Plant Physiology, November 1999, Vol. 121, pp. 695–703, www.plantphysiol.org © 1999 American Society of Plant Physiologists

Update on Biochemistry

Protein Folding in the Plant Cell

Jan A. Miernyk*

Mycotoxin Research Unit, United States Department of Agriculture, Agricultural Research Service, National Center for Agricultural Utilization Research, Peoria, Illinois 61604–3902

Whole-genome sequencing projects have drastically changed the landscape of biological research. The lexicon of contemporary biology contains a plethora of new terms, including: genomics, research pertaining to the genome; proteomics, description of the protein complement of an organism; and bioinformatics, the collection and interpretation of biological information (primarily nucleic acid and amino acid sequence data) (Bouchez and Höfte, 1998).

However, obtaining sequence information is not an end unto itself. It is essential that the products of these genes be identified and their function and physiological significance discovered (Bork et al., 1998; Saier, 1998). In complex eukaryotes such as flowering plants, as many as 5×10^4 genes can be selectively expressed in individual cells. It is the products of these genes, the proteins, that determine the fate and function of the cells.

Protein function is determined by how the protein folds to form a specific three-dimensional structure. The way that a protein folds is determined by the free energy of the constituent amino acid residues (Levitt et al., 1997). As much as 50% of the primary amino acid sequence is necessary just to define the three-dimensional structure of a typical protein (Dobson et al., 1998).

PROTEIN FOLDING IN THE CELL IS NOT AUTONOMOUS

The classic in vitro studies of Anfinsen, which resulted in his receipt of the 1972 Nobel prize in chemistry, demonstrated that the primary amino acid sequence of a protein can contain all of the information necessary to direct the folding of a polypeptide chain to the correct final structure (for review, see Anfinsen, 1973). However, the conditions of temperature and pH, the salt concentration, and especially the total protein concentration found in vivo tend to promote a plethora of side reactions that compete with the single pathway that will lead to the correct final structure.

Unfolded and partially folded proteins tend to aggregate when present at the concentrations found in vivo, which are estimated to be as high as 340 mg mL^{-1} in *Escherichia coli*. Molecular chaperones, proteins that prevent inappropriate association or aggregation of exposed hydrophobic surfaces of unfolded or partially folded proteins and direct them into productive folding, transport, or degradation pathways, function to minimize protein aggregation and can promote dissociation of aggregates that have formed (Boston et al., 1996; Miernyk, 1997; Netzer and Hartl, 1998; Sigler et al., 1998). Protein-folding catalysts, conventional enzymes that accelerate the rate-limiting steps in protein folding, allowing folding intermediates to avoid aggregation and non-productive interactions with other proteins, also assist cellular proteins to avoid aggregation by accelerating the rate of correct folding (Schmid, 1993; Boston et al., 1996; Huppa and Ploegh, 1998). The molten globule state is an intermediate stage where a protein is "partly unfolded," and it is thought that proteins are in the molten globule state when recognized by chaperones or for membrane translocation.

MOLECULAR CHAPERONES PREVENT NON-PRODUCTIVE INTERACTIONS

Molecular chaperones assist in the assembly/disassembly of proteins but are not themselves components of the final structures. Molecular chaperones do not have an active role in protein folding, do not accelerate the folding reactions, nor do they provide steric information directing protein folding. Rather, they serve to reduce the divergence of folding intermediates into non-productive side reactions. Molecular chaperones reversibly bind to and shield unfolded segments of polypeptides that would otherwise act as loci for aggregation. For in vitro analysis, it is assumed the protein folding and unfolding are equivalent. Molecular chaperones are often characterized by their ability to prevent aggregation of proteins unfolded by mild acid or heat treatment (Fig. 1). Under stress conditions such as heat shock, it is the synthesis of molecular chaperones that allows cellular proteins to avoid and/or recover from aggregation. A brief description of each of the major classes of stress proteins/molecular chaperones from the large to small subunit size is presented.

THERE ARE SIX MAJOR FAMILIES OF CHAPERONES, EACH WITH UNIQUE HOMOLOGS IN DIFFERENT SUBCELLULAR COMPARTMENTS

Stress100/Clp

The 100-kD stress protein is found in all organisms, with the actual size ranging from 84 to 104 kD. There are two major subclasses: class 1 proteins (A, B, C, and D) have two ATP-binding sites, and class 2 proteins (M, N, X, and Y) have a single ATP-binding site (Schirmer et al., 1996). Class

* E-mail miernykj@mail.ncaur.usda.gov; fax 309–681–6686.

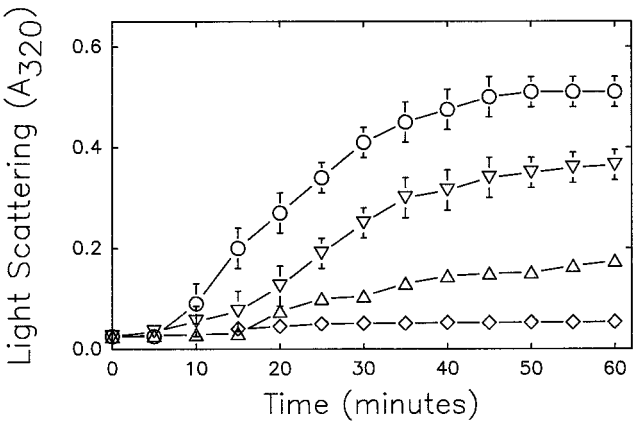

Figure 1. The Stress70 chaperone machine prevents thermal aggregation of the model protein malate dehydrogenase. Malate dehydrogenase (300 nM) was incubated at 45°C with no chaperone (○), with recombinant Arabidopsis Stress70 (▽), Stress70 plus the chaperone activating protein AtJ2 (△), or Stress70 plus AtJ2 plus the nucleotide exchange factor AtE1 (◇). Light scattering, measured at 320 nm, increased as the enzyme was unfolded during heat treatment. Data are the means ± SE of three measurements.

1 Stress100/Clp proteins have a coiled-coil secondary structure separating the ATP-binding domains (Nieto-Sotelo et al., 1999). Transmission electron microscopy of Stress100 reveals ring-shaped particles with a 6-fold rotational symmetry. Side views show two rings stacked together (a dodecamer).

The HSP100 proteins have been extensively studied in *E. coli* as subunits of an ATP-dependent protease (caseinolytic protease [Clp]) (Hoskins et al., 1998). Clp consists of two distinct subunits: ClpP is the actual protease, while the ClpA/B/C/X chaperone subunits designate target specificity. The ClpP sequence is unrelated to ClpA/B/C/X, and ClpP is not a chaperone. Either as subunits of the homomeric ring structure or as subunits of the protease, HSP100/Clp employs ATP hydrolysis to promote changes in protein folding and assembly. Thus, the HSP100/Clp proteins constitute a class of molecular chaperones (Boston et al., 1996; Hoskins et al., 1998).

In plant cells HSP100/Clp is both cytoplasmic and organellar (within plastids and/or mitochondria). HSP100/Clp is expressed in developmental and organ-specific patterns and is up-regulated by a variety of environmental stress conditions (heat, cold, high salt, and heavy metals) (Schirmer et al., 1994).

Stress90

The 90-kD stress proteins, which actually range in size from 82 to 96 kD, are highly conserved and abundant in the cytoplasm of both eukaryotic and prokaryotic cells (Csermely et al., 1998). Despite the high levels of expression under non-stress conditions, the term HSP90 is widely employed. Although considered controversial for many years, the question of ATP binding to Stress90 was answered unequivocally when the x-ray structure was solved (Fig. 2A). Thus, HSP90 is an ATP-dependent molecular chaperone that binds transiently to late, probably highly structured folding intermediates, preventing aggregation. While the majority of cellular HSP90 is in the cytoplasm, there are distinct organellar forms found in the rough endoplasmic reticulum (ER), plastids, and mitochondria (Boston et al., 1996; Møgelsvang and Simpson, 1998).

HSP90 can function independently as a chaperone; however, it also acts in concert with a group of other proteins that together comprise the foldosome, or cytoplasmic chaperone heterocomplex (CCH). The CCH has been studied most extensively in mammalian cells, where it plays an important role in signal transduction via interaction with steroid hormone receptors and protein kinases including the Src and Raf components of the MAP kinase system (Buchner, 1999). HSP90 is associated with at least six partner proteins complexed with the hormone-free receptor. Formation of this complex is essential for subsequent hormone binding. After binding of the ligand and dissociation of the CCH, the activated hormone-bound receptor functions as a transcription factor. In the absence of ligands, the receptors interact with HSP70 to start a new cycle. Two important factors have led to the characterization of CCH function: (a) the CCH can be assembled in vitro from isolated components, and (b) this assembly can be prevented by geldanamycin, a specific inhibitor of Stress90 function (Buchner, 1999).

HSP90 partner proteins in the CCH include: Stress70, Hip (HSP70 interacting protein; p48), Hop (HSP70/HSP90 organizing protein; p60 or Sti1p), a DnaJ homolog, the folding catalyst prolyl-isomerase (PPI), and p23 (Sba1p in yeast). An additional component, p50 (Cdc37), has only been detected in complexes of the CCH with protein kinases.

Although the CCH has been best characterized in mammalian systems, a complex that is very similar in both composition and function can also be found in both yeast and filamentous fungi (e.g. Brunt et al., 1998). Plant cells also contain a CCH capable of activating the mammalian glucocorticoid receptor in vitro. HSP70, HSP90, and an FKBP-type prolyl-isomerase have been identified as components of the wheat CCH (Reddy et al., 1998). The maize CCH also contains a DnaJ homolog and a cyclophilin-type prolyl-isomerase (J.A. Miernyk, unpublished data). Plant and yeast homologs of the mammalian steroid hormone receptors have not yet been reported, and the native targets for CCH function in these cells are at this time unknown.

Stress70

The 70-kD stress proteins comprise a ubiquitous set of highly conserved molecular chaperones that range in actual size from 68 to 110 kD (Vierling, 1991; Boston et al., 1996; Miernyk, 1997). Some family members are constitutively expressed and are often referred to as HSC70 (70-kD heat shock cognate). Other family members are expressed only when the organism is challenged by environmental stresses such as temperature extremes, anoxia, heavy metals, and predation. These family members are generally referred to as HSP70 (70-kD heat shock protein), even, for example, when they are induced by cold shock. No differences in actual chaperone function have been described

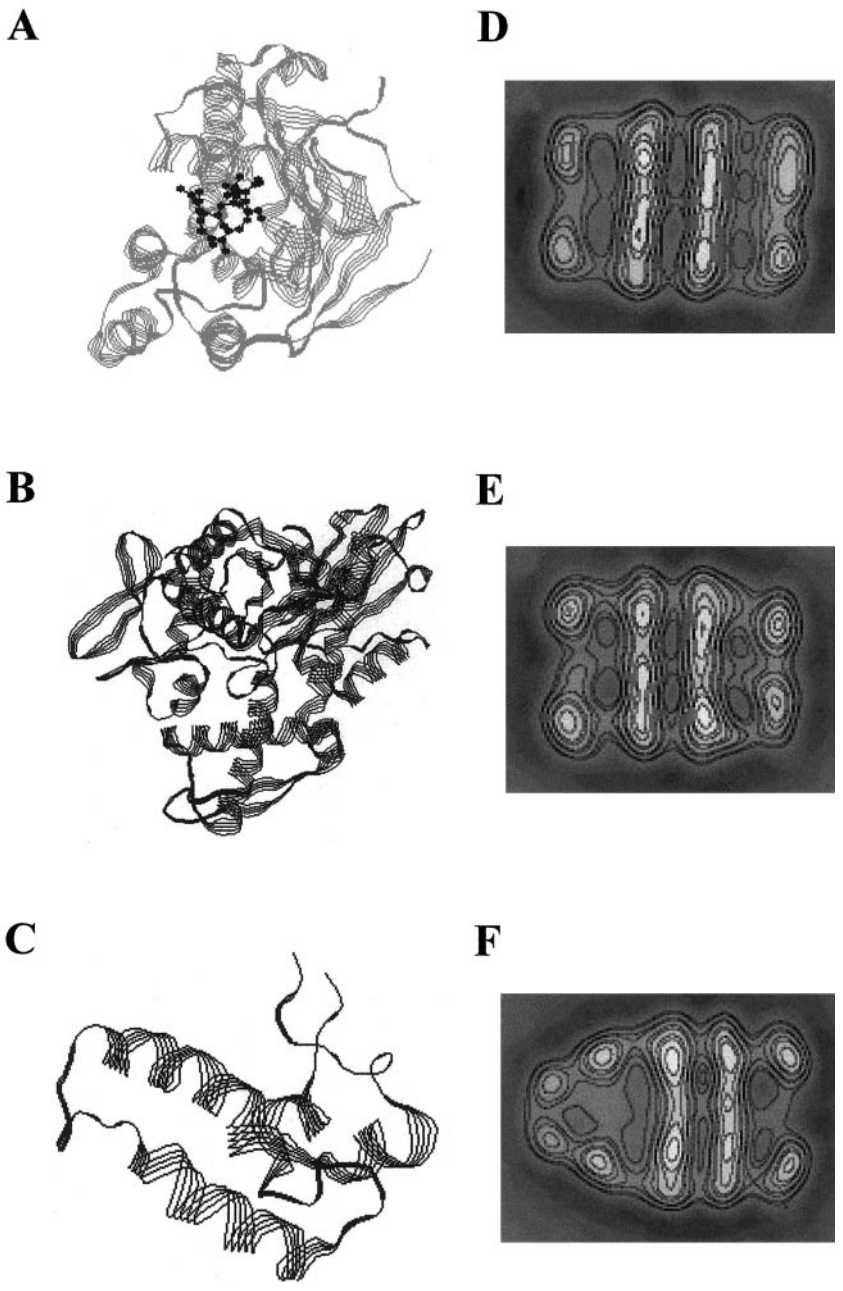

Figure 2. Structures of selected molecular chaperones. A, Ribbon diagram derived from the 1.8-Å x-ray structure of the tetragonal form of the N-terminal domain of *Saccharomyces cerevisiae* Stress90 complexed with the specific inhibitor geldanamycin. This inhibitor binds to the ATP-binding site. B, Ribbon diagram derived from the 1.7-Å x-ray structure of the 44-kD N-terminal ATP-binding domain of bovine Stress70. C, Ribbon diagram derived from the NMR structure of the J-domain of the human DnaJ homolog HdJ1. D through F, Structures of *E. coli* GroE reconstructed from cryo-electron microscopy. D, GroEL; E, GroEL plus 30 mM ATP; F, GroEL-ATP-GroES (http://bioc09.uthscsa.edu/approximately seale/chap/em1.html).

between the constitutive and stress-induced proteins, and hereafter they will be discussed collectively as Stress70. Several excellent recent reviews cover the molecular details of stress protein induction in plants (e.g. Schöffl et al., 1998), which are outside the scope of this review.

Specific species of the Stress70 proteins are found in all subcellular compartments (Boston et al., 1996; Miernyk, 1997). While plant cytoplasmic Stress70 proteins are more closely related to their mammalian isologs, the mitochondrial and plastidic proteins are more similar to their prokaryotic counterparts. The rough ER luminal resident form of Stress70 is variously referred to as BiP (from early studies on its function as the IgG-binding protein of mammalian cells), GRP78 (78-kD Glc-regulated protein), and in yeast, KAR2 (karyogamy).

The prokaryotic and organellar Stress70 proteins do not function as chaperones by themselves, but rather in concert with two accessory or co-chaperone proteins. The functional association of these components is often referred to as the Stress70 chaperone machine (Miernyk, 1997; Bukau and Horwich, 1998) (Fig. 3). DnaK (Stress70) is the central component of the machine, and functions as a chaperone in association with DnaJ, a chaperone-activating protein (CAP), and GrpE, a nucleotide exchange factor (NEF). Eukaryotic counterparts of DnaJ are widespread and are referred to variously as DnaJ homologs, Hsp40, or Stress40.

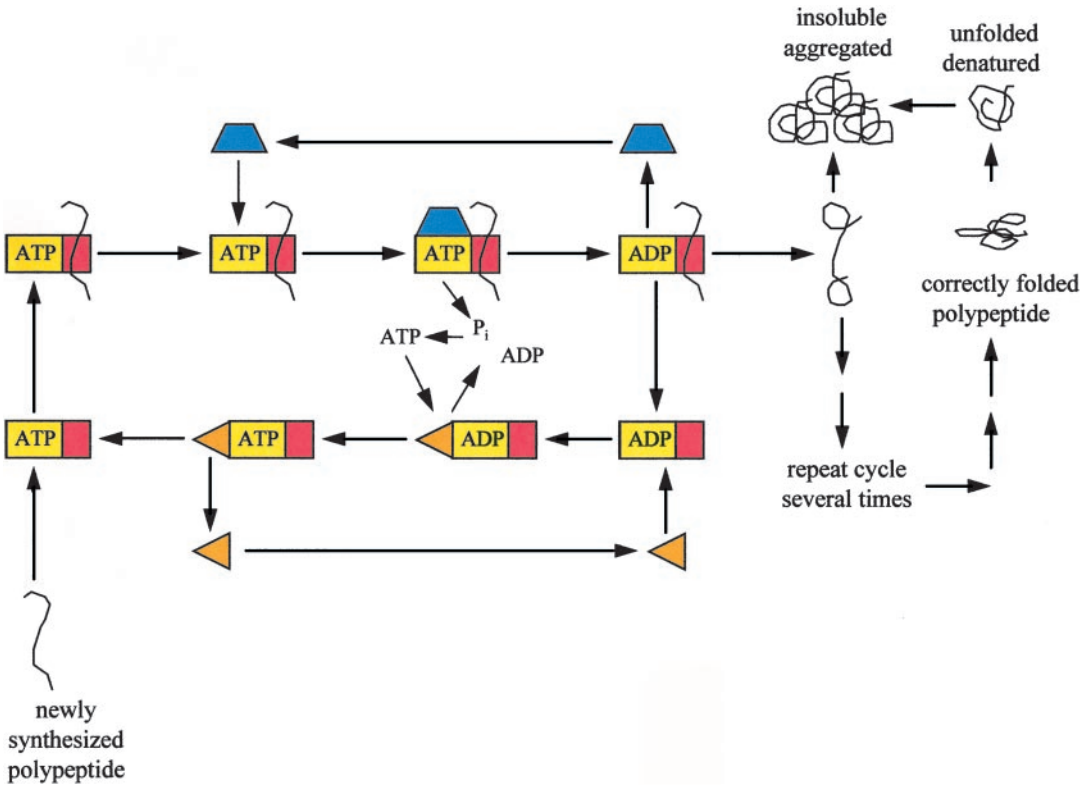

Figure 3. Schematic presentation of the Stress70 chaperone machine showing the interactions among the central ATP-dependent chaperone, the chaperone-activating protein/DnaJ, and the nucleotide exchange factor/GrpE. The 44-kD N-terminal ATPase domain of Stress70 is indicated as a gray rectangle; the peptide-binding domain is indicated in black. A newly synthesized polypeptide can be recognized and bound by the ATP-ligated form of Stress70 while still nascent or immediately after release from the ribosome. This binary complex is recognized by the chaperone-activating protein (blue trapezoid) that binds and stimulates ATP hydrolysis. The ADP-ligated form of Stress70 has a lower affinity for the polypeptide, which is then released from the machine. If the polypeptide folds incorrectly, it is prone to aggregation. Similarly, a mature correctly folded protein will aggregate when unfolded/denatured. The ADP-ligated form of Stress70 is recognized by the nucleotide exchange factor (orange triangle) that promotes the exchange of ADP for ATP, allowing another cycle by the machine. A typical protein would require several cycles of binding and release before reaching the final correctly folded conformation. The polypeptide chain is not folded by the machine, but rather "held" in a conformation that allows expression of the folding information present in the primary sequence.

Thus far the occurrence of eukaryotic cytoplasmic nucleotide exchange proteins has been controversial, although structurally related GrpE homologs have been found in mitochondria and plastids.

While both the primary sequence and structural organization of the Stress70 proteins are highly conserved, results from recent studies suggest that the Stress70 reaction cycle might be significantly different in prokaryotes and the eukaryotic cytoplasm. The eubacterial paradigm begins with DnaJ binding to an unfolded peptide and this binary complex subsequently interacts with ATP-ligated DnaK. When ATP occupies the nucleotide site of DnaK, the chaperone is in the open state and can effectively bind the extended peptide (Fig. 2B). DnaJ promotes the hydrolysis of ATP by DnaK, and the concomitant conformational change drives the release of a more structurally organized portion of the polypeptide chain. GrpE subsequently facilitates the exchange of ADP for ATP. This cycle repeats until a stable structure is achieved. In the mammalian Stress70 cycle, the unfolded peptide reacts directly with HSP70 (Fig. 3). Next, the DnaJ homolog binds to the polypeptide/HSP70 complex and stimulates the ATPase activity. The HSP70 complex then adopts the more stable ADP-ligated state. This state is further stabilized by association with Hip, the HSP70 interacting protein. In the absence of a nucleotide exchange factor, the cycle ends with slow dissociation of ADP, acquisition of the next ATP molecule, and release of the polypeptide from the ATP-bound, open form of Stress70.

In addition to Hip and Hop, which are mentioned in the section on HSP90, there are at least two other proteins that bind to the mammalian Stress70 complex and can potentially control chaperone activity. In keeping with the Hip/Hop vernacular, these have been termed Hap and Hup. Hap, the HSP70-associating protein, also known as BAG-1 or Rap46, is a negative modulator of Stress70 chaperone activity. Hup, the HSP70-unbinding protein, also known as p16 or Nm23, is ostensibly a nucleoside diphosphate ki-

nase. Hup is also a stress protein and modulates HSP70 chaperone activity by promoting dissociation (unbinding) of the chaperone-target complex. The elaborate, but until now comprehensible, model of the mammalian Stress70 chaperone machine became a bit more difficult to fathom with a recent report that Hop interacts with TriC and modulates the protein folding activity of the chaperonin (molecular chaperone proteins that are related to GroE either by primary sequence homology or overall structure) by affecting nucleotide exchange (Gebauer et al., 1998). This would represent an unprecedented economy of function, with Hop regulating two structurally unrelated chaperone complexes.

Less is known about the Stress70 chaperone machine in the cytoplasm of plant cells. The HSP70 proteins and DnaJ homologs of plant cells are structurally similar to their microbial or mammalian counterparts (Boston et al., 1996; Miernyk, 1997). While GrpE is a well-defined nucleotide exchange factor for the prokaryotic Stress70 chaperone machine, it is currently believed that the mammalian cytoplasmic machine functions without such a component. A small Arabidopsis protein (GB U64825) that stimulates ATP/ADP exchange by a plant Stress70/DnaJ-homolog complex has recently been characterized (B. Kroczyńska and J.A. Miernyk, unpublished observations). This protein (which should possibly be named the HSP70-exchange protein) has no obvious sequence relationship with GrpE, Hip, Hop, Hap, or Hup. To date, only one other plant component, a soybean Hop homolog (accession no. GBX79770), has been cloned or studied. Thus, it appears that the plant cytoplasmic Stress70 chaperone machine has some features in common with both bacterial and mammalian complexes, and perhaps unique features as well.

The Chaperonins

The chaperonins comprise the best-studied family of molecular chaperones (Boston et al., 1996; Netzer and Hartl, 1998; Sigler et al., 1998). They are double-ring, oligomeric structures that provide a closed compartment that shields folding proteins from the cellular environment (Fig. 2, D–F). Based upon evolutionary relationship, there are two distinct groups of chaperonins. The archetype group I chaperonin is GroEL, which occurs in the cytoplasm of eubacteria and within the mitochondria and plastids of eukaryotic cells, where it is referred to as cpn60 (Boston et al., 1996). Group II chaperonins include the archaeal thermosome and the eukaryotic cytoplasmic chaperonin complex, variously referred to as C-cpn, TCP-1 (T-complex protein-1), CCT (chaperonin containing TCP-1), or TriC (TCP-1 ring complex).

Members of both groups require energy derived from ATP hydrolysis to drive protein folding, and share the same basic mode of action. However, details at the molecular level differ considerably (Sigler et al., 1998). Group I chaperonins depend upon a partner protein, chaperonin 10 (cpn10, or GroES in bacteria), while a distinct cpn10 protein is not part of the group II system.

In group I chaperonin-mediated protein folding, an asymmetrical double-ring structure, in which GroES/cpn10 is bound to only one end, acts as the polypeptide acceptor (Fig. 2F). Binding occurs exclusively to the GroEL/cpn60 ring not occupied by GroES/cpn10. After polypeptide binding, a round of ATP hydrolysis in the occupied GroEL/cpn60 ring induces release of GroES/cpn10 (Fig. 2D). After release, GroES/cpn10 binds to the GroEL/cpn60 ring, where the polypeptide is bound. In this form the polypeptide is sequestered under a GroES/cpn10 "lid" in an enlarged central cavity. ATP then binds to the unoccupied ring, initiating polypeptide folding (Fig. 1E). ATP hydrolysis drives GroES/cpn10 release and allows the polypeptide to exit the GroEL/cpn60 central cavity. This cycle is repeated until the polypeptide reaches the final native structure (Bakau and Horwich, 1998; Sigler et al., 1998).

Nuclear-encoded group I chaperonins are found within the plastids and mitochondria of plant cells, and it was based upon studies of the chloroplast cpn60 (or Rubisco large-subunit-binding protein, as it was previously known) that the molecular chaperone concept was revived and extended (Hemmingsen et al., 1988). The structures and reaction cycles of the plant mitochondrial chaperonins are highly homologous to those of GroEL/S. The reaction cycle of the plastid chaperonins is also homologous to that of the prokaryotic counterparts. There are, however, some intriguing structural differences.

Plastids contain two distinct sets of GroEL homologs, cpn60α (61 kD) and cpn60β (60 kD). It remains unclear if there are distinct α- and β-ring structures in vivo or if there are structures composed of varying proportions of α- and β-subunits. The existence of distinct subunits implies distinct structures and distinct protein target specificity, however, this has not yet been experimentally demonstrated.

Initial analyses of chloroplast cpn10 revealed that the protein was unusually large at 21 kD. The spinach cpn10 cDNA encodes an open reading frame that consists of two cpn10 sequences fused head to tail, explaining this apparent anomaly (Boston et al., 1996). It has been suggested that the two halves of the protein might play subtly different roles in vivo. The large fused-dimer form of cpn10 is found only in photosynthetic eukaryotes (Boston et al., 1996). It is noteworthy that cyanobacterial cpn10 is of the usual size, suggesting that the apparent gene-fusion event took place after the presumptive endosymbiosis that gave rise to chloroplasts.

In mammalian cells, TriC forms a double-ring structure that is very similar in appearance to GroEL. The TriC is hetero-oligomeric; nine different subunits have been characterized (α–λ). TriC has been best characterized as a chaperone for actin and tubulin in mammalian cells (Boston et al., 1996; Sigler et al., 1998). In contrast to the organellar chaperonins, relatively little is known about TriC in plant cells. The larger subunit (e.g. α or 1) has been cloned from Arabidopsis and oat, and the deduced amino acid sequences are similar to those of their mammalian and yeast counterparts. In plant cells the TriC subunits are of relatively low abundance, and there have not yet been any reports of specific chaperone activity.

The Small Stress Proteins

The low-molecular-mass (small) heat shock proteins (smHSPs) (15–30 kD) are ubiquitous among eukaryotes, and represent a particularly important class of molecular chaperones in plant cells (Vierling, 1991). The smHSPs share a conserved C-terminal domain with the mammalian α-crystallin proteins. Biochemical analyses indicate that in vivo the smHSPs are found not as monomer/dimers but, rather, in high-molecular-mass complexes of 200 to 400 kD. Despite the occurrence of multiple species of smHSPs, they appear to exist as separate high-molecular-mass homo-complexes rather than as mixed-subunit hetero-complexes. This suggests subtly different functions for the different smHSPs, even when present in the same cellular compartment.

In vitro, the smHSPs can both facilitate the reactivation of chemically denatured proteins and prevent heat-induced protein aggregation (Boston et al., 1996). Based on these observations, it is likely that the smHSPs can also act in vivo as a type of molecular chaperone. In marked contrast to members of the other classes of molecular chaperones, the activity of the smHSPs is ATP independent.

In higher plants, six nuclear gene families encoding smHSPs have been described. Each family encodes proteins localized within distinct cellular compartments, including the cytoplasm, plastids, rough ER, and mitochondria (Waters et al., 1996). Additionally, specific smHSPs are expressed during various phases of plant development (Boston et al., 1996).

Calnexin and Calreticulin

Calnexin (CNX) is a rough ER-localized chaperone that is also a low-affinity, high-capacity calcium-binding protein (Helenius et al., 1997; Crofts and Denecke, 1998). CNX is a non-glycosylated type I integral membrane protein with a relatively short cytoplasmic domain and a large rough ER lumenal domain. Adjacent to the CNX transmembrane region is the P domain, a Pro-rich sequence that participates in glycan recognition and chaperone function. CNX binds ATP, although no ATPase activity has been reported, and ATP binding promotes a monomer to oligomer shift (Chapman et al., 1997). CNX interacts with a wide range of newly synthesized proteins, then dissociates from these proteins upon folding prior to export or further assembly (Helenius et al., 1997). Thus, CNX is a component of the rough ER quality control system that allows only proteins that have acquired a final native conformation to move on through the secretory pathway (Boston et al., 1996).

Calreticulin (CRT) is a soluble lumenal protein of the rough ER. Like CNX, CRT is a low-affinity, high-capacity calcium-binding protein (Borisjuk et al., 1998). CRT has been found associated with other proteins in the rough ER lumen, and has similar behavior and target specificity to that of CNX. Possibly, CNX and CRT cooperate in the rough ER as chaperones in the folding of secretory proteins (Helenius et al., 1997; Crofts and Denecke, 1998).

Because of the environment with the lumen of the rough ER (Huppa and Ploegh, 1998), protein folding is a particularly complex process involving multiple molecular chaperones and folding catalysts. Many of the interactions must occur sequentially in order for proteins to achieve the correct native structure. While CNX binds to all folding intermediates, CRT associates preferentially with the earliest oxidative species either immediately prior to or coincident with disulfide bond formation (Helenius et al., 1997). Thus, recognition and binding by CRT seems to precede interaction with the rough ER-localized folding catalysts. Overall, CNX and CRT promote correct protein folding in the rough ER by inhibiting aggregation, preventing premature oxidation and oligomerization, and suppressing degradation of incompletely or incorrectly folded intermediates.

FOLDING CATALYSTS ACCELERATE THE RATE OF FOLDING

In contrast to molecular chaperones, the folding catalysts are conventional enzymes. Protein-folding catalysts accelerate the slow chemical reactions in protein folding that might otherwise be rate limiting. Without the acceleration of the rate-limiting reactions, cellular proteins would be trapped in intermediate states, and most folding intermediates are vulnerable to aggregation and to non-productive interactions with other proteins (Dobson et al., 1998). The two best-studied protein-folding catalysts are protein disulfide isomerase (EC 5.3.4.1) and peptidyl-prolyl cis-trans-isomerase (EC 5.2.1.8).

Protein Disulfide Isomerase

Protein disulfide isomerase (PDI) is a member of the thioredoxin superfamily of proteins, and contains two copies of the characteristic active-site motif CXXC. Protein disulfide isomerase facilitates folding through its ability to reduce or oxidize disulfide bridges in the presence of an oxidizing or reducing agent such as glutathione (Huppa and Ploegh, 1998). To accomplish this, PDI expedites disulfide interchange by shuffling the disulfide bonds to quickly find the most thermodynamically stable pairing (Laboissière et al., 1995).

In eukaryotic cells, PDI is localized exclusively within the lumen of the rough ER. This is a unique environment for protein folding in terms of pH, redox conditions, and high calcium ion concentrations (Huppa and Ploegh, 1998; Møgelsvang and Simpson, 1998).

Peptidyl Prolyl Cis-Trans Isomerases

Attempts to follow the literature on PPI are complicated by the diverse terminology employed by various researchers. In addition to the formal enzyme name, PPIs are variously referred to as: cyclophilins, immunophilins, rotamase, FKBPs, and parvulins (Dolinski and Heitman, 1997). The term immunophilin comes from the ability of PPI to bind immunosuppressive drugs. The cyclophilins are immunophilins that bind the cyclic undecapeptide cyclosporin A, while FKBPs are immunophilins that bind the macrolide drug FK506.

Whatever the term used, PPI participates in protein folding by accelerating the cis-trans isomerization of prolyl peptide bonds. Pro isomerization can be the rate-limiting step in overall protein folding, and this reaction is accelerated 300-fold by PPI (Schmid, 1993). After isomerization, the correct form of Pro is stabilized by the polypeptide secondary or tertiary structure.

While most PPI proteins are localized in the cytoplasm/nucleus, there are unique forms present within the rough ER/nucleus, the plastid stroma, and the mitochondrial matrix (Boston et al., 1996; Kurek et al., 1999).

What's in a Name?

The distinctions between chaperones and folding catalysts have in some instances become blurred. The folding catalysts bind to their target proteins, and in some in vitro experiments it has been observed that this binding can prevent aggregation. Can the folding catalysts also therefore be chaperones? Furthermore, site-directed mutagenesis has been used to modify catalytically essential residues in PDI or PPI, and the resultant mutant proteins still had chaperone activity in vitro. However, such results must be interpreted with care. Recently it was shown that a mutant form of PPI that had no activity in vitro with an artificial substrate, remained fully active when assayed with an unfolded polypeptide as the substrate (Scholz et al., 1997). The distinction between chaperones and folding catalysts remains that the latter accelerate the rate of target polypeptide folding.

PROTEIN FOLDING IN THE CYTOPLASM

The majority of nuclear-encoded proteins reside in the cytoplasm throughout their lifetime. It is important to identify which molecular chaperones and folding catalysts assist the newly synthesized cytoplasmic proteins to their final native state. This is a particularly interesting question in plant cells that are uniquely complex in their array of cytoplasmic chaperones.

Some proteins, especially those that are small, single-domain, and monomeric, are able to fold spontaneously in the cytoplasm without any external assistance, or perhaps only with the assistance of a folding catalyst. Other proteins are recognized by the Stress70 chaperone machine and bound while still nascent. There are three potential fates for this complex: the polypeptide might fold to the correct final structure and be released from the chaperone, it might be maintained in an unfolded translocation-competent conformation until transferred to the protein machinery associated with an organellar outer membrane, or the polypeptide might be transferred to a different chaperone for further folding, assembly, or oligomerization. Among the other known cytoplasmic chaperones, the chaperonins, the CCH, and the smHSPs are the best understood.

For the most part, the actual native target proteins for each of the chaperone systems remain unidentified. The inner chamber of the chaperonin structure, or "Anfinsen cage" as it is sometimes referred to, is limited to polypeptides no larger than 55 kD (Netzer and Hartl, 1998; Sigler et al., 1998). A precise structural model of the CCH has not yet been defined; however, in principle, the HSP70 and HSP90 components are capable of assisting the folding of proteins considerably larger than 55 kD. In mammalian cells the only CCH target polypeptides identified to date are various receptors and some protein kinases (Pratt and Toft, 1997). The oligomeric small HSPs are thought to bind target polypeptides on the external surfaces and should be capable of assisting in the folding of any size protein (Vierling, 1991; Boston et al., 1996).

For the most part, HSP100/Clp are thought to mediate the unfolding/refolding of damaged proteins (Schirmer et al., 1996; Hoskins et al., 1998), but a potential role in the initial stages of protein folding cannot be excluded. Protein refolding by HSP100/Clp is typically assisted by the Stress70 machine and/or HSP90.

The complexity of the various interactions has not yet been fully defined. While chaperone-mediated protein folding is generally considered to be unidirectional, there is some in vitro evidence for retrograde transfer between chaperones (Netzer and Hartl, 1998). Additionally, there are some instances in which more than two of the chaperone systems seem to be involved in the folding of a single target protein (i.e. the prion protein has been observed during its "normal" lifetime to be associated with HSP70, the chaperonins, and HSP90).

FROM THE CYTOPLASM INTO THE ER

Molecular chaperones have a dual interaction with proteins destined to reside within the ER or any of other compartment of the endo-membrane system. These proteins are first inserted into a channel in the ER membrane that is aligned with the ribosome (Vitale and Denecke, 1999). Translocation must occur in a luminal direction. The directionality and driving force for polypeptide translocation into the ER lumen are provided by the resident Stress70 chaperone, BiP, which acts as a "molecular ratchet" (Matlack et al., 1999). Multiple BiP molecules interact with the translocating polypeptide to minimize any retrograde movement. The BiP:target polypeptide complex interacts with the J-domain of an integral membrane protein (Sec63p in yeast), which activates the ATPase function of the chaperone (Miernyk, 1997).

Once within the ER lumen, polypeptide folding begins immediately. As in the cytoplasm, non-productive folding is forestalled by interaction with the Stress70 chaperone machine (Møgelsvang and Simpson, 1998). BiP is absolutely essential for protein translocation and folding in the ER (Li et al., 1993). In addition to variously interacting with the specific ER homologs of Stress70 (BiP), DnaJ, PDI and PPI, the unique environment found within the rough ER lumen (Huppa and Ploegh, 1998) has given rise to the CNX and CRT chaperones and the PDI-folding catalyst (Chapman et al., 1997; Borisjuk et al., 1998; Crofts and Denecke, 1998; Møgelsvang and Simpson, 1998).

While all of the Stress70 proteins are subject to multiple posttranslational modifications (phosphorylation, *N*-methylation) (e.g. Miernyk et al., 1992), BiP is unique in that it is additionally ADP-ribosylated (Boston et al., 1996).

Specific roles have not been defined for any of the modifications, however, it has been shown that the extent of phosphorylation and ADP-ribosylation varies with physiological state or in some mutants, suggesting regulatory functions (Fontes et al., 1991).

PROTEIN FOLDING IN OTHER SUBCELLULAR COMPARTMENTS

Most organellar proteins are nuclear encoded and synthesized as precursors in the cytoplasm. As such, they will interact with various cytoplasmic chaperones prior to membrane translocation. After membrane translocation, protein folding within other subcellular organelles of a plant cell likely follows pathways similar to those in the cytoplasm or ER, and specific forms of HSP70, HSP90, the chaperonins, and the small HSPs have been found in plastids, mitochondria, and peroxisomes (Boston et al., 1996; Waters et al., 1996; Corpas and Trelease, 1997; Miernyk, 1997).

PROSPECTUS: WHERE DO WE GO FROM HERE?

I have attempted to provide an overview of the roles played by several of the better-understood chaperones in protein folding. Such a presentation is by definition limited in two ways: by space and by knowledge. SecB is a general chaperone for proteins that will be secreted by bacteria. MSF is a mammalian ATP-dependent protein that specifically chaperones mitochondrial precursors. Atx1p from yeast and UreE from bacteria are metal ion chaperones for copper and nickel, respectively. PapD and CssC chaperone pilus formation in prokaryotes. This list could go on, but the unifying feature of these otherwise diverse chaperones is that no homolog has yet been identified in plants. Thus, it seems likely that many different molecular chaperones remain, awaiting discovery.

Necessarily, most of the early studies on molecular chaperones and folding catalysts have been conducted in vitro, based upon mammalian and microbial paradigms, and used model target proteins. It is crucially important that we bridge the gap between these studies and actual in vivo activities. Only then can we be said to have truly "cracked the second half of the genetic code" (Ellis, 1991). A promising step in this direction has recently been provided by Forreiter et al. (1997). While this experimental system employs a model protein as the target, folding and the roles of molecular chaperones in this process can be studied in vivo in real time. Developing further methods to identify and study the bona fide targets of specific chaperone/folding systems in vivo remains a substantial challenge to the ingenuity of plant cell biologists.

ACKNOWLEDGMENTS

Thanks to my friends and colleagues in the field of protein folding in plant cells, A. Breiman, C. Guy, K. Keegstra, P. Krishna, G. Lee, and E. Vierling, for providing information prior to publication and in some instances for their constructive suggestions. N.E. Coop was instrumental in the preparation of the figures.

Received February 1, 1999; accepted July 14, 1999.

LITERATURE CITED

Anfinsen CB (1973) Principles that govern the folding of protein chains. Science **181:** 223–230

Borisjuk N, Sitailo L, Adler K, Malysheva L, Tewes A, Borisjuk L, Manteuffel R (1998) Calreticulin expression in plant cells: developmental regulation, tissue specificity and intracellular distribution. Planta **206:** 504–514

Bork P, Dandekar T, Diaz-Lazzcoz Y, Eisenhaber F, Huynen M, Yuan Y (1998) Predicting function: from genes to genomes. J Mol Biol **283:** 707–725

Boston RS, Vittanen PV, Vierling E (1996) Molecular chaperones and protein folding in plants. Plant Mol Biol **32:** 191–222

Bouchez D, Höfte H (1998) Functional genomics in plants. Plant Physiol (1998) **118:** 725–732

Brunt SA, Perdew GH, Toft DO, Silver JC (1998) Hsp90-containing multiprotein complexes in the eukaryotic microbe *Achlya*. Cell Stress Chaperones **3:** 44–56

Buchner J (1999) Hsp90 & Co.: a holding for folding. Trends Biochem Sci **24:** 136–141

Bukau B, Horwich AL (1998) The Hsp70 and Hsp60 chaperone machines. Cell **92:** 351–366

Chapman KD, McAndrew RS, Huynh TT (1997) Cottonseed calnexin: identification and isolation of a membrane-bound molecular chaperone. Plant Physiol Biochem **35:** 483–490

Corpas FJ, Trelease RN (1997) The plant 73 kDa peroxisomal membrane protein (PMP73) is immunorelated to molecular chaperones. Eur J Cell Biol **73:** 49–57

Crofts AJ, Denecke J (1998) Calreticulin and calnexin in plants. Trends Plant Sci **3:** 396–399

Csermely P, Schnaider T, Sôti C, Prohászka Z, Nardai G (1998) The 90 kDa molecular chaperone family: structure, function, and clinical applications. Pharmacol Ther **79:** 129–168

Dobson CM, Sali A, Karplus M (1998) Protein folding: a perspective from theory and experiment. Angew Chem Int Ed **37:** 868–893

Dolinski K, Heitman J (1997) Peptidyl-prolyl isomerases. *In* J Sambrook, S Tooze, eds, Guidebook to Molecular Chaperones and Protein Folding Catalysts. Oxford University Press, London, pp 359–369

Ellis RJ (1991) Chaperone function: cracking the second half of the genetic code. Plant J **1:** 9–13

Fontes EBP, Shank BB, Wrobel RL, Moose SP, O'Brian GR, Wurtzel ET, Boston RS (1991) Characterization of an immunoglobulin binding protein homolog in the maize *floury*-2 endosperm mutant. Plant Cell **3:** 483–496

Forreiter C, Kirschner M, Nover L (1997) Stable transformation of an Arabidopsis cell suspension culture with firefly luciferase providing a cellular system for analysis of chaperone activity *in vivo*. Plant Cell **9:** 2171–2181

Gebauer M, Melki R, Gehring U (1998) The chaperone cofactor Hop/p60 interacts with the cytosolic chaperonin-containing TCP-1 and affects its nucleotide exchange and protein folding activities. J Biol Chem **273:** 29475–29480

Helenius A, Trombetta ES, Hebert DN, Simons JF (1997) Calnexin, calreticulin and the folding of glycoproteins. Trends Cell Biol **7:** 193–200

Hemmingsen SM, Woolford C, Van der Vies SM, Tilly K, Dennis DT, Georgopoulos CP, Hendrix RW, Ellis RJ (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. Nature **333:** 330–334

Hoskins JR, Pak M, Maurizi MR, Wickner S (1998) The role of the ClpA chaperone in proteolysis by ClpP. Proc Natl Acad Sci USA **95:** 12135–12140

Huppa JB, Ploegh HL (1998) The eS-Sence of -SH in the ER. Cell **92:** 145–148

Kurek I, Herman E, Erel N, Blecher O, Devos K, Breiman A (1999) A novel heat shock induced wheat FKBP-peptidyl prolyl*cis-trans* isomerase: developmental and heat induced expression. Plant Physiol **119:** 693–703

Laboissière MCA, Sturley SL, Raines RT (1995) The essential function of protein-disulfide isomerase is to unscramble nonnative disulfide bonds. J Biol Chem **270:** 28006–28009

Levitt M, Gerstein M, Huang E, Subbiah S, Tsai J (1997) Protein folding: the endgame. Annu Rev Biochem **66:** 549–579

Li X, Wu Y, Zhang D-Z, Gillikin JW, Boston RS, Franceschi V, Okita TW (1993) Rice prolamine protein body biogenesis: a BiP-mediated process. Science **262:** 1054–1056

Matlack KES, Misselwitz B, Plath K, Rapoport TA (1999) BiP acts as a molecular ratchet during posttranslational transport of prepro-α factor across the ER membrane. Cell **97:** 553–564

Miernyk JA (1997) The 70 kDa stress-related proteins as molecular chaperones. Trends Plant Sci **2:** 80–87

Miernyk JA, Duck NB, David NR, Randall DD (1992) Autophosphorylation of the pea mitochondrial heat shock protein 70 homolog. Plant Physiol **100:** 965–969

Møgelsvang S, Simpson DJ (1998) Protein folding and transport from the endoplasmic reticulum to the Golgi apparatus in plants. J Plant Physiol **153:** 1–15

Netzer WJ, Hartl FU (1998) Protein folding in the cytosol: chaperonin-dependent and -independent mechanisms. Trends Biochem Sci **23:** 68–73

Nieto-Sotelo J, Kannan KB, Martinez LM, Segal C (1999) Characterization of a maize heat-shock protein 101 gene, HSP101, encoding a ClpB/Hsp100 protein homologue. Gene **230:** 187–195

Reddy RK, Kurek I, Silverstein AM, Chinkers M, Breiman A, Krishna P (1998) High-molecular-weight FK506-binding proteins are components of heat-shock protein 90 heterocomplexes in wheat germ lysate. Plant Physiol **118:** 1395–1420

Saier MH Jr (1998) Genome sequencing and informatics: new tools for biochemical discoveries. Plant Physiol **117:** 1129–1133

Schirmer EC, Glover JR, Singer MA, Lindquist S (1996) HSP100/Clp proteins: a common mechanism explains diverse functions. Trends Biochem Sci **21:** 289–296

Schirmer EC, Lindquist S, Vierling E (1994) An Arabidopsis heat shock protein complements a thermotolerance defect in yeast. Plant Cell **6:** 1899–1909

Schmid FX (1993) Prolyl isomerase: enzymatic catalysis of slow protein folding reactions. Annu Rev Biophys Biomol Struct **22:** 123–143

Schöffl F, Prändl R, Reindl A (1998) Regulation of the heat-shock response. Plant Physiol **117:** 1135–1141

Scholz C, Schlinder T, Dolinski K, Heitman J, Schmid FX (1997) Cyclophilin active site mutants have native prolyl isomerase activity with a protein substrate. FEBS Lett **414:** 69–73

Sigler PB, Xu Zhaohui, Rye HS, Burston SG, Fenton WA, Horwich AL (1998) Structure and function in GroEL-mediated protein folding. Annu Rev Biochem **67:** 581–608

Vierling E (1991) The roles of heat shock proteins in plants. Annu Rev Plant Physiol Plant Mol Biol **42:** 579–620

Vitale A, Denecke J (1999) The endoplasmic reticulum: gateway of the secretory pathway. Plant Cell **11:** 615–628

Waters ER, Lee GJ, Vierling E (1996) Evolution, structure and function of the small heat shock proteins in plants. J Expl Bot **47:** 325–338