate actin (red) obtained by cryoelectron microscopy and sliced vertically along the central axis. (Structures and diagram modified from References 19, 28, 93 and, 107 with permission.)

Group II Chaperonins The group II chaperonins are more heterogeneous in sequence and structure than members of group I (reviewed in 28, 99–101). Though group II members also have a double ring structure, they are hetero-oligomeric and the number of subunits varies between eight and nine per ring (Figure 3*c*). Archaeal forms have two or three different subunits per complex (arranged respectively in eight- or nine-fold symmetrical rings). The eukaryotic chaperonin, named TRiC (for tailless complex polypeptide-1 [TCP-1] ring complex) or CCT (for chaperonin containing TCP-1), is also ring shaped and consists of eight different, yet homologous, subunits per ring, ranging between 50 and 60 kDa (102–104). Crystal structure of the archetype group II chaperonin, the thermosome complex from *Thermoplasma acidophilum*, revealed that individual subunits have a domain arrangement similar to those in GroEL (105, 106). Indeed, the equatorial (ATP-binding) domain is relatively conserved among all chaperonins. Most sequence divergence between subunits is found in the apical domains, which probably contain the substrate binding sites (99). Strikingly, the backbone trace of the apical domains of the thermosome is almost identical to that of GroEL, but it has an insertion that extends a large protrusion toward the central cavity (105, 106). Since a major difference between group I and group II chaperonins is the lack of a GroES-like cofactor for members of group II chaperonins, this protrusion is thought to function as a built-in lid (Figure 3*c*). Thus, a functional equivalent to GroES may be integral to the primary sequence of group II apical domains.

Substrate binds in the central cavity of group II chaperonins (Figure 3*c*, *right*) (107) and is folded in an ATP-dependent manner. The mechanism by which group II chaperonins mediate folding is very poorly defined. On the basis of the crystal structure of the thermosome, the complex appears to close into an Anfinsen cage (105), but it is not clear whether the substrate is completely engulfed during productive folding. The eukaryotic chaperonin can assist the folding of proteins, such as actin, that cannot be folded by any other chaperone system (108).

One important difference from GroEL is that several subunits of TRiC lack hydrophobic residues within the regions in the apical domains that correspond to the substrate binding site in GroEL. This difference led to the proposal that the substrate binding site in TRiC is located at the apical protrusions, which contain an obvious hydrophobic surface that faces the central cavity (106). This view is based on the assumption, still untested, that association of the substrate with TRiC is mediated primarily by hydrophobic interactions. The substrate binding site within the individual subunits remains to be defined.

The origin of the subunit heterogeneity in group II chaperonins has not been systematically addressed. One intriguing possibility is that heterogeneity is directly linked to substrate specificity. A number of biochemical studies using endogenous model substrates suggest that each subunit contributes to the recognition of specific motifs within the substrates (107, 109–112). Experiments with truncations and peptide libraries indicate that TRiC interacts with defined regions within actin, tubulin, and the tumor suppressor protein, VHL. For instance, deletion analysis of actin suggested that stable chaperonin binding requires at least

three discrete regions in the polypeptide (109, 112). In addition, a structural analysis of the chaperonin-actin complex using immunoelectron microscopy supports the idea that the polypeptide interacts with specific subunits in the chaperonin (107). Cross-linking experiments that monitor the interaction of nascent actin chains emerging from the ribosome indicate that the elongating polypeptide interacts specifically with individual subunits of the chaperonins (110). In these experiments, the extent to which a nascent chain was cross-linked to multiple TRiC subunits was correlated with the stability of the TRiC–nascent chain complexes to undergo immunoprecipitation, which supports the idea that the frequency and number of different cross-links indeed reflects subunit-specific interactions with different binding sites within the nascent chains. Thus, as in GroEL, stable interactions between a folding polypeptide and the eukaryotic chaperonin may also result from a multivalent set of weak interactions between defined motifs in the substrate and individual chaperonin subunits.

The Hsp90 System

Hsp90 is the central component of a complex chaperone system whose cellular functions and mechanism are still poorly understood (reviewed in 30, 113–116). In eukaryotic cells, Hsp90 is highly abundant (∼2% of cytosolic proteins) and is essential for viability. In contrast, the eubacterial homolog of Hsp90, HtpG, is dispensable. Elucidation of the crystal structure of Hsp90 (117, 118), together with recognition that its activity is regulated by ATP binding and hydrolysis (119, 120), opens the way to a clearer mechanistic understanding of Hsp90 function. Hsp90 works in the context of a complex multiprotein assembly that has been termed the foldosome. This machinery assembles in a stepwise fashion (reviewed in 113, 121).

The first insights into the function of Hsp90 came from higher eukaryotes, where Hsp90 participates in the conformational regulation of signal transduction molecules, such as tyrosine kinases and steroid hormone receptors (reviewed in 30, 113–116). For instance, the oncogenic tyrosine kinase $p60^{v\text{-}src}$ requires Hsp90 to become activated upon myristoylation and attachment to the plasma membrane. Similarly, steroid hormone receptors must associate with Hsp90 to adopt the conformation competent for hormone binding (30, 113–116). Genetic evidence and in vitro reconstitution experiments indicate that Hsp90 function requires its sequential interaction with different subsets of cofactors (122–127), some of which have also been shown to have chaperone activity. For instance, a steroid hormone receptor appears to enter the maturation pathway through a cotranslational or posttranslational interaction with Hsp70, which probably induces conformational changes of the receptor molecule essential for its subsequent recognition by Hsp90. Formation of this early complex is ATP dependent and apparently involves the action of Hdj1 and Hip (124, 125, 128).

Substrate transfer from Hsp70 to Hsp90 is facilitated by the action of Hop (Hsp70-Hsp90 organizing protein) (126, 129). Hop possesses binding sites for both

Hsp70 and Hsp90 and thus promotes the formation of an Hsp70-Hop-Hsp90 complex. In this way, Hop may mediate the efficient transfer of a receptor from Hsp70 to Hsp90.

At the final stage of receptor maturation, an additional set of proteins comprising p23 and different peptidylprolyl isomerases are recruited to the Hsp90-receptor complex (130–134). Tyrosine kinases, such as src, also recruit an additional cofactor, Cdc37 (135, 136). Hsp90-associated peptidylprolyl isomerases, such as FKBP52 and CyP-40 may induce additional conformational changes of the receptor molecule at this stage (130, 131). The interactions of the isomerases, as well as Hip and Hop, with either Hsp90 or Hsp70 are mediated by structurally related tetratricopeptide (TPR) domains that are present in all these cofactors (137–139). Binding to Hsp70, Hsp90, or both is determined by the identity of the TPR domains in a given cofactor. For example, Hip possesses only TPR domains specific for Hsp70 binding, whereas Hop contains a TPR domain that confers binding to Hsp70 and another specific for Hsp90.

p23 associates with Hsp90 only in its ATP-bound state (132–134) and stimulates polypeptide release from Hsp90 upon ATP hydrolysis (134a). Although its role is still unclear, the importance of the Hsp90-p23 interaction for Hsp90 action is underscored by the fact that antitumor drugs of the ansamycin class, such as herbimycin A or geldanamycin, block the p23-Hsp90 interaction and interfere with the Hsp90-mediated maturation of src-like kinases and steroid hormone receptors (118, 132, 140).

COTRANSLATIONAL PROTEIN FOLDING

In the cell, proteins are synthesized on ribosomes. Thus, another important difference between in vitro refolding and folding in vivo arises from the vectorial nature of the translation process. During translation (and also during translocation), the N terminus of the polypeptide is available for folding before the C terminus. In contrast, in vitro refolding experiments examine folding under conditions wherein the entire polypeptide is available at the beginning of the reaction. In principle, the N-terminal portion of a polypeptide could fold as it emerges from the ribosome, or as it is translocated into the ER or the mitochondrion. Although the idea that folding can begin cotranslationally was suggested as early as 1961 (141), its relevance for folding in vivo was disputed on several grounds. First, in vitro refolding experiments can produce native proteins, indicating that cotranslational folding is not always essential. In addition, it is possible to fold circularly-permuted proteins, that is, proteins where the normal N and C termini were covalently linked and new termini introduced by breaking the backbone elsewhere (142). In the resulting protein, the positions of the N and C termini are relocated to different regions of the polypeptide. In several cases, these circular permutations result in folded proteins with a structure or activity similar to that of the wild-type protein (143, 144). Finally, native structures typically are stabilized by multiple highly cooperative

interactions. Thus, folding into a proper tertiary structure can occur only after a complete folding domain (∼50–300 amino acids) has emerged from the ribosome (16, 17). During translation, about 40 residues of a nascent chain are sequestered within the exit channel of the ribosome and are not available for folding; so, it is clear that proteins have to complete their folding posttranslationally.

Unlike in vitro refolding of small single domain proteins, the refolding of larger proteins in vitro is generally inefficient or unsuccessful, particularly in the case of multidomain proteins. It has been observed that, when expressed individually, domains often refold spontaneously and with high efficiency, whereas the same domain within the full-length polypeptide may be unable to refold, ultimately leading to aggregation. For example, the isolated 45-kDa N-terminal kinase domain of aspartokinase-homoserine dehydrogenase refolds spontaneously with good kinetics (145). However, renaturation of the complete 89-kDa protein under the same conditions is considerably slower and more inefficient, and does not yield an active kinase (145). These experiments suggest that folding of an individual domain may be hindered by unfavorable interactions with the rest of the protein.

The difficulties in refolding large proteins raises the possibility that, in vivo, efficient folding of larger, multidomain proteins is promoted by the cotranslational formation of folded structures. The existence of such cotranslational folding has recently been examined using cell-free translation systems derived from eukaryotic and prokaryotic cells.

The 62-kDa protein, firefly luciferase, has been a useful substrate as a model of a large, multidomain protein. In the absence of chaperones, refolding of luciferase upon dilution from denaturant is exceedingly slow and inefficient and leads to aggregation (146–149). In contrast, denatured luciferase is efficiently refolded by several different chaperone systems. For example, chaperone-mediated refolding of chemically denatured luciferase in rabbit reticulocyte lysate occurs with a half time of ∼8 minutes (146–148). Even more striking, when luciferase is translated in reticulocyte or wheat germ lysate, it folds within 1 minute (150, 151). Analysis of the folding intermediates formed during translation revealed that the 22-kDa N-terminal domain of luciferase (residues 1 to 190) folds cotranslationally on ribosome-bound nascent chains (151, 152). Importantly, comparative analysis of folding intermediates indicated that the cotranslational folding pathway of luciferase differs from that of luciferase refolding (152), even though both processes are chaperone-assisted. During translation, the majority of the luciferase chains transit through the intermediate that contains the folded N-terminal domain. In contrast, refolding of chemically denatured full-length luciferase appears to proceed through an ensemble of different conformations, most of which are probably unproductive (152). Although the unproductive intermediate(s) are eventually rescued by the action of molecular chaperones, this reaction is still 8–10 times slower than cotranslational folding. By folding during its biosynthesis, luciferase presumably avoids kinetically trapped intermediates, thereby accounting for the rapid speed of luciferase folding during translation. These results further suggest that the folding pathway of any multidomain proteins could be dictated by establishing

a hierarchy in the folding of its individual domains. Moreover formation of an N-terminal folded structure may provide a scaffold for further folding, thereby preventing the formation of kinetic traps and facilitating rapid folding in vivo.

Cotranslational folding has also been demonstrated for several other proteins (reviewed in 153–155). For instance, ribosome-bound α-globin chains can bind heme after synthesis of only 86 amino acids, possibly reflecting the formation of a folded heme-binding pocket (156). The tailspike protein from phage P22, as well as the subunits of bacterial luciferase (which are completely different from firefly luciferase) also seem to be folded cotranslationally (157, 158). As in the case of firefly luciferase, the folding intermediates observed for newly translated P22 tailspike protein are different from the in vitro refolding intermediates (157). Sequential, cotranslational domain folding was also observed in vivo for the alphavirus capsid protein (159) and for an artificial fusion protein composed of two domains, ras and dihydrofolate reductase (DHFR) (160). These in vivo experiments are valualble because they indicate that cotranslational folding does not result simply from the fact that translation rates in cell-free systems are slower than in the cell. Interestingly, oligomerization can also occur cotranslationally, as indicated by studies of the trimeric reovirus cell attachment protein, sigma 1 (161). When full-length sigma 1 and a C-terminally truncated sigma 1 are translated together in a cell-free system, each newly made protein assembles preferentially into the corresponding homotrimers, particularly at low mRNA concentrations, which favor polysome formation. These results suggest that nascent chains on neighboring ribosomes within a polysome are close enough to interact physically and form higher-order structures.

Interestingly, cotranslational folding has also been demonstrated for proteins that are translocated into the ER (162, 163). This indicates that some of the principles governing cellular folding, such as the cotranslational formation of folded structures, are common to different cellular compartments. Although cotranslational folding has been observed in both eukaryotic and prokaryotic systems, a difference was found in the ability of these systems to fold multidomain proteins (160). The bacterial two-domain protein OmpR and the ras-DHFR fusion protein fold cotranslationally when synthesized in eukaryotic cell-free systems but posttranslationally when expressed in bacterial translation lysates. Because posttranslational folding of the ras-DHFR fusion is inefficient, expression in *E. coli* did not result in an active protein, despite the fact that both ras and DHFR fold efficiently when expressed individually in *E. coli*. These results have led to the proposal that eukaryotic cells have optimized their translation and/or folding machinery to enhance cotranslational folding events. This suggestion falls in line with the observation that eukaryotic proteins tend to be larger than the corresponding prokaryotic proteins (160). Thus, in several instances, the subunits of bacterial oligomeric complexes are fused to form multidomain proteins in their eukaryotic homologs.

Cotranslational domain formation appears important for the efficient folding of multidomain proteins. It is not clear whether cotranslational formation of

secondary and supersecondary structure also plays a role in the folding of individual domains. This question is more difficult to address, given the complexity of even partially purified translation systems. However, the finding that a protein can follow a different folding pathway during translation or during in vitro refolding underscores the importance of characterizing the pathways of protein folding in vivo, the effect of molecular chaperones on the formation of these intermediates, and their relationship to the folding pathways observed in vitro.

CONTRIBUTION OF CYTOSOLIC CHAPERONES TO FOLDING IN VIVO

How many proteins need chaperone assistance to fold in vivo? This question has been actively debated ever since the role of chaperones in mediating cellular folding was realized. Studies of chaperone function using model proteins that fail to fold spontaneously in vitro (such as luciferase) demonstrate that chaperones are indeed required for folding some newly translated or translocated polypeptides. However, these studies do not consider whether the chaperone requirements observed for these model proteins can be generalized to all proteins in the cell. Although the behavior of model proteins provides insight into the general principles of cellular folding, it has been proposed that chaperones are needed to assist the folding of only a few aggregation-prone proteins (164). In this view, the vast majority of proteins reach the native state without chaperone assistance.

A number of recent studies have attempted to define experimentally the contribution of molecular chaperones to overall cellular folding in both prokaryotic and eukaryotic cells (18, 72, 73, 76, 97, 165–167). These studies were designed to estimate the fraction of newly made proteins that fold with the assistance of the different chaperone systems. Addressing this problem experimentally is complicated by the rapid kinetics of translation and folding in living cells. Two major approaches are used to examine this question, one primarily genetic and the other biochemical. Although the problems inherent to each type of study preclude accurate quantitative assessments, their conclusions are remarkably consistent and have greatly clarified our views on the contribution of chaperones to cellular folding. The primary focus has been on the Hsp70 and chaperonin systems, but studies on other chaperone systems will probably follow.

Genetic Approaches These studies have primarily examined chaperone function in prokaryotic cells (72, 76, 166, 167). The rationale behind these experiments is the presumption that proteins that require a given chaperone to fold will aggregate if the cellular levels of this chaperone are reduced. For instance, when GroEL synthesis, driven by a regulatable promoter, is turned off, there is a marked increase in the amount of protein that aggregates and forms inclusion bodies. Identity of the proteins in such aggregates can then be determined by sequencing or mass spectrometry. This approach is quite powerful and allows identification of classes

of proteins that exhibit an absolute requirement for chaperone function. However, it also presents a number of drawbacks. First, it does not establish a direct physical interaction between the putative substrate and the chaperone. Second, the experiments are performed under conditions that lead to cell death, as the loss of chaperone function is ultimately lethal. Finally, proteins that are not themselves substrates of the chaperone but require such substrates in order to fold or assemble correctly (e.g. subunits of an oligomeric complex) could also be found in such aggregates, whereas true substrates could possibly escape aggregation either by being rapidly degraded or by folding with the assistance of an alternate chaperone pathway. These shortcomings can be overcome by confirming the findings of these experiments using a biochemical approach.

Biochemical Approaches The rationale for this approach is that during their folding newly translated polypeptides associate transiently with their cognate chaperones. To determine the population of cellular proteins that associate with a certain chaperone, cells are subjected to a short labeling pulse and then chased. Most of the labeled proteins will complete their synthesis, bind to chaperones, and fold during the chase period. Consequently, the cells are rapidly lysed at different times during the chase and the chaperone-substrate complexes are isolated by immunoprecipitation with antichaperone antibodies (18, 72, 73, 97, 165). This method has a complementary set of drawbacks to the genetic approach. For example, immunoprecipitation typically detects only high-affinity interactions. Thus, weakly bound or rapidly dissociated substrates will be overlooked, and quantitative analysis will underestimate the fraction of cellular proteins that transit through a given chaperone. Unlike the genetic approach, these studies do not establish that folding of a given substrate requires a given chaperone but only that an interaction occurs in vivo.

Contribution of Hsp70 and Other Small Chaperones to de Novo Folding

A role for Hsp70 in de novo folding was originally suggested by several lines of evidence. Because cytoplasmic Hsp70s associate with ribosome-bound nascent chains in eukaryotic cells (33, 35, 151, 168, 169) and mitochondrial and ER Hsp70s bind to translocating polypeptides (56, 170, 171), it was suggested that Hsp70s play a general role in stabilizing translating or translocating polypeptides to prevent their premature misfolding. Genetic and biochemical studies in *S. cerevisiae* also showed that the yeast Hsp70 homologs Ssa1–4 assist the in vivo folding of model proteins (172, 173). However, these experiments do not provide information on the overall contribution of Hsp70 to de novo protein folding in vivo. Given that unfolded polypeptides contain, on average, high-affinity binding sites for Hsp70 every 40 amino acids, it can be assumed that every protein contains at least one Hsp70 binding site. Thus, it is possible in principle that Hsp70 interacts with, and stabilizes, all newly translated polypeptides as they emerge from the ribosome. The

substrate spectrum of Hsp70 was initially examined in mammalian cells, where the flux of newly translated polypeptides through Hsp70 was assessed by quantitative immunoprecipitation (18). These experiments demonstrated that Hsp70 associates transiently with a broad spectrum of polypeptides larger than 20 kDa. Quantitative analysis indicated that at least 15–20% of newly synthesized proteins associate with Hsp70 during their biogenesis. Maximal association with Hsp70 was observed at early chase times, and only a small fraction of labeled polypeptides remained associated after 30 min of chase. The kinetics of dissociation from Hsp70 varied for different substrates, implying that some proteins may require multiple cycles of binding and release. Interestingly, a large fraction of the Hsp70 substrates were proteins greater than 50 kDa in size. Smaller proteins may have a more limited requirement or weaker affinity for Hsp70. Because individual domains in cytosolic proteins are approximately 25–30 kDa, multidomain proteins that fold cotranslationally may be overrepresented among the substrates of this chaperone. Thus, an important function of Hsp70 may be to prevent the interdomain surfaces of folded domains in newly translated proteins from engaging in intermolecular domain-swapping that could lead to aggregation.

Early studies of the major bacterial Hsp70, DnaK, did not support a direct role in de novo folding. *E. coli* strains lacking DnaK (*dnaK*Δ) are viable, albeit heat-sensitive (174). Furthermore, their viability does not arise from a functional overlap with the other Hsp70 homolog, HscA (heat-shock cognate protein A), as *E. coli* lacking both DnaK and HscA are viable (175). These findings suggested that Hsp70s do not play an essential role in prokaryotic folding. However, a direct role for DnaK in chaperoning bacterial nascent chains has now been established (72, 73). Pulse-chase analysis indicated that DnaK interacts transiently with newly synthesized polypeptides over a broad size range, from 14 kDa to well over 90 kDa, binding preferentially to chains ranging from 30–75 kDa. Overall, ∼10% of all soluble polypeptides associate with DnaK at the earliest chase times and dissociate within 2 minutes. The association of DnaK with nascent chains occurs cotranslationally, as demonstrated by experiments that take advantage of the fact that puromycin-released nascent chains become C-terminally tagged with puromycin. At least 20% of DnaK-bound polypeptides could be immunoprecipitated using antipuromycin antibodies (73). This finding argues for a general role of Hsp70 in preventing protein misfolding at the ribosome.

If DnaK does indeed associate with nascent chains, why are cells viable in its absence? Only one other chaperone component, the TF protein, is known to bind nascent chains in *E. coli* (176, 177). The functional significance of this interaction was also unclear, as *E. coli* cells lacking TF (*tig*Δ) are also viable (178). However, the absence of TF results in a two- to threefold increase in the amount of nascent polypeptides that associate with DnaK, which suggests that TF and DnaK cooperate in chaperoning nascent chains (72, 73). In *tig*Δ cells, the proteins associated with DnaK shift to include low-molecular-weight species, which suggests that TF may associate with nascent chains prior to DnaK (73, 76). This apparent functional overlap was confirmed by genetic experiments showing that

*tig*Δ *dnaK*Δ double mutants are inviable (72, 73). Such double-mutant strains contain aggregates derived from both newly synthesized and preexisting proteins. Interestingly, identification of the proteins aggregated in the double mutants indicates that, as observed for mammalian Hsp70, substrates of the DnaK/TF system are predominantly large proteins that probably have more than one folding domain (76).

These studies indicate that DnaK and TF together ensure the productive folding of a substantial fraction of proteins in *E. coli*. An interesting lesson provided by these studies is the existence of functional redundancy between structurally different chaperones. Thus, absence of DnaK in vivo is not compensated by an Hsp70 homolog but rather by TF, an altogether different class of small chaperone that shares with Hsp70 the ability to stabilize nonnative polypeptides. Studies comparing the binding specificity of TF and DnaK indicate that, despite differences in their structure and mechanism, both chaperones bind to very similar motifs comprising short linear sequences enriched in hydrophobic residues (76). This similar substrate selectivity may partly explain how TF and DnaK recognize and promote the folding of the same protein subset in vivo.

Although the role of Hsp70 in de novo folding appears to be conserved in evolution, eukaryotic and prokaryotic cells have some interesting differences. For instance, nascent chains in the eukaryotic system remain bound to Hsp70 for longer times than in bacteria, with a half time for dissociation of ∼10 minutes. In addition, a greater proportion of nascent polypeptides associates with mammalian Hsp70. These findings imply a more prominent role for Hsp70s in eukaryotic protein folding, which may relate to a greater preponderance of large, multidomain proteins in these cells. Although eukaryotic homologs of TF have not been described, it is possible that other small chaperone complexes, such as the nascent chain–associated complex (NAC) (179) and the Gim complex (GimC, for genes involved in microtubules complex; also called prefoldin) may replace or cooperate with Hsp70. For instance, several archaea lack Hsp70 homologs (180) but have a GimC-like complex that might fulfill an Hsp70-like function. In addition, prefoldin/GimC has also been proposed to fulfill a similar function in stabilizing newly translated actin (78, 79). However, another study suggests that GimC acts at a later posttranslational stage (181); therefore, the exact function of the GimC complex remains a subject for future investigation.

Contribution of Chaperonin Complexes to de Novo Folding

The bacterial chaperonin, GroEL, is essential; and loss of GroEL function results in aggregation of ∼30% of cellular proteins (166). Experiments that directly analyzed the flux of newly synthesized proteins through GroEL indicated that it associates transiently with ∼12% of all polypeptides; this estimate increases two- to threefold during heat shock (97). The majority of putative GroEL substrates range in size between 10 and 55 kDa, and are composed of a specific subset of ∼300 polypeptides (97, 165). Given the size constraint estimated for the central cavity of GroEL,

the upper size limit observed for physiological substrates is remarkably consistent with the hypothesis that polypeptide folding occurs within the chaperonin chamber. Overexpression of GroEL increases the fraction of chaperonin-associated polypeptides but does not change the overall size distribution of chaperonin-bound substrates. This observation suggests that the cellular GroEL concentration is limiting, and only a fraction of available substrates can interact with the chaperonin at any given time. Several proteins associate with GroEL throughout their lifetime, which indicates that, in addition to folding, the chaperonin may also play a role in the maintenance of the structural integrity of mature cellular proteins. Examination of over 50 natural GroEL substrates for which structural information is available revealed a significant preference for proteins composed of multiple α/β domains (165). As β-sheets are assembled from discontinuous regions in a polypeptide, binding of their hydrophobic surfaces to GroEL might facilitate the correct packing of strands within a β-sheet as well as the packing of α-helices against neighboring β-sheets.

The role of the chaperonin system of yeast mitochondria in protein folding was also recently examined, using temperature-sensitive alleles of both Hsp60 and the GroES homolog, Hsp10 (167). As observed for GroEL, the loss of Hsp60 results in a pronounced increase in the aggregation of a wide range of mitochondrial proteins. Interestingly, the subsets of proteins that aggregated in *hsp10* and *hsp60* mutants were not identical, which suggests that some polypeptides require only the assistance of Hsp60 for folding.

Despite its similarity to bacterial chaperonins, the substrate spectrum of the eukaryotic cytosolic chaperonin, TRiC/CCT, has been a matter of controversy. It has been suggested that TRiC is a specialized chaperone that folds only a few cytoskeletal proteins (182). This suggestion was based primarily on analysis of TRiC/CCT mutants in *S. cerevisiae*, which exhibit cytoskeletal defects characteristic of defective actin and tubulin function (183). However, direct examination of the cohort of cellular proteins that associate with TRiC/CCT using pulse-chase analysis in mammalian cells demonstrated interaction of 9–15% of newly synthesized proteins with the chaperonin (18). As observed for Hsp70 and GroEL, the kinetics of dissociation from TRiC varied for different proteins, which suggests differences in their requirements for cycles of binding and release. Most TRiC-bound proteins were between 30 and 60 kDa in size, in striking parallel to the size range of GroEL substrates. This size restriction lends further support to the hypothesis that chaperonin-mediated folding occurs within the enclosed central cavity (see above, Ring-Shaped Chaperones: The Chaperonins). Nonetheless, several proteins of 100–120 kDa also transit through the chaperonin, raising the possibility that TRiC may also participate in the folding of individual domains of large proteins. Analysis of TRiC-associated substrates on two-dimensional gels identified at least 70 distinct substrate polypeptides (18). These endogenous substrates remain to be identified. However, studies using model proteins have expanded the list of TRiC substrates to include, in addition to actin and tubulin-related proteins, luciferase (151), G alpha transducin (184), cyclin E (185), the

EBNA1 viral protein (186), myosin (187), and the tumor suppressor protein, VHL (111). Given the limited substrate set, it is difficult to speculate on the structural features that characterize TRiC substrates. Judging by the structures of the known examples, TRiC substrates may have a complex domain organization that results in folding intermediates with an increased tendency to aggregate. Alternatively, they may share a requirement for binding to either a cofactor or an oligomeric partner to complete folding. Given that most of the heterogeneity between TRiC subunits resides in the putative domain for substrate binding (see above, Ring-Shaped Chaperones), it is possible that different subunits in the complex have evolved to recognize different motifs in substrate proteins.

Contribution of Hsp90 to Cellular Folding

As discussed in the section above on Hsp90, many lines of evidence demonstrate that Hsp90 is required for the activity of steroid hormone receptors and some protein kinase families, including src-like and raf-like kinases, as well as the yeast MEKK, Ste11 (188). Association of Hsp90 with these substrates is highly dynamic. In addition to binding Hsp90 after translation, these substrates rebind to the chaperone after activation to return to their inactive conformation. In addition, Hsp90 is also implicated in the maturation and activation of the cystic fibrosis transmembrane regulator (CFTR) (189), nitric oxide synthase (190), telomerase (191), and the hepadnavirus polymerase complex (192).

Considering the abundance of Hsp90 in the eukaryotic cytosol (1–2%), its role is unlikely to be limited to regulation of a restricted number of signal transduction proteins. In vitro, Hsp90 demonstrates chaperone activity toward model proteins. For example, purified Hsp90 stimulates refolding of β-galactosidase by purified Hsp70 and Hsp40 (193). Furthermore, when denatured luciferase is diluted into a mammalian cell extract, it associates with the Hsp90 system (22, 147), which is required for luciferase refolding in the lysate. The possibility that Hsp90 plays a general role in the folding of newly translated proteins was examined in yeast, using a conditional mutant (194). When expressed at the nonpermissive temperature, the folding of two known Hsp90 substrates, v-src (195) and glucocorticoid hormone receptor (196), was impaired. The loss of Hsp90 function, however, did not significantly affect folding of newly synthesized endogenous proteins, nor did it result in aggregation of any distinct subset of cellular proteins (194). Surprisingly, loss of Hsp90 function did not affect de novo folding of firefly luciferase or β-galactosidase, in sharp contrast with the requirement for Hsp90 when these proteins are refolded from denaturant. These results underscore the differences between in vitro and in vivo folding and suggest that Hsp90 does not play a general role in de novo folding.

What then is the role of Hsp90 in the cell? One possible function is assisting the recovery of protein activity following thermal stress. For instance, loss of Hsp90 decreases the kinetics, but not the yield, of refolding of thermally denatured luciferase in vivo. However, the idea that this function is the primary role of Hsp90

is also problematic. First, Hsp90 is essential at all temperatures and not just under stress; second, reduction of Hsp90 levels does not significantly affect the thermotolerance of yeast cells (194). An exciting recent hypothesis proposes that Hsp90 functions as a buffer for protein conformational diversity in the cell. Inactivation of Hsp90 in the fruit fly *Drosophila melanogaster* by conditional mutations or by treatment with geldanamycin uncovers multiple phenotypic variation among the flies (197). The phenotypes arise from preexisting mutations in various regulatory pathways, which were kept silent by the action of Hsp90. These findings led to the suggestion that Hsp90 serves to preserve protein function in the face of genetic variation, presumably by maintaining these proteins in a functional wild-type conformation. This proposal ascribes to the Hsp90 system a pivotal role in evolution: allowing the accumulation of a reservoir of genetic diversity that will permit the emergence of adaptive changes in the face of selective pressure.

In any case, it appears that the restricted subset of proteins that require Hsp90 to fold following translation, such as the steroid receptors and src-related kinases, have coopted a major cellular chaperone pathway for regulatory purposes. Further investigation is needed for a better understanding of the complete range of substrates of Hsp90, as well as of the other cellular functions of this complex system.

ORGANIZATION OF THE CYTOSOLIC CHAPERONE MACHINERY

The dramatic advances in our understanding of mechanistic and structural aspects of individual chaperone systems, together with the realization that chaperones play a significant role in de novo folding, raise a new set of questions concerning how chaperones function in vivo (reviewed in 19, 24, 198). In principle, the cytosol could behave like a bag of chaperones; in this scenario, folding in the cell would be equivalent to diluting a denatured polypeptide into a cell lysate, with the exception that the polypeptide would enter vectorially during translation. Alternatively, there may be important functional and organizational differences between chaperone-mediated folding in vivo and in vitro. Notably, chaperones such as Hsp70 or GroEL display a very broad specificity toward unfolded proteins in vitro, but bind selectively to specific subsets of proteins in the cell. Recent years have witnessed a spirited debate on the extent of functional integration between the various chaperone systems in the cell. Given its novelty and complexity, many aspects of this problem are far from resolved. However, several lines of evidence suggest that chaperone action in the cell is governed by three fundamental principles:

1. Chaperone-mediated folding is processive and functionally coupled to translation or translocation. This allows the chaperone machinery to function as a protective folding compartment that sequesters newly synthesized, nonnative proteins from the bulk cytosol.

2. Molecular chaperones are recruited to bind newly translated (or translocated) polypeptides by specific interactions with the translational (or translocation) machinery.
3. Folding in the cell involves sequential cooperation between different classes of chaperones.

Compartmentalization of Protein Folding in Vivo

A mechanistic hallmark of chaperone-mediated folding is that it occurs via cycles of substrate binding and release, driven by ATP binding and hydrolysis. Multiple cycles are often required for efficient folding, as shown in assays in which chemically denatured rhodanese or luciferase are diluted into a solution containing purified chaperone components. These in vitro studies led to the view that the main function of chaperones is to unfold incorrectly folded or kinetically trapped intermediates to prepare them for another trial of folding in the bulk solution (184, 199, 200). However, whether the interaction of chaperones with newly translated proteins in vivo is governed by the same principles is a matter of debate. Two models have been proposed to describe de novo folding mediated by chaperones in the cell (18, 181, 184, 198, 200–202). According to one model, folding of newly synthesized proteins is a highly coordinated process involving a sequential and processive interaction of different chaperone systems with nonnative folding intermediates (18, 181, 201, 202). An alternative view proposes that chaperones interact with substrate proteins in a stochastic manner and that nonnative folding intermediates partition freely through the cytosol, cycling between a network of available chaperones (184, 198, 200). Because only a small fraction of polypeptides reach the native state in each cycle, a major difference between these models lies in the state of the polypeptide that is released to the bulk cytosol. According to the partitioning model, nonnative folding intermediates are repeatedly discharged in full into the bulk cytosol before reaching the native state (184, 200). Thus, the bulk solution would be the major site of folding for newly translated polypeptides. In contrast, the processive model proposes that the newly translated polypeptide is released into the bulk cytosol only after it adopts a conformation that is committed to fold. An additional aspect of the debate concerns the transfer of a polypeptide between different chaperone systems. Under the partitioning model, a polypeptide released into the bulk cytosol rebinds to either the same or another kind of chaperone. In contrast, under the processive model, folding polypeptides are handed over from one chaperone to another without being released to the solution.

To discriminate between these models, the processivity of de novo folding was examined in both yeast and mammalian cells by introduction of a GroEL mutant (D87K-GroEL) that acts as a trap for nonnative folding intermediates (18, 181). D87K-GroEL binds promiscuously to nonnative proteins but is unable to release them (25, 199). Expression of this trap form of GroEL was used to measure the extent of exposure or release of the nonnative folding polypeptides into the bulk cytosol. Indeed, when expressed in the cytosol of yeast or mammalian

cells, D87K-GroEL was fully capable of binding stress-denatured proteins as well as newly translated polypeptides that were unable to fold (18, 181). However, the D87K-GroEL trap was unable to bind to the folding intermediates generated during protein synthesis, which associated instead with endogenous cytoplasmic chaperones. These experiments support the view that folding in vivo is mediated by a highly organized chaperone machinery that is functionally coupled to translation. They also suggest that the mechanisms that determine the fate of misfolded or stress-denatured proteins do involve cycling of nonnative forms between cellular components and the cytosol, as proposed by the partitioning model.

Interestingly, experiments performed in mammalian cells using an ATPase-defective trap form of BiP, the ER Hsp70, also indicate that newly translocated intermediates are not released into the bulk solution of the ER lumen in vivo (203). Upon translocation into the ER, the heavy chain (HC) of immunoglobulin molecules associates with BiP, and remains in a nonnative conformation until binding to the light chain (LC) to yield the folded immunoglobulin complex. In the absence of LC, the HC remains associated with BiP in a relatively long-lived complex. Expression of the trap form of BiP in these cells did not result in transfer of the nonnative HC from endogenous wild-type BiP to the trap-BiP, which indicates that the unassembled HCs do not cycle on- and off-BiP in vivo but require LCs to trigger their release.

Sequestration of newly translated polypeptides that prevents their binding to the prokaryotic GroEL trap, and presumably prevents their release into the cellular milieu, suggests that the endogenous eukaryotic chaperone machinery functions as a folding compartment that protects the newly made polypeptides from unproductive interactions. The nature of this folding compartment is poorly understood. An important element of the compartmentalization of folding in vivo appears to be the recruitment of chaperones to the translational machinery (see below). Another important factor may be the conditions of macromolecular crowding in the cell, which increase the association constants between substrates and chaperones. Finally, polypeptide transfer between cooperating chaperone systems may also be a highly coordinated and regulated process (see below, Cooperation Between Different Chaperone Systems). Possible transfer mechanisms may include direct interactions between chaperone systems, formation of ternary complexes directed by the bound substrate, or the action of transfer or coupling factors as described for p60/HOP, which coordinates the transfer of substrates between Hsp70 and Hsp90.

Recruitment of Chaperones to the Sites of Translation and Translocation

Elucidation of the crystal structure of the large ribosomal subunit of *Haloarcula marismuorti* at 2.4-Å resolution showed that the polypeptide exit channel averages 15 Å in diameter and is approximately 100 Å long (204, 205). The channel is lined almost entirely by RNA, which provides an unreactive Teflon-like surface that prevents interactions between the ribosome and the nascent polypeptide. Given the

narrow diameter of the exit channel, the polypeptide cannot begin folding while inside the ribosome and emerges unfolded into the cellular milieu where it can associate with chaperones. Whereas association of chaperones with ribosome-bound polypeptides has been demonstrated for both prokaryotic and eukaryotic systems (35, 73, 79, 110, 151, 168, 169, 177, 201, 206, 207), the mechanism directing binding of chaperone proteins to newly synthesized polypeptides is unclear. In principle, the polypeptide emerging from the ribosome can attempt spontaneous folding and chaperones would interact only with those polypeptide chains that have misfolded. However, the experiments indicating that newly translated polypeptides are sequestered from the cellular milieu suggest that the chaperone machinery is specifically recruited to the nascent chain before any misfolding has occurred. Indeed, several chaperone components associate with ribosomes via specific interaction domains.

In prokaryotes, TF and DnaK associate with nascent chains (73, 177), whereas GroEL interacts with substrates posttranslationally. TF has an N-terminal charged domain that mediates its binding to bacterial ribosomes (208). How DnaK associates ribosomes remains unclear. Analysis with chemical cross-linkers suggests that short nascent chains associate first with TF, and only later with DnaK, which indicates that TF factor is positioned closer to the exit site than DnaK (76).

Eukaryotic cells have no TF homolog. However, several cytosolic chaperones were shown to interact with ribosome-bound nascent chains. These factors include NAC (179), the Hsp/Hsc70 system (35, 151, 168, 169), TRiC/CCT (110, 151, 201, 206), GimC/prefoldin complex (79), and Hsp90 (207).

NAC was identified using a cross-linking approach, where a photoactivatable probe is incorporated into nascent chains by in vitro translation in the presence of a modified aminoacyl-tRNA. Cross-links between the NAC heterodimer and nascent chains are lost if the polypeptide is released from the ribosomes or if the probe is distant from the exit site (i.e. in longer nascent chains), which indicates that NAC is in close proximity to the ribosomal exit site (179). Because NAC does not bind unfolded polypeptides and lacks demonstrable chaperone activity in vitro, its role in assisting folding or chaperone binding remains unclear. Deletion mutants in yeast are viable but appear to have a slight defect in mitochondrial import (209).

Binding of Hsp70 family members to ribosome–nascent chain complexes has been demonstrated in mammalian cells and in yeast (33–35, 58, 151, 168, 169, 201, 210). In *S. cerevisiae*, two families of Hsp70 homologs bind to ribosomes: the Ssb proteins (34, 35, 210) and Pdr13 (also named Ssz1) protein (33, 58). The Ssb class of Hsp70 is found only in lower eukaryotes, such as *S. cerevisiae*. The association of Ssb with ribosomes is salt-sensitive in the absence of translation, but becomes resistant to 2 M NaCl in the presence of nascent chains. Although Ssb does not bind to unfolded polypeptides in vitro, it can be cross-linked to ribosome-bound nascent chains of 52 amino acids or longer. Since approximately 40 amino acids at the C terminus of the nascent chain are sequestered inside the exit tunnel of the ribosome, this result suggests that Ssb is located very close to the exit site (210).

The structural basis of the association of Ssb with ribosomes is not known. The binding of Ssb to ribosomes was examined using chimeras of Ssb and Ssa, another Hsp70 that does not bind to ribosomes (34, 210). These experiments indicated that the ribosome-binding site of Ssb is not restricted to one specific element but appears to be distributed throughout the protein.

Proteins containing a J-domain (see above, The Hsp70 System) are thought to facilitate the interaction of Hsp70s with nascent chains (48). For instance, cytosolic J-domain proteins associate with ribosomes as well as with nascent chains (151, 211, 212). The protein, zuotin, is another intriguing link between the chaperone machinery and the ribosome. In addition to containing a J-domain, zuotin contains a charged domain that mediates its ribosome binding (59). It has recently been found that zuotin forms a stable complex with the Hsp70 homolog Prd13 (58). Interestingly, association of Prd13 with ribosomes is dependent on the presence of zuotin, consistent with a role for zuotin as an Hsp70-recruiting factor.

The eukaryotic chaperonin, TRiC/CCT, also interacts cotranslationally with ribosome–nascent chain complexes (110, 151, 201, 206, 213). TRiC was found to associate with ribosome–nascent chain complexes upon size fractionation of cell extracts both in reticulocyte lysate (110, 151, 206) and in P19 embryonic carcinoma cells (213). In this cell line, a significant fraction (5–20%) of the cellular TRiC appeared to be ribosome-associated. Importantly, the TRiC-ribosome interaction was confirmed by coimmunoprecipitation of ribosomes and the chaperonin (213).

TRiC can interact with nascent chains very soon after they emerge from the ribosome. An analysis of the chain length dependence of TRiC binding to actin nascent chains detected photo-cross-links between TRiC and actin nascent chains as short as 133 amino acids (110). This length would leave only ~90–100 amino acids outside the peptide channel, exposed to the cytosol and available for chaperone binding. Similarly, TRiC could be cross-linked to luciferase nascent chains as short as 77 amino acids, i.e. with only ~40 amino acids exposed outside the channel (110). It is therefore conceivable that TRiC is already positioned in close proximity to the nascent chain, perhaps as a result of a specific recruitment mechanism. At present, there is no evidence for a direct physical interaction between TRiC and any component of the ribosome. However, it is possible that TRiC recruitment to newly translated polypeptides is facilitated by other components of the folding machinery, such as the Hsp70 system (151) or the recently described GimC/prefoldin complex (79). Cross-links of ribosome-bound chains to TRiC are observed only for chains longer than those required for cross-linking to the Ssb Hsp70 (210), consistent with the idea that nascent chains interact first with an Hsp70 homolog when they emerge from the ribosome. The finding that incubation with ATP enhanced cross-linking between TRiC and short nascent chains (110) further suggests that TRiC-recruitment may be ATP-mediated, as expected if binding of nascent chains to the chaperonin is promoted by Hsp70. A recent study also proposes that the GimC/prefoldin complex binds first to nascent actin and tubulin chains and delivers them to TRiC (79). However, GimC/prefoldin is not essential for substrate binding to TRiC (112, 181). Thus, actin truncations that

do not bind to GimC can efficiently interact with TRiC after cell-free translation (112). Furthermore, actin binding to TRiC is not affected, and is even slightly enhanced, in an *S. cerevisiae* mutant lacking GimC (181). Since GimC is important for efficient actin folding, these in vivo experiments led to the converse suggestion—that GimC acts on TRiC-bound substrates (181).

The specific recruitment of molecular chaperones to the site where polypeptides first emerge appears to be a general principle of cellular folding. Thus, the study of protein targeting into organelles and into the ER has produced several examples in which chaperone components are physically recruited to the translocon, which is the site of entry of newly translated proteins into mitochondria, chloroplasts, and the ER. These situations are topologically similar to that of a polypeptide entering the cytosol (56, 170, 214–216). The translocons in both mitochondria and the ER contain subunits exposing a J-domain to the luminal side. This J-domain serves as a localization signal that recruits an Hsp70 homolog to bind the incoming polypeptide. For instance, Tim44, a component of the mitochondrial import machinery, binds to Ssc1, the mitochondrial Hsp70 (170, 214), whereas Sec63, a component of the ER translocon, binds to BiP/Kar2 (erHsp70) (56, 216). On the other hand, calnexin, a transmembrane ER protein without cytosolic homologs, has a luminal chaperone domain and a cytoplasmic tail that recruits it to the ribosome (217), thereby promoting the interaction of this chaperone with translocating nascent chains (163, 218). Yet another recruitment mechanism appears to be functional in chloroplasts, where IAP100, a component of the translocation machinery, directly recruits Hsp60 (215).

What is the functional relevance of evolving a mechanism to link chaperones to the translation/translocation machinery? The specific recruitment of chaperones to bind to translating polypeptides may spatially organize the folding machinery. This localization may serve to couple folding and translation, thereby protecting the nascent polypeptide from misfolding, aggregation, and premature degradation. In the case of posttranslational translocation, binding of chaperones to the translocon also serves to drive the import reaction. Preferential chaperone binding to emerging nascent chains may establish a hierarchy in which folding is the preferred fate of a newly translated polypeptide. This may be important as a quality control mechanism: Only after chaperones have attempted to fold the newly synthesized protein will components of the degradation machinery have the opportunity to interact with the polypeptide (further discussed in the section below on degradation).

A better understanding of the architectural interface between the ribosome and the chaperone machinery is needed to determine how chaperones select their substrates. The first step in substrate selection is probably determined by kinetic partitioning between the different chaperones associated with the exit site of the ribosome. In this scenario, the variables that determine binding would be which chaperones are present at the exit site, the affinity of the nascent chain for these chaperones, and the kinetics of formation of folding intermediates that either mask or present chaperone binding sites.

Importantly, chaperone recruitment to the ribosome provides a mechanistic explanation for the observed coupling between translation and folding observed in intact eukaryotic cells, which probably contributes to formation of a protected folding environment for nascent chains (18, 181).

Cooperation Between Different Chaperone Systems

Many newly translated or translocated proteins interact with more than one class of chaperone. Several experiments indicate that these interactions occur in a defined sequence. The first examples were found in experiments that examined folding of model proteins imported into mitochondria or chloroplasts (171, 219, 220). For instance, firefly luciferase imported into mitochondria interacts sequentially with Hsp70 and then Hsp60 as does the mitochondrial matrix protein, Mas2 (171, 220). Subsequent experiments using purified chaperone components (42, 221) proposed a structural basis for this sequential interaction. Given that Hsp70 binds polypeptides in an extended conformation, this protection from aggregation may be sufficient for some proteins for folding, whereas others would require transfer to a chaperonin. Evidence for a transfer from Hsp70 to GroEL has also been obtained in vivo in *E. coli* (73). Analysis of the transit of newly made polypeptides through bacterial chaperones indicated that overexpression of GroEL increases the flux of substrates through DnaK, as expected if GroEL is downstream of DnaK in the folding pathway. Likewise, TF, which functionally replaces DnaK in *dnaK*Δ strains, also appears to cooperate with GroEL in substrate binding (222).

In eukaryotes, TRiC and Hsp70 associate in vivo, which suggests they cooperate functionally (104, 111). Sequential interaction of nascent firefly luciferase and actin with Hsp70 and TRiC was observed after translation in mammalian cell-free lysates (151, 201). In these translation systems, it has been shown both for luciferase and actin that short nascent chains interact with Hsp70, whereas longer ones also interact with TRiC. Chaperone immuno depletion experiments also support the idea of sequential interaction of a substrate with Hsp70 and TRiC (151). As mentioned above, sequential cooperation between the GimC complex and TRiC has also been proposed for actin and tubulin folding (78, 79). Yet another example of sequential cooperation between different chaperone systems is the interaction between Hsp70 and Hsp90 in the folding of steroid hormone receptors and kinases (see above, The Hsp90 System).

Considering all of these examples, it appears that the cell has evolved redundant pathways of polypeptide transfer from small or holding chaperones (i.e. Hsp70, TF, and GimC) to chaperonins (i.e. GroEL, Hsp60, and TRiC). The pathway of chaperone interactions always appears to start with a small chaperone that has a holding function to prevent polypeptide aggregation, such as Hsp70, TF, and possibly prefoldin/GimC. If necessary, the substrate will then bind to a chaperone that recognizes more compact structures and can promote folding, such as the chaperonins or the Hsp90 system. It is not clear how substrate polypeptides are transferred between chaperone systems. The opportunity for sequential interaction

could be determined by several factors: progress along the folding pathway (i.e. formation of compact folding intermediates), on- and off-rates for substrate binding (i.e. from initial binding to a chaperone with a high on-rate to a chaperone with a higher affinity), generation of new chaperone binding sites (i.e. by synthesis), and finally, possible spatial constraints on chaperone binding (i.e. one chaperone is located at the exit site but is unable to complete polypeptide folding).

Cooperation between different chaperones appears to be a central principle of cellular folding. In addition to the examples described above, polypeptides translocating into the ER also interact sequentially with different luminal chaperones (218, 223). However, the mechanistic basis and functional relevance of these observations are the subjects of intense debate (71, 198, 200, 202). Although several lines of evidence suggest that different chaperone systems are organized in pathways, this concept has not been conclusively established. An alternative model has been proposed (200) in which substrates move bidirectionally between different chaperones until folded. In this model, the chaperone machinery has no organization, but rather the substrate partitions among a network of different chaperones. In support of this model, experiments using denatured firefly luciferase as a folding substrate and purified bacterial Hsp70 and GroEL systems demonstrated that the folding polypeptide can indeed be transferred in both directions (224). However, bidirectional transfer between Hsp70 and the chaperonin system was not observed upon translocation of firefly luciferase into isolated mitochondria (220). Mitochondria contain chaperones similar to those of bacteria, and imported luciferase interacts sequentially with mitochondrial (mt)Hsp70 and mtHsp60. It appears that imported luciferase undergoes only one round of interaction with mtHsp70, and cannot return to it once bound to Hsp60, in contrast to the results obtained in vitro (220). Although this experiment supports the idea of sequential interaction of a substrate with chaperones, it is not definitive. It is conceivable that under some circumstances a sequential interaction with Hsp70 (or Hsp70-like proteins) is not essential for a substrate to reach GroEL, TRiC, or Hsp90. Hopefully, future studies will elucidate this fundamental question.

DEGRADATION OF NEWLY TRANSLATED PROTEINS

Given that ribosome-bound polypeptides are unable to fold stably, sequestration of nascent chains in a complex with a specific set of chaperones serves to protect them from aggregation and misfolding. However, cells also possess a machinery to recognize nonnative polypeptides and target them for degradation (reviewed in 225–228). This situation raises the question, When do newly translated chains become accessible to the cellular degradation machinery? If polypeptides were exposed to degradation while bound to the ribosome, the efficiency of folding would be greatly reduced for longer polypeptides or for proteins that have slow folding kinetics in vivo. On the other hand, polypeptides that are unable to fold must eventually be released from the nascent chain–chaperone complex to permit their

degradation. Little is known about the balance between folding and degradation of newly translated proteins. A very small number of studies have addressed this problem in eukaryotic cells. One study examined the behavior of C-terminal truncations of actin (actin-ΔC) lacking a stop codon, which remained ribosome-bound after translation (201). These nascent chains are topologically identical to a nascent chain of bona fide actin, but are unable to fold once released from the ribosome. Upon release from the ribosome by puromycin treatment, these polypeptides remained unfolded, and were rapidly degraded through the ubiquitin-proteasome pathway (225–227). Strikingly, ribosome-associated actin-ΔC was ubiquitinated and degraded much more slowly, indicating a significant degree of protection of the nascent chains. These experiments suggest that association of chaperones with ribosome-bound nascent chains protect the elongating polypeptide from premature ubiquitination and degradation. Interestingly, cotranslational ubiquitination was observed when similar experiments were performed with CFTR, a large polytopic membrane protein that is mutated in cystic fibrosis patients (229). What determines the different behavior of actin and CFTR? Two major differences between these proteins are their size and their domain structure. The 42-kDa actin monomer associates cotranslationally with the chaperonin TRiC and folds into two discontinuous domains upon release from the ribosome. In contrast, the 140-kDa CFTR is a multidomain protein that folds into three independent cytosolic and two membrane-spanning domains. The domains begin to fold cotranslationally and are stabilized by interaction with Hsp70, Hsp90, and Hsp40 (189, 211). Assuming that chaperones in the vicinity of the ribosomal exit site protect polypeptides from proteolysis, the ubiquitin-proteasome system of protein degradation would have access to the newly synthesized polypeptide only after it has been given a chance to interact with chaperones. As the elongating polypeptide grows away from the ribosome, it may leave the protective chaperone-rich environment and become accessible to the ubiquitination machinery. Thus, domain-wise folding may help establish a functional hierarchy of folding and degradation, which may serve as an effective mechanism of quality control. This idea is supported by experiments that examined the extent of cotranslational degradation of model proteins carrying an N-terminal degradation signal (230). A significant degree of cotranslational degradation (almost 50%) was observed in the case of large multidomain proteins, such as the 115-kDa β-galactosidase and the 69-kDa nsP4 polymerase, bearing a destabilizing N-terminal amino acid. In contrast, degradation of the 34-kDa protein Ura3 carrying the same N-terminal degradation signal occurred primarily in a posttranslational manner (230).

Degradation of unproductive newly translated polypeptides by the proteasome may provide a mechanism to generate a spectrum of peptides for presentation by major histocompatibility complex (MHC) class I-molecules that is unbiased by the half-lives of the folded proteins (231). This mechanism would provide the immune system with an overview of the proteins that are currently expressed in any given cell and could explain why peptides from stable proteins, such as actin, are efficiently presented. One important question concerns the fraction of newly

translated proteins that reaches the native state upon translation. A recent set of experiments using cultured mammalian cells led to the surprising conclusion that a very large fraction (approximately 30%) of newly synthesized proteins is degraded shortly after translation (232). These experiments compared the amount of newly translated proteins observed during a pulse-chase experiment in the presence or absence of proteasome inhibitors. Inhibition of the proteasome pathway yielded a substantial increase in the amount of newly translated proteins, which led to the proposal that these proteins would have otherwise been degraded. Considering that the tRNA charging machinery and the translational apparatus have evolved highly sophisticated and stringent proofreading mechanisms (the measured misincorporation frequency is less than 10^{-4}) (232a), it is highly unlikely that such a large fraction of newly translated proteins incorporate mutations that impair folding. The proposal that these degraded polypeptides are defective translating products or errors of the chaperone machinery remains to be investigated (232). Alternatively, the proteins in question could represent surplus subunits of oligomeric proteins, as several examples have been documented in which the level of an oligomeric complex is controlled by the regulated synthesis of one critical subunit and the unregulated expression of the remaining subunits; in such cases, the unassembled subunits are degraded by the ubiquitin-proteasome pathway (233–235).

CONCLUSION: The Complexity of Folding in Vivo and the Regulation of Protein Function

Current View of de Novo Folding

As a result of the recent focus on understanding protein folding as it occurs in the cell, a more coherent picture is beginning to emerge. Despite important differences between prokaryotic and eukaryotic protein folding, the two groups also have striking parallels (Figure 4). In vivo analysis of chaperone interactions revealed that a large fraction of newly translated proteins transit through the major chaperone systems in the cell. Newly translated polypeptides interact first with so-called small or holding chaperones, including Hsp70 and TF (Figure 4*a*). The ability of these chaperones to prevent aggregation probably suffices to promote the folding of a large subset of polypeptides, including multidomain proteins that fold like beads on a string. However, a considerable number of polypeptides also require the protected folding environment provided by the central cavity of prokaryotic and eukaryotic chaperonin complexes (Figure 4*b*). Perhaps these proteins have a more complex, aggregation-prone domain structure that requires extensive interactions between noncontiguous regions. Most chaperonin substrates are medium-size proteins, between 25 and 60 kDa. This observed size distribution suggests that very small proteins do not need the protected environment of the chaperonin cavity to fold. Conversely, proteins too large to fit are presumably composed of smaller individual domains that can fold cotranslationally. The lack

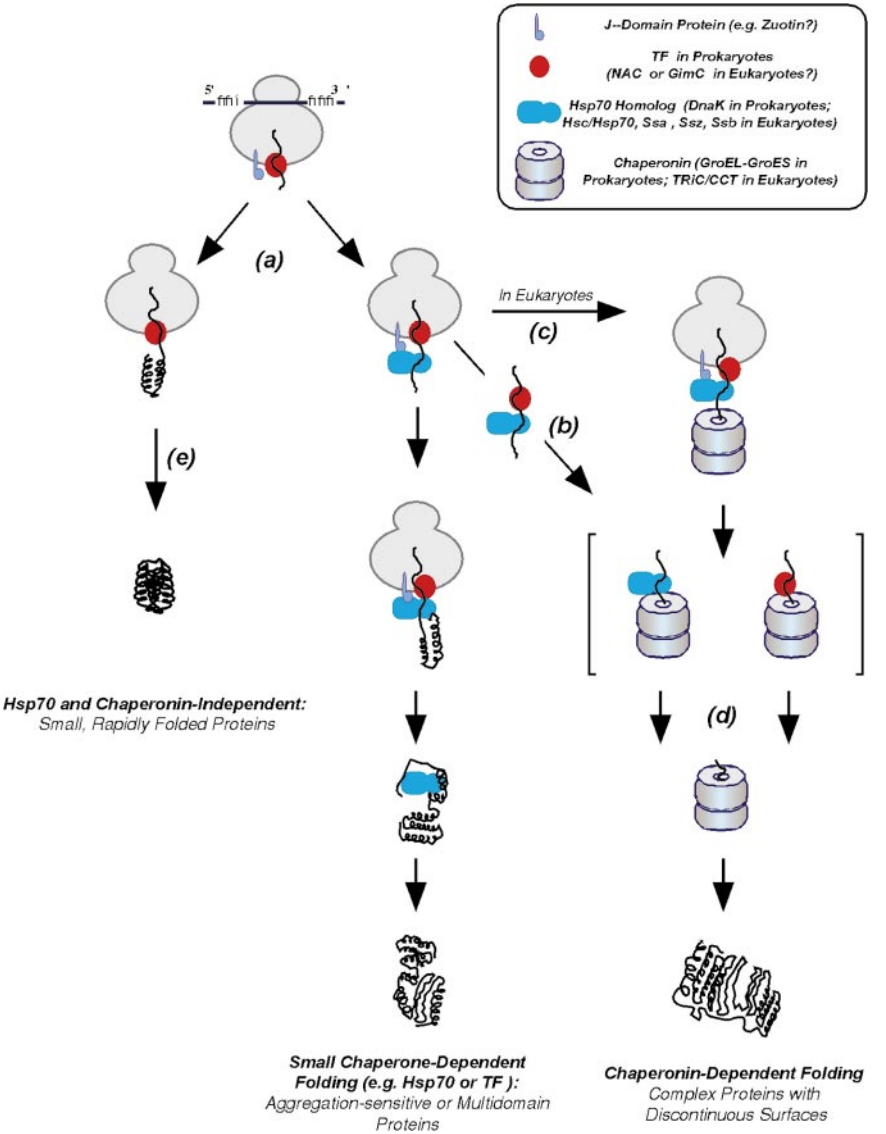


Figure 4 Schematic representation of de novo protein folding in the cytosol of eubacterial and eukaryotic cells. The model emphasizes the evolutionarily conserved characteristics of the folding process. However, some aspects are specific to either eubacterial or eukaryotic cells. For instance, cotranslational domain folding, as well as cotranslational association of the chaperonin complex with nascent chains, is favored in eukaryotes. Conversely, no homolog of TF has been identified in eukaryotes (although several candidates exist). For simplicity, Hsp70 and chaperonin cofactor proteins (e.g. DnaJ, Hdj1/2, or GroES) are not shown; alternative folding pathways involving other chaperones (e.g. Hsp90) are also not represented. See text for details. (Modified from Reference 24 with permission.)

of a GroES-like cofactor in eukaryotes might allow for the cotranslational binding of a single domain rather than the whole protein to the chaperonin (Figure 4*c*); such might be the case for firefly luciferase, whose N-terminal domain folds cotranslationally, and for myosin, whose N-terminal motor domain also associates cotranslationally with the chaperonin. Although sequential interaction of newly synthesized polypeptides with small and large chaperones has been observed in both prokaryotes and eukaryotes (Figure 4*d*), it is possible that some proteins bind directly to the chaperonins.

It appears that a substantial fraction of cellular proteins folds without the assistance of either Hsp70 or the chaperonins (Figure 4*e*). How do these proteins reach the native state? Folding of these proteins might be carried out by novel, as yet uncharacterized chaperone systems. For example, Hsp90 does not appear to play a general role in de novo folding, but is required for folding a restricted class of proteins that includes steroid hormone receptors and Src-like tyrosine kinases (not shown in Figure 4). It is also possible that folding of certain cytosolic proteins occurs in an unassisted manner. Given that intermediates exposing hydrophobic surfaces are not released into the bulk cytosol, except under stress conditions, it is likely that newly translated polypeptides that expose hydrophobic surfaces are targeted to the chaperone machinery. In contrast, small proteins with rapid folding kinetics, as well as proteins consisting of small domains that fold cotranslationally, may not engage in stable or detectable interactions with cytosolic chaperones.

Perspectives and Future Directions

Recent research into how proteins fold in vivo has uncovered a fascinating protein machinery that is highly organized, yet versatile and dynamic. Molecular chaperones have the capacity to fold proteins but also to maintain them in an unfolded state or to target them for degradation. The complexity of protein folding in vivo raises many exciting and important questions.

An important area of future research is to understand how cotranslational folding events influence the folding pathway of proteins. The concept established by pioneering studies of protein folding in vitro, namely, that the three-dimensional conformation of a protein is determined by its amino acid sequence, still holds true. However, folding pathways and their kinetics in vivo appear to be determined in part by the conformation adopted by nascent polypeptides as they emerge from the ribosome. Furthermore, to protect folding intermediates from nonproductive interactions in the crowded cellular environment, the chaperone machinery appears to be coupled to translation. The interface between the ribosome and the chaperone machinery will be an exciting area for future research, particularly in view of the recent elucidation of the ribosome at atomic resolution.

Studies defining the contribution of Hsp70 and the chaperonins to folding in vivo raise the critical question of what determines the discrepancy between the substrate repertoire observed in vivo and in vitro. For instance, both GroEL and Hsp70 interact promiscuously with most unfolded proteins in vitro. In contrast,

only a discrete fraction of the large constellation of cellular polypeptides actually interacts in vivo with either of these chaperones, raising the question of how this specificity is achieved. Identification and characterization of the in vivo substrates of these chaperones will be essential to understand this question. This task is challenging but can perhaps exploit recent advances in proteomics approaches.

Given the ability of chaperones to influence protein conformation, it is tempting to speculate that the expression of protein function might be regulated at the level of folding, as exemplified by the role of Hsp90 in tyrosine kinase and steroid hormone receptor activity. Recent findings indicate that defective folding of key regulatory proteins, as well as the cellular accumulation of incorrectly folded proteins, is the molecular basis of many diseases, including cancer, Alzheimer's disease, prion diseases, and Huntington's disease. This pathology underscores the importance of understanding the in vivo mechanisms of folding (236, 237). Equally as important, it is becoming clear that some proteins do not have a unique native structure in the cell. In the case of prion-like proteins, such as Sup35 in yeast, the newly translated protein adopts a different conformation and function, depending on the conformational state of preexisting Sup35 molecules (238). Tellingly, the outcome of this process is affected by the presence or absence of different molecular chaperones (239). Thus, knowledge of how proteins actually fold in the cell should eventually provide the basis for controlling protein function under normal conditions, and during abnormal conditions of environmental stress and disease.

ACKNOWLEDGMENTS

I thank Drs. Elizabeth Craig, Bernd Bukau, Patricia Clark, Linda Hendershot, and Sabine Rospert for sharing unpublished data. I am also grateful to Drs. John Ellis, Elizabeth Craig, Ulrich Hartl, Doug Cyr, Jörg Höhfeld, Raul Andino-Pavlovsky, Deborah Gordon, and members of the Frydman lab for helpful suggestions and comments on the manuscript. Work in the Frydman lab is supported by an award from the WM Keck Foundation and a grant from the National Institutes of Health (GM56433).

This chapter is dedicated to the memory of Dr. Rosalía B Frydman, my teacher and friend.

LITERATURE CITED

1. Stevens FJ, Argon Y. 1999. *Semin. Cell Dev. Biol.* 10:443–54
2. Trombetta ES, Helenius A. 1998. *Curr. Opin. Struct. Biol.* 8:587–92
3. Parodi AJ. 2000. *Annu. Rev. Biochem.* 69:69–93
4. Anfinsen CB. 1973. *Science* 181:223–30
5. Dill KA, Chan HS. 1997. *Nat. Struct. Biol.* 4:10–19
6. Dobson CM, Sali A, Karplus M. 1998. *Angew. Chem. Int. Ed. Engl.* 37:868–93
7. Jaenicke R. 1998. *Biol. Chem.* 379:237–43
8. Dinner AR, Sali A, Smith LJ, Dobson CM,

Karplus M. 2000. *Trends Biochem. Sci.* 25:331–39
9. Zimmerman SB, Trach SO. 1991. *J. Mol. Biol.* 222:599–620
10. Minton AP. 2000. *Curr. Biol.* 10:R97–99
11. Zimmerman SB, Minton AP. 1993. *Annu. Rev. Biophys. Biomol. Struct.* 22:27–65
12. Martin J, Hartl FU. 1997. *Proc. Natl. Acad. Sci. USA* 94:1107–12
13. van den Berg B, Wain R, Dobson CM, Ellis RJ. 2000. *EMBO J.* 19:3870–75
14. van den Berg B, Ellis RJ, Dobson CM. 1999. *EMBO J.* 18:6927–33
15. Ellis RJ. 1997. *Curr. Biol.* 7:R531–33
16. Jaenicke R. 1991. *Biochemistry* 30:3147–60
17. Creighton TE. 1990. *Biochem. J.* 270:1-16
18. Thulasiraman V, Yang CF, Frydman J. 1999. *EMBO J.* 18:85–95
19. Ellis RJ, Hartl FU. 1999. *Curr. Opin. Struct. Biol.* 9:102–10
20. Seksek O, Biwersi J, Verkman AS. 1997. *J. Cell Biol.* 138:131–42
21. Arrio-Dupont M, Foucault G, Vacher M, Devaux PF, Cribier S. 2000. *Biophys. J.* 78:901–7
22. Schneider C, Sepplorenzino L, Nimmesgern E, Ouerfelli O, Danishefsky S, et al. 1996. *Proc. Natl. Acad. Sci. USA* 93:14536–41
23. Ellis RJ, van der Vies SM. 1991. *Annu. Rev. Biochem.* 60:321–47
23a. Ostermann J, Horwich AL, Neupert W, Hartl FU. 1989. *Nature* 341:125–130
24. Feldman DE, Frydman J. 2000. *Curr. Opin. Struct. Biol.* 10:26–33
25. Bukau B, Horwich AL. 1998. *Cell* 92:351–66
26. Hartl FU. 1996. *Nature* 381:571–79
27. Hartl FU, Martin J. 1995. *Curr. Opin. Struct. Biol.* 5:92–102
28. Gutsche I, Essen LO, Baumeister W. 1999. *J. Mol. Biol.* 293:295–312
29. Fink AL. 1999. *Physiol. Rev.* 79:425–49
30. Caplan AJ. 1999. *Trends Cell Biol.* 9:262–68
31. Mayer MP, Bukau B. 1998. *Biol. Chem.* 379:261–68
32. Craig E, Baxter B, Becker J, Halladay J, Ziegelhoffer T. 1994. In *Heat Shock Proteins: Structure, Function and Regulation*, ed. R Morimoto, A Tissieres, C Georgopoulos, pp. 31–52. Cold Spring Harbor, NY: Cold Spring Harbor Lab.
33. Hallstrom TC, Moye-Rowley WS. 2000. *Mol. Microbiol.* 36:402–13
34. Nelson RJ, Ziegelhoffer T, Nicolet C, Werner-Washburne M, Craig EA. 1992. *Cell* 71:97–105
35. Pfund C, Lopez-Hoyo N, Ziegelhoffer T, Schilke BA, Lopez-Buesa P, et al. 1998. *EMBO J.* 17:3981–89
36. Flaherty KM, DeLuca-Flaherty C, McKay DB. 1990. *Nature* 346:623–28
37. Zhu XT, Zhao X, Burkholder WF, Gragerov A, Ogata CM, et al. 1996. *Science* 272:1606–14
38. McCarty JS, Buchberger A, Reinstein J, Bukau B. 1995. *J. Mol. Biol.* 249:126–37
39. Gamer J, Multhaup G, Tomoyasu T, McCarty JS, Rudiger S, et al. 1996. *EMBO J.* 15:607–17
40. Szabo A, Langer T, Schroder H, Flanagan J, Bukau B, Hartl FU. 1994. *Proc. Natl. Acad. Sci. USA* 91:10345–49
41. Szabo A, Korszun R, Hartl FU, Flanagan J. 1996. *EMBO J.* 15:408–17
42. Langer T, Lu C, Echols H, Flanagan J, Hayer MK, Hartl FU. 1992. *Nature* 356:683–89
43. Suh WC, Lu CZ, Gross CA. 1999. *J. Biol. Chem.* 274:30534–39
44. Harrison CJ, Hayer-Hartl M, DiLiberto M, Hartl FU, Kuriyan J. 1997. *Science* 276:431–35
45. Rudiger S, Germeroth L, Schneider-Mergener J, Bukau B. 1997. *EMBO J.* 16:1501–7
46. Blond-Elguindi S, Cwirla SE, Dower WJ, Lipshutz RJ, Sprang SR, et al. 1993. *Cell* 75:717–28
47. Flynn GC, Pohl J, Flocco MT, Rothman JE. 1991. *Nature* 353:726–30

48. Kelley WL. 1998. *Trends Biochem. Sci.* 23:222–27
49. Pellecchia M, Szyperski T, Wall D, Georgopoulos C, Wuthrich K. 1996. *J. Mol. Biol.* 260:236–50
50. Qian YQ, Patel D, Hartl FU, McColl DJ. 1996. *J. Mol. Biol.* 260:224–35
51. Lu Z, Cyr DM. 1998. *J. Biol. Chem.* 273:27824–30
52. Minami Y, Höhfeld J, Ohtsuka K, Hartl FU. 1996. *J. Biol. Chem.* 271:19617–24
53. Ziegelhoffer T, Lopez-Buesa P, Craig EA. 1995. *J. Biol. Chem.* 270:10412–19
54. Cyr DM, Langer T, Douglas MG. 1994. *Trends Biochem. Sci.* 19:176–81
55. Sha BD, Lee S, Cyr DM. 2000. *Struct. Fold. Des.* 8:799–807
56. Brodsky JL, Schekman R. 1993. *J. Cell Biol.* 123:1355–63
57. Ungewickell E, Ungewickell H, Holstein SE, Lindner R, Prasad K, et al. 1995. *Nature* 378:632–35
58. Gautschi M, Lilie H, Fünfschilling U, Mun A, Ross S, et al. 2001. *PNAS.* In press
59. Yan W, Schilke B, Pfund C, Walter W, Kim SW, Craig EA. 1998. *EMBO J.* 17:4809–17
60. Höhfeld J, Minami Y, Hartl FU. 1995. *Cell* 83:589–98
61. Höhfeld J. 1998. *Biol. Chem.* 379:269–74
62. Takayama S, Bimston DN, Matsuzawa S, Freeman BC, AimeSempe C, et al. 1997. *EMBO J.* 16:4887–96
63. Takayama S, Sato T, Krajewski S, Kochel K, Irie S, et al. 1995. *Cell* 80:279–84
64. Höhfeld J, Jentsch S. 1997. *EMBO J.* 16:6209–16
65. Gross M, Hessefort S. 1996. *J. Biol. Chem.* 271:16833–41
66. Takayama S, Xie ZH, Reed JC. 1999. *J. Biol. Chem.* 274:781–86
67. Stuart JK, Myszka DG, Joss L, Mitchell RS, McDonald SM, et al. 1998. *J. Biol. Chem.* 273:22506–14
68. Terada K, Mori M. 2000. *J. Biol. Chem.* 275:24728–34
69. Luders J, Demand J, Papp O, Höhfeld J. 2000. *J. Biol. Chem.* 275:14817–23
70. Luders J, Demand J, Höhfeld J. 2000. *J. Biol. Chem.* 275:4613–17
71. Ellis RJ. 1999. *Curr. Biol.* 9:R137–39
72. Deuerling E, Schulze-Specking A, Tomoyasu T, Mogk A, Bukau B. 1999. *Nature* 400:693–96
73. Teter SA, Houry WA, Ang D, Tradler T, Rockabrand D, et al. 1999. *Cell* 97:755–65
74. Crooke E, Wickner W. 1987. *Proc. Natl. Acad. Sci. USA* 84:5216–20
75. Scholz C, Mucke M, Rape M, Pecht A, Pahl A, et al. 1998. *J. Mol. Biol.* 277:723-32
76. Deuerling E, Patzelt H, Rudiger S, Schaffitzel E, Mogk A, et al. 2001. *Nature.* Submitted
77. Geissler S, Siegers K, Schiebel E. 1998. *EMBO J.* 17:952–66
78. Vainberg IE, Lewis SA, Rommelaere H, Ampe C, Vandekerckhove J, et al. 1998. *Cell* 93:863–73
79. Hansen WJ, Cowan NJ, Welch WJ. 1999. *J. Cell Biol.* 145:265–77
80. Leroux MR, Fandrich M, Klunker D, Siegers K, Lupas AN, et al. 1999. *EMBO J.* 18:6730–43
80a. Siegert R, Leroux MR, Scheufler C, Hartl FU, Moarefi I. 2000. *Cell* 103:621–32
81. Saibil H. 2000. *Curr. Opin. Struct. Biol.* 10:251–58
82. Gottesman ME, Hendrickson WA. 2000. *Curr. Opin. Microbiol.* 3:197–202
83. Mayhew M, da Silva ACR, Martin J, Erdjument-Bromage H, Tempst P, Hartl FU. 1996. *Nature* 379:420–26
84. Weissman JS, Rye HS, Fenton WA, Beechem JM, Horwich AL. 1996. *Cell* 84:481–90
85. Sigler PB, Xu ZH, Rye HS, Burston SG, Fenton WA, Horwich AL. 1998. *Annu. Rev. Biochem.* 67:581–608
86. Weber F, Keppel F, Georgopoulos C, Hayer-Hartl MK, Hartl FU. 1998. *Nat. Struct. Biol.* 5:977–85
87. Wang JD, Michelitsch MD, Weissman

JS. 1998. *Proc. Natl. Acad. Sci. USA* 95:12163–68
88. Braig K, Otwinowski Z, Hegde R, Boisvert DC, Joachimiak A, et al. 1994. *Nature* 371:578–86
89. Hayer-Hartl MK, Ewbank JJ, Creighton TE, Hartl FU. 1994. *EMBO J.* 13:3192–202
90. Farr GW, Furtak K, Rowland MB, Ranson NA, Saibil HR, et al. 2000. *Cell* 100:561–73
91. Chen LL, Sigler PB. 1999. *Cell* 99:757–68
92. Hunt JF, Weaver AJ, Landry SJ, Gierasch L, Deisenhofer J. 1996. *Nature* 379:37–45
93. Xu ZH, Horwich AL, Sigler PB. 1997. *Nature* 388:741–50
94. Ellis RJ. 1994. *Curr. Biol.* 4:633–35
95. Rye HS, Burston SG, Fenton WA, Beechem JM, Xu ZH, et al. 1997. *Nature* 388:792–98
96. Martin J, Langer T, Boteva R, Schramel A, Horwich AL, Hartl FU. 1991. *Nature* 352:36–42
97. Ewalt KL, Hendrick JP, Houry WA, Hartl FU. 1997. *Cell* 90:491–500
98. Shtilerman M, Lorimer GH, Englander SW. 1999. *Science* 284:822–25
99. Kim S, Willison KR, Horwich AL. 1994. *Trends Biochem. Sci.* 19:543–48
100. Willison KR, Horwich AL. 1996. In *Cell Biology Series: The Chaperonins*, ed. RJ Ellis, pp. 107–36. London: Academic
101. Leroux MR, Hartl FU. 2000. *Curr. Biol.* 10:R260–64
102. Gao Y, Thomas JO, Chow RL, Lee GH, Cowan NJ. 1992. *Cell* 69:1043–50
103. Frydman J, Nimmesgern E, Erdjument-Bromage H, Wall JS, Tempst P, Hartl FU. 1992. *EMBO J.* 11:4767–78
104. Lewis VA, Hynes GM, Dong Z, Saibil H, Willison K. 1992. *Nature* 358:249–52
105. Ditzel L, Lowe J, Stock D, Stetter KO, Huber H, et al. 1998. *Cell* 93:125–38
106. Klumpp M, Baumeister W, Essen LO. 1997. *Cell* 91:263–70
107. Llorca O, McCormack EA, Hynes G, Grantham J, Cordell J, et al. 1999. *Nature* 402:693–96
108. Tian GL, Vainberg IE, Tap WD, Lewis SA, Cowan NJ. 1995. *Nature* 375:250–53
109. Hynes GM, Willison KR. 2000. *J. Biol. Chem.* 275:18985–94
110. McCallum CD, Do H, Johnson AE, Frydman J. 2000. *J. Cell Biol.* 149:591–601
111. Feldman DE, Thulasiraman V, Ferreyra RG, Frydman J. 1999. *Mol. Cell* 4:1051–61
112. Rommelaere H, De Neve M, Melki R, Vandekerckhove J, Ampe C. 1999. *Biochemistry* 38:3246–57
113. Pratt WB. 1998. *Proc. Soc. Exp. Biol. Med.* 217:420–34
114. Buchner J. 1999. *Trends Biochem. Sci.* 24:136–41
115. Smith DF. 2000. *Semin. Cell Dev. Biol.* 11:45–52
116. Pearl LH, Prodromou C. 2000. *Curr. Opin. Struct. Biol.* 10:46–51
117. Prodromou C, Roe SM, O'Brien R, Ladbury JE, Piper PW, Pearl LH. 1997. *Cell* 90:65–75
118. Stebbins CE, Russo AA, Schneider C, Rosen N, Hartl FU, Pavletich NP. 1997. *Cell* 89:239–50
119. Panaretou B, Prodromou C, Roe SM, O'Brien R, Ladbury JE, et al. 1998. *EMBO J.* 17:4829–36
120. Obermann WMJ, Sondermann H, Russo AA, Pavletich NP, Hartl FU. 1998. *J. Cell Biol.* 143:901–10
121. Frydman J, Höhfeld J. 1997. *Trends Biochem. Sci.* 22:87–92
122. Morishima Y, Murphy PJM, Li DP, Sanchez ER, Pratt WB. 2000. *J. Biol. Chem.* 275:18054–60
123. Dittmar KD, Pratt WB. 1997. *J. Biol. Chem.* 272:13047–54
124. Dey B, Caplan AJ, Boschelli F. 1996. *Mol. Biol. Cell* 7:91–100
125. Kimura Y, Yahara I, Lindquist S. 1995. *Science* 268:1362–65
126. Johnson BD, Schumacher RJ, Ross ED,

Toft DO. 1998. *J. Biol. Chem.* 273:3679-86
127. Johnson JL, Craig EA. 2000. *Mol. Cell. Biol.* 20:3027–36
128. Prapapanich V, Chen SY, Smith DF. 1998. *Mol. Cell. Biol.* 18:944–52
129. Chen SY, Smith DF. 1998. *J. Biol. Chem.* 273:35194–200
130. Prapapanich V, Chen S, Nair S, Rimerman RA, Smith DF. 1996. *Mol. Cell. Biol.* 16:6200–7
131. Ratajczak T, Carrello A. 1996. *J. Biol. Chem.* 271:2961–65
132. Bohen SP. 1998. *Mol. Cell. Biol.* 18:3330–39
133. Grenert JP, Johnson BD, Toft DO. 1999. *J. Biol. Chem.* 274:17525–33
134. Fang YF, Fliss AE, Rao J, Caplan AJ. 1998. *Mol. Cell. Biol.* 18:3727–34
134a. Young JC, Hartl FU. 2000. *EMBO J.* 19:5930–40
135. Kimura Y, Rutherford SL, Miyata Y, Yahara I, Freeman BC, et al. 1997. *Genes Dev.* 11:1775–85
136. Stepanova L, Leng X, Parker SB, Harper JW. 1996. *Genes Dev.* 10:1491–502
137. Liu FH, Wu SJ, Hu SM, Hsiao CD, Wang C. 1999. *J. Biol. Chem.* 274:34425–32
138. Young JC, Obermann WMJ, Hartal FU. 1998. *J. Biol. Chem.* 273:18007–10
139. Scheufler C, Brinker A, Bourenkov G, Pegoraro S, Moroder L, et al. 2000. *Cell* 101:199–210
140. Whitesell L, Mimnaugh EG, De Costa B, Myers CE, Neckers LM. 1994. *Proc. Natl. Acad. Sci. USA* 91:8324–28
141. Chantrenne H. 1961. In *Modern Trends in Physiological Science*, ed. P Alexander, Z Bacq, pp. 122. Oxford, UK: Pergamon
142. Pan T, Uhlenbeck OC. 1993. *Gene* 125:111–14
143. Iwakura M, Nakamura T, Yamane C, Maki K. 2000. *Nat. Struct. Biol.* 7:580–85
144. Graf R, Schachman HK. 1996. *Proc. Natl. Acad. Sci. USA* 93:11591–96
145. Garel JR. 1992. See Ref. 240, pp. 405–54
146. Nimmesgern E, Hartl FU. 1993. *FEBS Lett.* 331:25–30
147. Schumacher RJ, Hurst R, Sullivan WP, McMahon NJ, Toft DO, Matts RL. 1994. *J. Biol. Chem.* 269:9493–99
148. Freeman BC, Myers MP, Schumacher R, Morimoto RI. 1995. *EMBO J.* 14:2281–92
149. Herbst R, Schafer U, Seckler R. 1997. *J. Biol. Chem.* 272:7099–105
150. Kolb VA, Makeyev EV, Spirin AS. 1994. *EMBO J.* 13:3631–37
151. Frydman J, Nimmesgern E , Ohtsuka K, Hartl FU. 1994. *Nature* 370:111–17
152. Frydman J, Erdjument-Bromage H, Tempst P, Hartl FU. 1999. *Nat. Struct. Biol.* 6:697–705
153. Hardesty B, Tsalkova T, Kramer G. 1999. *Curr. Opin. Struct. Biol.* 9:111–14
154. Fedorov AN, Baldwin TO. 1997. *J. Biol. Chem.* 272:32715–18
155. Netzer WJ, Hartl FU. 1998. *Trends Biochem. Sci.* 23:68–73
156. Komar AA, Kommer A, Krasheninnikov IA, Spirin AS. 1997. *J. Biol. Chem.* 272:10646–51
157. Clark PL, King J. 2001. *J. Biol. Chem.* Submitted
158. Fedorov AN, Baldwin TO. 1995. *Proc. Natl. Acad. Sci. USA* 92:1227–31
159. Nicola AV, Chen W, Helenius A. 1999. *Nat. Cell Biol.* 1:341–45
160. Netzer W, Hartl F. 1997. *Nature* 388:343–49
161. Gilmore R, Coffey MC, Leone G, McLure K, Lee PWK. 1996. *EMBO J.* 15:2651–58
162. Braakman I, Helenius J, Helenius A. 1992. *Nature* 356:260–62
163. Chen W, Helenius J, Braakman I, Helenius A. 1995. *Proc. Natl. Acad. Sci. USA* 92:6229–33
164. Lorimer GH. 1996. *FASEB J.* 10:5–9

165. Houry WA, Frishman D, Eckerskorn C, Lottspeich F, Hartl FU. 1999. *Nature* 402:147–54
166. Horwich AL, Low KB, Fenton WA, Hirshfield IN, Furtak K. 1993. *Cell* 74: 909–17
167. Dubaquie Y, Looser R, Funfschilling U, Jeno P, Rospert S. 1998. *EMBO J.* 17:5868–76
168. Hansen WJ, Lingappa VR, Welch WJ. 1994. *J. Biol. Chem.* 269:26610–13
169. Beckmann RP, Mizzen LA, Welch WJ. 1990. *Science* 248:850–54
170. Voos W, von Ahsen O, Muller H, Guiard B, Rassow J, Pfanner N. 1996. *EMBO J.* 15:2668–77
171. Manning KU, Scherer PE, Schatz G. 1991. *EMBO J.* 10:3273–80
172. Crombie T, Boyle JP, Coggins JR, Brown AJP. 1994. *Eur. J. Biochem.* 226:657–64
173. Kim S, Schilke B, Craig EA, Horwich AL. 1998. *Proc. Natl. Acad. Sci. USA* 95:12860–65
174. Paek KH, Walker GC. 1987. *J. Bacteriol.* 169:283–90
175. Hesterkamp T, Bukau B. 1998. *EMBO J.* 17:4818–28
176. Stoller G, Rucknagel KP, Nierhaus KH, Schmid FX, Fischer G, Rahfeld JU. 1995. *EMBO J.* 14:4939–48
177. Hesterkamp T, Hauser S, Lutcke H, Bukau B. 1996. *Proc. Natl. Acad. Sci. USA* 93:4437–41
178. Guthrie B, Wickner W. 1990. *J. Bacteriol.* 172:5555–62
179. Wang S, Sakai H, Wiedmann M. 1995. *J. Cell Biol.* 130:519–28
180. Gribaldo S, Lumia V, Creti R, deMacario EC, Sanangelantoni A, Cammarano P. 1999. *J. Bacteriol.* 181:434–43
181. Siegers K, Waldmann T, Leroux MR, Grein K, Shevchenko A, et al. *EMBO J.* 18:75–84
182. Lewis SA, Tian G, Vainberg IE, Cowan NJ. 1996. *J. Cell Biol.* 132:1–4
183. Stoldt V, Rademacher F, Kehren V, Ernst JF, Pearce DA, Sherman F. 1996. *Yeast* 12:523–29
184. Farr GW, Scharl EC, Schumacher RJ, Sondek S, Horwich AL. 1997. *Cell* 89:927–37
185. Won KA, Schumacher RJ, Farr GW, Horwich AL, Reed SI. 1998. *Mol. Cell. Biol.* 18:7584–89
186. Kashuba E, Pokrovskaja K, Klein G, Szekely L. 1999. *J. Hum. Virol.* 2:33–37
187. Srikakulam R, Winkelmann DA. 1999. *J. Biol. Chem.* 274:27265–73
188. Louvion JF, Abbas-Terki T, Picard D. 1998. *Mol. Biol. Cell* 9:3071–83
189. Loo MA, Jensen TJ, Cui LY, Hou YX, Chang XB, Riordan JR. 1998. *EMBO J.* 17:6879–87
190. Garcia-Cardena G, Fan R, Shah V, Sorrentino R, Cirino G, et al. 1998. *Nature* 392:821–24
191. Holt SE, Aisner DL, Baur J, Tesmer VM, Dy M, et al. 1999. *Genes Dev.* 13:817–26
192. Hu JM, Toft DO, Seeger C. 1997. *EMBO J.* 16:59–68
193. Freeman BC, Morimoto RI. 1996. *EMBO J.* 15:2969–79
194. Nathan DF, Vos MH, Lindquist S. 1997. *Proc. Natl. Acad. Sci. USA* 94:12949–56
195. Xu Y, Lindquist S. 1993. *Proc. Natl. Acad. Sci. USA* 90:7074–78
196. Picard D, Khursheed B, Garabedian B, Fortin MG, Lindquist S, Yamamoto KR. 1990. *Nature* 348:166–68
197. Rutherford SL, Lindquist S. 1998. *Nature* 396:336–42
198. Bukau B, Deuerling E, Pfund C, Craig EA. 2000. *Cell* 101:119–22
199. Weissman JS, Kashi Y, Fenton WA, Horwich AL. 1994. *Cell* 78:693–702
200. Bukau B, Hesterkamp T, Luirink J. 1996. *Trends Cell Biol.* 6:480–86
201. Frydman J, Hartl FU. 1996. *Science* 272:1497–502
202. Ellis RJ, Hartl FU. 1996. *FASEB J.* 10:20–26
203. Vanhove M, Usherwood Y-K, Hendershot LM. 2001. *Immunity.* Submitted

204. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. *Science* 289:920–30
205. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. *Science* 289:905–20
206. Dobrzynski J, Sternlicht M, Farr G, Sternlicht H. 1996. *Biochemistry* 35:15870–82
207. Uma S, Hartson SD, Chen JJ, Matts RL. 1997. *J. Biol. Chem.* 272:11648–56
208. Hesterkamp T, Deuerling E, Bukau B. 1997. *J. Biol. Chem.* 272:21865–71
209. George R, Beddoe T, Landl K, Lithgow T. 1998. *Proc. Natl. Acad. Sci. USA* 95:2296–301
210. Pfund C, Huang P, Lopez-Hoyo N, Craig EA. 2001. Submitted
211. Meacham GC, Lu Z, King S, Sorscher E, Tousson A, Cyr DM. 1999. *EMBO J.* 18:1492–505
212. Zhong T, Arndt KT. 1993. *Cell* 73:1175–86
213. Roobol A, Carden MJ. 1999. *Eur. J. Cell Biol.* 78:21–32
214. Horst M, Oppliger W, Rospert S, Schonfeld HJ, Schatz G, Azem A. 1997. *EMBO J.* 16:1842–49
215. Kessler F, Blobel G. 1996. *Proc. Natl. Acad. Sci. USA* 93:7684–89
216. Scidmore MA, Okamura HH, Rose MD. 1993. *Mol. Biol. Cell* 4:1145–59
217. Chevet E, Wong HN, Gerber D, Cochet C, Fazel A, et al. 1999. *EMBO J.* 18:3655–66
218. Hammond C, Helenius A. 1994. *Science* 266:456–58
219. Tsugeki R, Nishimura M. 1993. *FEBS Lett.* 320:198–202
220. Heyrovska N, Frydman J, Höhfeld J, Hartl FU. 1998. *Biol. Chem.* 379:301–9
221. Petit MA, Bedale W, Osipiuk J, Lu C, Rajagopalan M, et al. 1994. *J. Biol. Chem.* 269:23824–29
222. Kandror O, Sherman M, Moerschell R, Goldberg AL. 1997. *J. Biol. Chem.* 272:1730–34
223. Melnick J, Dul JL, Argon Y. 1994. *Nature* 370:373–75
224. Buchberger A, Schroder H, Hesterkamp T, Schonfeld HJ, Bukau B. 1996. *J. Mol. Biol.* 261:328–33
225. Hershko A, Ciechanover A. 1998. *Annu. Rev. Biochem.* 67:425–79
226. Voges D, Zwickl P, Baumeister W. 1999. *Annu. Rev. Biochem.* 68:1015–68
227. Ciechanover A, Orian A, Schwartz AL. 2000. *BioEssays* 22:442–51
228. Wickner S, Maurizi MR, Gottesman S. 1999. *Science* 286:1888–93
229. Sato S, Ward CL, Kopito RR. 1998. *J. Biol. Chem.* 273:7189–92
230. Turner GC, Varshavsky A. 2000. *Science* 289:2117–20
231. Reits EAJ, Vos JC, Gromme M, Neefjes J. 2000. *Nature* 404:774–78
232. Schubert U, Anton LC, Gibbs J, Norbury CC, Yewdell JW, Bennink JR. 2000. *Nature* 404:770–74
232a. Kurland CG. 1992. *Annu. Rev. Genet.* 26:29–50
233. Moritz M, Paulovich AG, Tsay YF, Woolford JL. 1990. *J. Cell Biol.* 111:2261–74
234. Keller SH, Lindstrom J, Taylor P. 1996. *J. Biol. Chem.* 271:22871–77
235. Johnson PR, Swanson R, Rakhilina L, Hochstrasser M. 1998. *Cell* 94:217–27
236. Dobson CM. 1999. *Trends Biochem. Sci.* 24:329–32
237. Taubes G. 1996. *Science* 271:1493–95
238. Serio TR, Lindquist SL. 1999. *Annu. Rev. Cell Dev. Biol.* 15:661–703
239. Chernoff YO, Newnam GP, Kumar J, Allen K, Zink AD. 1999. *Mol. Cell. Biol.* 19:8103–12
240. TE Creighton, ed. 1992. *Protein Folding.* New York: Freeman

Annual Review of Biochemistry
Volume 70, 2001

CONTENTS