For Genetic Engineering Principals & Methods
Vol 24 2002.

BNL-68837

PROTEIN QUALITY CONTROL IN BACTERIAL CELLS: INTEGRATED NETWORKS OF CHAPERONES AND ATP-DEPENDENT PROTEASES.

John M. Flanagan and Maria C. Bewley

Biology Department, Brookhaven National Laboratory
Upton NY, 11973

OFFICIAL FILE COPY

INTRODUCTION

It is generally accepted that the information necessary to specify the native, functional, three-dimensional structure of a protein is encoded entirely within its amino acid sequence; however, efficient reversible folding and unfolding is observed only with a subset of small single-domain proteins. Refolding experiments often lead to the formation of kinetically-trapped, misfolded species that aggregate, even in dilute solution. In the cellular environment, the barriers to efficient protein folding and maintenance of native structure are even larger due to the nature of this process. First, nascent polypeptides must fold in an extremely crowded environment where the concentration of macromolecules approaches 300-400 mg/mL and on average, each ribosome is within its own diameter of another ribosome (1-3). These conditions of severe molecular crowding, coupled with high concentrations of nascent polypeptide chains, favor nonspecific aggregation over productive folding (3). Second, folding of newly-translated polypeptides occurs in the context of their vectorial synthesis process. Amino acids are added to a growing nascent chain at the rate of ~5 residues per sec, which means that for a 300 residue protein its N-terminus will be exposed to the cytosol ~1 min before its C-terminus and be free to begin the folding process. However, because protein folding is highly cooperative, the nascent polypeptide cannot reach its native state until a complete folding domain (50-250 residues) has emerged from the ribosome. Thus, for a single-domain protein, the final steps in folding are only completed post-translationally since ~40 residues of a nascent chain are sequestered within the exit channel of the ribosome and are not available for folding (4). A direct consequence of this limitation in cellular folding is that during translation incomplete domains will exist in partially-folded states that tend to expose hydrophobic residues that are prone to aggregation and/or misfolding. Thus it is not surprising that, in cells, the protein folding process is error prone and organisms have evolved "editing" or quality control (QC) systems to assist in the folding, maintenance and, when

necessary, selective removal of damaged proteins. In fact, there is growing evidence that failure of these QC-systems contributes to a number of disease states (5-8). This chapter describes our current understanding of the nature and mechanisms of the protein quality control systems in the cytosol of bacteria. Parallel systems are exploited in the cytosol and mitochondria of eukaryotes to prevent the accumulation of misfolded proteins.

CHAPERONE SYSTEMS IN EUBACTERIA

Misfolded/aggregated proteins can arise at numerous stages in the life cycle of all cells (Figure 1). These proteins are potentially toxic to cells, since they can act as nucleation centers for non specific aggregation of essential proteins, or perhaps act as dominant-negative effectors by binding to their normal cellular targets without being able to perform their intended functions. The most basic method for minimizing the concentration of misfolded proteins is to ensure that all proteins fold correctly in the first place. In Eubacteria, two major ATP-dependent chaperone systems have been implicated in *de novo* protein folding, the Hsp70-system (DnaK/DnaJ/GrpE) (Figure 2) (9-11) and the cylindrical chaperonin complexes (GroEL/GroES, Cpn60/Cpn10 or Hsp60/Hsp10) (Figure 3) (12, 13). In both systems, nucleotide binding and hydrolysis regulate the binding and release of polypeptide substrates; however, the Hsp70 and chaperonin systems are structurally and functionally distinct and represent radically different approaches to enhancing the efficiency of folding (14). In addition to misfolded proteins generated during *de novo* folding, there are a significant number of marginally-stable proteins in bacterial cells, which as a result of environmental or genetic stresses unfold and aggregate after they have achieved their native state. For most of these proteins, the Hsp70 and Hsp100 classes of chaperones have been shown to be important for their reactivation *in vitro* and *in vivo* (10, 15-18).

Bacterial Homologues of the Hsp70-system

Escherichia coli possess three Hsp70 homologues, DnaK, HscA and Hsc62 (19-21): DnaK is the most functionally important in *de novo* folding for and in the rescue of damaged proteins. Like most Hsp70 homologues, it requires the cellular function of two additional proteins; the first is a member of the Hsp40 family, and the second is a nucleotide exchange factor (19, 22-25). In *E. coli*, there are three authentic Hsp40 homologues, DnaJ, CbpA and DjlA, and three additional proteins that contain one or more Hsp40-like domains, Hsc20, and the genes *ybeS* and *ybeV* (19, 21, 26-28), of which DnaJ is the cognate Hsp40 for DnaK. DnaJ and its bacterial and eukaryotic homologues are chaperones in their own right, since they bind to a wide range of unfolded polypeptides, preventing their aggregation and maintaining them in a foldable state (22, 29). The third member of the *E. coli* DnaK triumvirate is GrpE, a homodimer that serves as the ADP-release factor for DnaK and HscA, allowing these proteins to complete their catalytic cycle (24, 25, 30). Homologues of GrpE are found in all bacteria that have an Hsp70 system, as well as eukaryotes (31-36). Together, these three proteins, termed the Hsp70- or DnaK-system, are involved in *de novo* protein folding, refolding and degradation (9, 11, 17, 23). Over the last several years much has been learned about the structures and functions of DnaK/DnaJ/GrpE (Figure 2). DnaK, like all Hsp70 homologues, is a two-domain protein comprised of an amino terminal ~44 kDa ATPase domain, and a carboxyl-terminal ~18 kDa polypeptide binding domain (Figure 2a). Binding and release of substrates rely on modulating the intrinsic peptide affinity of DnaK by cycles of ATP binding and hydrolysis. In the ATP-bound state, DnaK binds and releases substrates rapidly, whereas in the ADP-bound form, it binds and releases substrates slowly (14). All known Hsp70 homologues, including DnaK, recognize short (6-8 residues) extended peptide segments that are enriched in hydrophobic amino acids (37, 38). This accounts for their ability to interact with most

unfolded polypeptides, while ignoring folded ones. Binding of DnaK to polypeptide substrates maintains them in an unfolded, but foldable, state even though it only interacts with a small portion of the amino acid sequence of the substrate (22). The ability of DnaK to stabilize the unfolded state of the bound substrate can be explained by the observation that DnaK interacts preferentially with regions in its substrates that are typically buried in their hydrophobic core when folded. Previous studies with a model protein indicate that removing as few as five residues from this hydrophobic core prevents folding (39, 40), thus bind of DnaK to its substrates has the same effect, and hence prevents refolding. Interestingly, DnaK prefers to bind short peptides compared with unfolded polypeptides (22, 24). Efficient binding of unfolded substrates both *in vitro* and *in vivo* requires the co-chaperone DnaJ (22, 24, 29). The *E. coli* DnaJ chaperone is the canonical type I Hsp40 ortholog. It is comprised of at least four domains: the N-terminal J-domain that contains the highly-conserved YHPD sequence which is present in all Hsp40 homologues, the G/F domain that is rich in glycine and phenylalanine residues, a zinc-finger domain and a less well conserved C-terminal domain (29) (Figure 2b). Biochemical studies show that the J- and G/F-domains are primarily responsible for mediating the DnaK-DnaJ interaction while the zinc-finger and C-terminal domains are required for substrate binding (24, 41-44). In Type II Hsp40 homologues the zinc-finger domain is replaced with a functionally similar, but structurally distinct, domain. Biochemical analysis suggests that DnaJ recognizes a binding motif in the substrate that consists of a hydrophobic stretch of approximately eight residues enriched in aromatic, large aliphatic and arginine side chains. The overall hydrophobicity of this motif probably accounts for the ability of DnaJ to inhibit substrate aggregation (45). Interestingly, both DnaJ and DnaK interact with similar types of peptides. However, unlike DnaK, DnaJ binds peptides that contain both D- and L-amino acids equally well, suggesting that binding is not restricted by backbone contacts. Consistent with this idea, there is no clear substrate-binding pocket in the structure of DnaJ (46, 47) that would allow extensive

contacts between DnaJ and the backbone of the bound substrates, suggesting that substrate binding may be achieved through interactions of the substrates' hydrophobic side chains with hydrophobic patches on the surface of DnaJ. These observations led Bukau and his colleagues to speculate that DnaJ acts as a scanning factor for DnaK by identifying hydrophobic protein surfaces and initiating the functional cycle of the DnaK chaperone by associating with exposed hydrophobic patches and subsequently transferring these or adjacent patches to the peptide-binding site on DnaK (45).

The major function of the DnaK system is to reduce the tendency of unfolded polypeptide chains to misfold, due to premature hydrophobic collapse, or to aggregate (4). Recent mechanistic studies have illuminated the basic details of the DnaK/DnaJ/GrpE reaction cycle (Figure 2d). Initially, DnaJ binds the nascent or unfolded polypeptide substrate, from which it is then transferred to DnaK in its ATP-bound state. Conversion of ATP to ADP stabilizes the substrate-DnaK interaction. Subsequently, both inorganic phosphate and DnaJ are released from the complex followed by GrpE-promoted ADP release. Rebinding of ATP to DnaK triggers dissociation of the bound substrate, which may either fold or undergo another round of interaction with the DnaK chaperone system (24, 45, 48-50). A key feature of this reaction cycle is that the bound substrate is not folded while in a complex with the DnaK system; instead it is maintained and released in a folding-competent form.

GroEL, the Bacterial Group I Chaperonin

The second major class of bacterial chaperones is the chaperonin, GroEL (Figure 3). Members of the GroEL family are termed Group I chaperonins and form large cylindrical homooligomeric protein complexes consisting of two, stacked heptameric rings (51-53). Group I chaperonins are found in all eubacteria and in eukaryotic mitochondria, and are distinguishable from Group II chaperonins, which

are heterooligomers, comprised of two, stacked octameric rings (4). Group II chaperonins are found in archaebacteria and in the eukaryotic cytosol. Both Group I and II chaperonins share a common overall architecture and function in *de novo* protein folding; however, their sequences are highly divergent (4). In contrast to the DnaK chaperone system, substrates are thought to fold while they are in association with the chaperonin. GroEL, the canonical Group I chaperonin, binds unfolded substrates in a central axial cavity, which is large enough to accommodate substrates up to ~50 kDa (Figure 3). The efficient folding of most GroEL substrates requires the participation of the heptameric GroES co-chaperone (54). Recently, the structures of GroEL and GroES alone and in their complex have been solved at high resolution (52, 53, 55, 56)(Figure 3a). Each GroEL subunit consists of two discrete domains, joined by a hinge-like intermediate domain. 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 central cavity and bind the unfolded substrate polypeptides through hydrophobic contacts. Unfolded GroEL substrates are bound entirely within this cavity, thereby sequestering them from the remainder of the cellular macromolecules (57-61). Unlike DnaK, GroEL does not appear to bind linear peptides, but interacts efficiently with non-native proteins. Binding presumably involves multivalent interactions between hydrophobic surfaces on the unfolded substrate with similar surfaces on multiple subunits of GroEL (59). Interestingly, the substrate-binding residues in the apical domain of GroEL are also responsible for interacting with the cofactor, GroES (61), a ring-shaped complex composed of seven identical 10 kDa subunits that is essential for GroEL-mediated folding (54, 55, 58). Based upon both structural and biochemical data, a mechanism has been proposed to explain GroEL-mediated protein folding (Figure 3b). Initially, unfolded substrates bind in the interior cavity of nucleotide-free GroEL. Binding of ATP and GroES to GroEL triggers a conformational change in the apical domains that displaces the substrate from its binding sites and releases it into the central cavity, which is now lined with hydrophilic side

chains (57). GroES also promotes the concerted hydrolysis of seven ATP molecules in the proximal (*cis*) ring of GroEL. Enclosure of the substrate polypeptide within the chamber of this asymmetric GroEL-GroES-ADP complex is essential for folding. The substrate remains enclosed for ~15 sec within the cavity, which functions as an Anfinsen cage, isolating the polypeptide under conditions of infinite dilution. Binding of seven ATP molecules to the *trans* ring in GroEL triggers the release of GroES and ADP from the *cis* ring. This returns the apical domains to a binding-competent conformation that exposes their hydrophobic sites toward the cavity, permitting the rebinding of a still unfolded substrate. However, if folding has been completed, the substrate will no longer expose sufficient hydrophobic surfaces to mediate binding and it will be released.

Additional Bacterial Chaperones Implicated in Quality Control

In addition to the GroEL/GroES chaperone system, most bacterial and eukaryotic cells possess another large, barrel-like, chaperone which is a Hsp100 family member. The *E. coli* protein ClpB is a homooligomeric ATPase comprised of 96 kDa subunits. Currently, the exact oligomeric state of this protein is under debate; however, it is most likely a hexamer, based upon the structures of the yeast homologue and that of the closely-related Clp ATPases (18, 62-64). The Clp/Hsp100 superfamily is comprised of two major classes of proteins that are currently divided into at least 8 subfamilies (65). The class I proteins (ClpA, ClpB/Hsp100, ClpC and ClpD), contain two ATPase domains separated by an intervening variable-length linker domain. The class II proteins (ClpM, ClpN, ClpX, ClpY/HslU) contain only a single ATPase domain. All members of the Clp/Hsp100 family possess the chaperone-like ability to bind protein substrates and unfold them. In addition, some, like ClpA, ClpC, ClpX and

ClpY(HslU), form complexes with proteolytic components to generate functional ATP-dependent proteases (Figure 4).

In vitro, ClpB possesses a protein-activated ATPase activity (66). Although little is known about the mechanism of ClpB action, it is thought to be involved in the recovery of misfolded proteins, based upon its ability to prevent/dissolve protein aggregates both *in vivo* and *in vitro*. Moreover, by analogy with the better-characterized homologues in the Clp/Hsp100 family, ClpA, ClpX and ClpY/HslU, ClpB probably acts as an unfoldase, binding misfolded and/or aggregated proteins and completely unfolding them (65, 67-71). By analogy with ClpA, ClpB may release bound unfolded substrates by translocating them through a narrow pore that runs through the center of the oligomer. Substrates are thought to exit from this complex in an extended conformation, which may approximate the conformation of a nascent chain exiting the ribosome. A more detailed description of the mechanism of this substrate unfolding and translocation by proteins of the Clp/Hsp100 family will be presented below in the context of their role in ATP-dependent proteolysis.

In addition to the chaperones discussed above there are a number of other less-well characterized, though widely distributed, bacterial chaperones, including trigger factor (TF), HtpG and two small heat shock proteins IbpA and IbpB. TF is a highly expressed (~20,000/cell) nonessential protein that binds stoichiometrically to the 50S subunit of actively-translating ribosomes (~15,000/cell). *In vitro* studies indicate that TF can be crosslinked to nascent polypeptide chains, and that its proline isomerase activity may be important in the folding of some proteins (72-74). HtpG is also a nonessential gene in *E. coli* (75-78). This protein is homologous to the Hsp90 chaperones in eukaryotes that are involved in the folding of a small number of very specific substrates. In eukaryotes, the Hsp90 chaperone does not act alone, but instead is involved in large complexes containing several other chaperones, including Hsp70 and Hsp40 (4, 79). It is possible that HtpG may serve a similar purpose in

prokaryotes, or perhaps may be involved in binding misfolded proteins. The small heat-shock proteins IbpA/B are even less well understood (80-84). These proteins appear to bind unfolded polypeptides, and may act as holding chaperones for the Hsp70-system and/or chaperonins. The recently determined X-ray structure of a small heat-shock protein from *Methanococcus janashii,* which is a member of this family, provides little additional insight into their mechanism (85-87). The structure is a hollow sphere of 24 subunits with octahedral symmetry. Current speculation is that unfolded polypeptides bind relatively nonspecifically to exposed hydrophobic surfaces on the outside of this complex, thereby preventing their aggregation.

In vivo Roles of Chaperones During *de novo* Protein Folding

Until very recently, the importance of chaperones in *de novo* protein folding could only be inferred from *in vitro* refolding studies and from genetic studies which implicated them in the folding of a few specific substrates (82, 88-90). In fact, for the Hsp70/DnaK-system, early indications suggested that it did *not* play a major role in *de novo* protein folding since it was dispensable under non-stress conditions (91). The initial difficulties in identifying specific cellular roles for these chaperones point to an important feature of the protein QC-systems in both prokaryotes and eukaryotes: they are highly redundant and there is a high degree of cooperation between the various components. Recent advances in whole-cell-based proteomics approaches have begun to clarify the role and mechanisms of these systems *in vivo* (9, 11, 92).

A direct role for DnaK/DnaJ in *de novo* folding was demonstrated in pulse-chase studies which show that DnaK interacts transiently with newly-synthesized polypeptides over a broad size range, from ~15 kDa to 167 kDa, binding preferentially to multidomain proteins with chains that range in size from

30-70 kDa (9, 11). Around 10% of all soluble polypeptides (340 cytosolic proteins) associated with DnaK at the earliest chase times were released within 2 minutes. Interaction with these nascent polypeptides occurred co-translationally, since after treatment with puromycin, which is incorporated into the nascent chain and affects its releases from the ribosome, ~20% of the DnaK-bound polypeptides could be immunoprecipitated with anti-puromycin antibodies. Similar studies in eukaryotes had previously implicated the Hsp70-system in binding nascent chains (89, 93, 94), suggesting an evolutionarily conserved role(s) for this system in preventing misfolding at the ribosome. However, if taken at face value, it is still unclear why DnaK is a nonessential gene, and why it could not be crosslinked to short nascent polypeptides *in vitro*. A solution to this dilemma was suggested by the observation that TF, the only other chaperone known to bind to nascent chains, is also not essential under normal growth conditions. Examination of the fate of nascent chains in *E. coli* lacking TF (*tig*Δ) revealed a two-to-threefold increase in the amount of nascent polypeptides associated with DnaK, suggesting that TF and DnaK cooperate in chaperoning nascent chains. Moreover, in *tig*Δ cells, the proteins associated with DnaK shift to include low-molecular-weight species, consistent with the idea that, under normal cellular conditions, TF associates with nascent chains prior to DnaK, preventing its interaction with very short nascent chains. This also explains why, in previous studies, DnaK was not crosslinked to short ribosome-associated nascent chains. This apparent functional overlap between TF and DnaK was confirmed genetically in experiments that showed that the *tig*Δ *dnaK*Δ double mutants were inviable (95). Such double-mutant strains contain protein aggregates formed from both newly-synthesized and preexisting proteins. Identification of the aggregated proteins in these strains revealed that they are predominantly large proteins that have more than one independent folding domain. This observation suggests an important role for the DnaK system in folding multidomain proteins. Typically, the isolated domains of multidomain proteins are efficiently refolded *in vitro*, while the intact proteins

fold inefficiently under similar conditions, suggesting interactions between domains in the unfolded state can inhibit their refolding. Interaction with the DnaK system may allow the domains to fold independently, as occurs with the isolated domains. This mechanism may be especially important in eukaryotic cells where a larger proportion of their proteins are expected to consist of two or more domains.

By contrast, GroEL and GroES are essential for growth of *E. coli* under all conditions (12). However, this requirement could result if GroEL facilitated the folding of just one or a few essential proteins. To begin understanding the extent of proteins in bacteria that require the GroEL/GroES chaperone during *de novo* folding, Hartl and his co-workers have identified *in vivo* substrates for this system by pulse labeling followed by immunoprecipitation of GroEL under native conditions (13). In these experiments, a subset of ~300 proteins appear to fold in a complex with GroEL. Unlike TF and the DnaK system, GroEL does not interact with ribosome-bound polypeptides, but binds polypeptides post-translationally. The set of newly-synthesized polypeptides interacting with GroEL is highly diverse, and accounts for ~15% of all cytoplasmic proteins under non-stress conditions (30-37°C). Heat shock at 42°C increased the fraction of GroEL-bound polypeptides to ~30% of the total. Most GroEL substrates are between 10 and 55 kDa in size with a clear cutoff for larger polypeptides, as expected from the size of the GroEL folding cavity. The majority of these substrates are released from GroEL in a chaperonin cycle (~15 sec) and are smaller than 25 kDa. Larger substrates, 25-55 kDa, remain in association with GroEL (100-150 sec on average) suggesting that these require multiple chaperonin cycles.

Sequential Roles For The DnaK-System and GroEL/GroES Chaperonins In Folding

These recent advances in our understanding of the mechanism and roles in *de novo* folding for the DnaK and GroEL systems beg the question: do these two systems act sequentially and cooperatively in folding substrates? In principal, the DnaK systems may act only on large multidomain proteins by allowing the individual folding domains to reach their native conformations separately. In this case, DnaK and GroEL would target different classes of substrates acting independently of each other. Alternatively, there could be a functional coupling between the DnaK and GroEL systems for some substrates. This seems to be a particularly attractive possibility since the DnaK system acts co-translationally and is present in the cell at concentrations that would allow it to interact with virtually every nascent polypeptide chain, whereas GroEL acts post-translationally and its oligomeric concentration is relatively low. Cooperation between these two systems would couple chaperone-mediated folding to translation and sequester the newly-synthesized non-native proteins from the bulk of the cytosol. Initial support for the sequential nature of chaperone interactions was suggested by experiments that examined the folding of model proteins that were either imported into mitochondria or chloroplasts, or translated in cell-free extracts, as well as by experiments with purified components (4, 22, 94, 96-99). In these systems, the polypeptide was initially bound and stabilized by DnaK/DnaJ or the Hsp70/Hsp40 homologues, and subsequently transferred to GroEL or its eukaryotic equivalent. Consistent with this model, overexpression of GroEL, which increases the flux of substrates through this chaperone, also increases the flux of substrates through DnaK/DnaJ, as expected if the DnaK-system acts upstream of GroEL (95). Currently there little direct *in vivo* evidence for this pathway in bacteria; however, several studies suggest that this it is important in *de novo* folding in yeast and mammalian cells (93). Figure 5 summarizes the salient features of our current understanding of the contributions and interactions of TF and the DnaK and GroEL systems to *de novo* protein folding.

Chaperone Catalyzed Rescue of Misfolded and Aggregated Polypeptides

The DnaK and GroEL chaperones play key roles in minimizing the amount of misfolding produced during synthesis and *de novo* protein folding. However, even under normal growth conditions, misfolding occurs both during and after *de novo* folding and increases dramatically as a result of genetic or environmental stresses, such as mutations or heat shock (17, 100-102). Under stress conditions, high levels of misfolded proteins that escape chaperone binding may reach an alternatively stable state as inactive, insoluble, aggregates. These aggregates are sometimes called inclusion, or Russell, bodies and are seen as electron-dense masses in electron micrographs of whole cells (103). A number of recent studies suggest the presence of one or more conserved, highly specialized pathways to solubilize and, when possible, reactivate misfolded/aggregated proteins that are operative under all growth conditions (15, 16, 23, 81, 83, 101, 103).

In vitro studies with purified proteins indicate that the DnaK system can protect several proteins from aggregation and irreversible thermal inactivation. Moreover, this system can reactivate some previously-denatured proteins when they are present in small aggregates *in vitro*, but it is inefficient in dissolving and reactivating proteins found in large aggregates (15). By contrast, several studies have suggested that *E. coli* ClpB can dissolve large protein aggregates (10, 104), although it does not efficiently reactivate these solublized proteins. Several laboratories have recently tested the possibility that the DnaK system and ClpB can cooperate in suppressing and/or dissolving and refolding protein aggregates *in vitro*. Alone, ClpB is capable of binding to proteins found in these large aggregates; however, it is inefficient in their refolding. By contrast, when both ClpB and the DnaK system are present together, the aggregated proteins can be dissolved and the proteins reactivated (15-17, 104). Mechanistic studies of this process suggest that ClpB binds protein aggregates directly and, through

ATP-induced conformational changes in its hexameric structure, increases the exposure of their hydrophobic surfaces, thereby allowing DnaK/DnaJ/GrpE to bind and mediate dissociation and refolding of solubilized polypeptides into native proteins. The mechanism by which ClpB accomplishes this task remains unclear, as is the question of whether ClpB increases the availability of proteins in the aggregate by unmasking DnaK/DnaJ binding sites or by transiently removing these from the aggregate. Based upon the mechanism of the highly homologous ClpA protein, it seems likely that it actually removes proteins from these aggregates and completely unfolds them; however, in the absence of DnaK/DnaJ, these solubilized proteins cannot refold.

Cooperation Between ClpB and the DnaK-System In Recovery of Misfolded and Aggregated Proteins

The importance of the DnaK/ClpB bi-chaperone network to the *in vivo* recovery of misfolded/aggregated proteins has recently been demonstrated (10, 17). Heat treatment of total soluble extracts for 15 min at 45°C caused the denaturation of 10-15% of the proteins despite the presence of endogenous chaperones. Addition of the DnaK system to these extracts prior to heat shock protected the proteins from aggregation in a concentration-dependent manner. However, addition of four other *E. coli* chaperones (GroEL/GroES, HtpG, ClpB, IbpA) did not prevent aggregation. Examination of the proteins protected by the DnaK system revealed ~250 protein species corresponding to ~30% of all the proteins detectable by 2-D gel electrophoresis. Similarly, *in vivo* experiments with *E. coli* knock-out mutations in various chaperone genes (Δ*clpB::kan*, Δ*htpG::lacZ*, Δ*ibpA::kan*, Δ*dnaK52::cat*) confirmed the preeminence of the DnaK system in preventing protein aggregation during heat-shock treatment (10). Heat shock at 42°C caused strong protein aggregation in Δ*dnaK52::cat* mutants, but no

aggregation in wild type or any of the other mutant cell lines tested. To identify the thermolabile DnaK substrates, mass spectrometry of the insoluble proteins found in the heat-shocked Δ*dnaK52::cat* mutant cells revealed that 80% of the large proteins (>90 kDa), but only 18% of the small (<30 kDa) cytoplasmic proteins were DnaK substrates, including several essential proteins. Interestingly, protein aggregation in Δ*dnaK52::cat* mutants could be reversed by induced synthesis of ClpB and the DnaK system, paralleling the *in vitro* studies indicating that ClpB and the DnaK system cooperate to dissolve and refold protein aggregates.

Figure 6 shows one model for the action of the DnaK/ClpB chaperone network. In addition to the DnaK system and ClpB, at least one additional chaperone system may be involved in protecting proteins from aggregation. Analysis of the protein content of inclusion bodies that result from the high-level expression of some non-foldable proteins in *E. coli* revealed that, in addition to the overexpressed protein, these aggregates were enriched in two small cellular proteins, termed IbpA and IbpB, that are members of the small heat-shock family of chaperones (81, 83). *In vitro,* these proteins prevent aggregation and may cooperate with the DnaK system as holding chaperones preventing aggregation prior to reactivation by the DnaK system.

ATP-DEPENDENT DEGRADATION OF DAMAGED PROTEINS

Despite the considerable effort that cells expend in preventing misfolding and aggregation of nascent polypeptides during *de novo* folding and in refolding the majority of misfolded and/or aggregated proteins, some of these proteins cannot be rescued. In addition, a large fraction (~20% in bacteria) of nascent chains are degraded immediately after synthesis (105, 106). This fraction rises to ~30% in eukaryotes (107), and if the polypeptide contains an amino-terminal degradation signal, more

than 50% of these are degraded co-translationally, never reaching their mature size before destruction (108). A fraction of the rapidly-degraded polypeptides are probably damaged due to errors in transcription or translation; however, based upon the known fidelity of these processes (1/10000 amino acids incorporated), these errors can only account for a small fraction of the total (109). In some cases, this rapid turnover may be an essential component of a regulatory pathway. For example, the quorum-sensing transcriptional regulator TraR cannot fold and is rapidly degraded in the absence of its cognate signaling ligand (110). In fact, this protein only binds its inducing ligand during its synthesis. The continuous synthesis and degradation of TraR, until its cognate ligand is present, provides an extremely rapid means of communication between bacterial cells. In both prokaryotic and eukaryotic cells the vast majority of the rapidly-degraded proteins, including newly-synthesized polypeptides, and damaged, misfolded or aggregated proteins, requires ATP hydrolysis and is conducted by large ATP-dependent proteolytic complexes (111, 112). As is the case for many of the chaperones, failure of these ATP-dependent proteases has been implicated in several human disease states (113). Interestingly, all of the ATP-dependent proteolytic complexes show a substantial degree of architectural similarity with the GroEL-like chaperonins, and overlap functionally with a wide range of chaperones (67, 114, 115). In particular, the bacterial ATP-dependent proteases use similar principals with both the GroEL and DnaK chaperones for selecting substrates to be degraded (90). This is not entirely unexpected since both chaperones and ATP-dependent proteases must recognize similar non-native substrates. In fact, recent studies have shown that the ATPase components of proteases can also function as molecular chaperones (116, 117). Together, these observations suggest a kinetic model for the fate of nascent and marginally stable polypeptides in cells which partition between the networks of chaperones that promote folding/refolding and the chaperone-like ATP-dependent proteases that degrade them. At any particular time, the cellular concentration of these proteins depends upon a combination of their rate of synthesis

and relative flux through the folding and degradation pathways. One important implication of this model is that, for a given protein, its cellular concentration is not necessarily coupled to its mRNA level, and thus attempts to estimate changes in the cellular concentration of individual proteins based upon changes in mRNA levels are necessarily flawed.

Four Major Families of Eubacterial Proteases

In Eubacteria, greater than 90% of cytoplasmic proteolysis is carried out by energy-dependent proteases, the most common of which are Lon, FtsH, Clp and HslUV (118). Lon and the integral membrane protease, FtsH, are homooligomeric complexes in which the ATPase and protease activities reside in a single polypeptide chain. By contrast, the Clp and HslUV proteases are heterooligomeric complexes in which the ATPase and peptidase subunits form independently-stable oligomers that associate in an ATP-dependent fashion during substrate degradation (Figure 4, and Figure 7) (119-121). Although each of these four enzymes represents a unique proteolytic system, they all appear to be architecturally and mechanistically related at the most basic level (122). For example, in their active complexes, all of the proteases form large barrel-like structures that bind, unfold and engulf substrates, utilizing homologous ATPase domains of the AAA (ATPases Associated with various cellular Activities) family of ATPases (123-125).

Moreover, although each of these proteases has a unique proteolytic core, they all degrade substrates processively to produce 5-15 residue peptides by cleaving between virtually any non-proline, non-glycine bond (118, 120). Since there is a strong architectural and functional conservation between the various proteases, we will only briefly describe the Lon and FtsH proteases, and focus our mechanistic discussions on the better-characterized Clp and HslU proteases.

The *E. coli lon* gene encodes a nonessential 88 kDa protein, the Lon protease, which probably forms a homo-hexameric complex, although the exact stoichiometry of the complex is still under debate (126, 127). Lon is a serine protease with its catalytic residues located in the carboxy-terminal half of the amino acid sequence (128). Limited proteolysis studies demonstrate that the proteolytic domain is an independently-stable folding unit that catalyzes the peptide bond hydrolysis of short peptide substrates in the absence of its ATPase domain but cannot degrade protein substrates. The amino-terminal half of Lon contains the AAA-ATPase domain, which is responsible for substrate binding, unfolding and translocation to the proteolytic active sites (129). Based upon the sequence homology between the various AAA-ATPases, the structure of this domain is likely to resemble that of the ATPase in the Clp (ClpA, ClpX) and HslUV (HslU) proteases, and also that in the more distantly-related FtsH protease (123-125, 130). These observations provide further support for the notion that Lon is architecturally and functionally similar to the two-component Clp and HslUV proteases.

The *ftsH* gene encodes a 71 kDa ATP-dependent protease, FtsH, that unlike the other three proteases encodes an integral membrane protein, although both the proteolytic and ATPase activities reside within cytosolically-exposed domains (130). The 200-residue ATPase domain also belongs to the AAA-ATPase superfamily (130, 131). The proteolytic domain, which resides near the carboxy-terminus, is a zinc metalloprotease (130). Homologues of FtsH are found in the mitochondria of most eukaryotes and are thought to play a critical role in protein degradation in these organelles (132, 133).

The *hslUV* operon, found in many eubacteria, encodes two proteins, HslU, a 50 kDa ATPase, and HslV, a 20 kDa threonine protease (120, 134, 135). The functional HslUV protease is comprised of two HslU hexamers bound at each end of the dodecameric HslV protease (two stacked rings of hexameric subunits) giving rise to a rotationally symmetric complex (136). The high-resolution structures of HslU and HslV have been recently determined alone and in their complex (Figure 7) (137-

139). The HslU ATPase adopts a barrel-like structure with the central core of the oligomer formed by the association of the AAA-ATPase domains. Similarly, the proteolytic component, HslV, also has a barrel-like structure comprised of subunits that share a common fold with the β-subunits in the archaebacterial and eukaryotic 20S proteasome protease. Assembly of these subunits into the functional dodecamer encloses active sites in a central chamber that is ~50 Å in diameter. Access to this chamber is restricted to two ~14Å axial pores that limit access of proteins to those targeted for degradation by the HslU ATPase. Interestingly, unlike the β-subunits of the proteasome, which form heptameric rings, subtle differences in the subunit packing result in HslV forming hexameric rings (140). Peptide bond hydrolysis by most HslV homologues goes by an Ntn-hydrolytic mechanism (N-terminal threonine nucleophile), which was first described for the β-subunits of the proteasome. However, recently at least one HslV homologue has been considered to use an N-terminal serine to catalyze this reaction (141). This observation underlies the fact that the details of peptide bond hydrolysis by these proteases do not play a significant role in degradation of protein substrates. The self-compartmentalization of the proteolytic active sites within a central chamber in the HslV oligomer is thought to be a common feature of all four bacterial ATP-dependent proteases (142). This feature prevents the degradation of substrates not specifically targeted by the proteolytic subunits. Association of HslU and HslV to form the active protease requires ATP, but not its hydrolysis, and involves interactions between residues near the C-terminus of HslU with the outer surface of the HslV cylinder (139). In the complex, HslU hexamers are axially stacked at each end of the HslV dodecamer. In this arrangement there is a small central channel (~5 Å minimum diameter) that passes through the HslU hexamer into the HslV proteolytic chamber. Degradation of folded substrates by HslUV requires ATP hydrolysis; however, unfolded model substrates can be degraded under some circumstances in the presence of the non-hydrolyzable analog AMPPNP (143). Currently, little is known about the *in vivo* substrates of HslUV.

The mechanistically best-characterized of the four proto-typical eubacterial ATP-dependent proteases is Clp (144). The functional Clp protease is comprised of ClpP in complex with either of two ATPase components, ClpX or ClpA [in some bacteria, and plant chloroplasts, ClpA is replaced with a highly homologous ATPase ClpC that also binds ClpP (145)]. The ATPases differ structurally in that ClpA/ClpC, like the ClpB chaperone, contains two nucleotide-binding domains (NBD's), while ClpX, like HslU, has only one (65) (Figure 4). The functional significance of this is not known, since both ATPases support the degradation of a range of substrates. More importantly, though, ClpA/C and ClpX target different substrates for degradation, and are thus alternative specificity factors for the protease (146). Currently there is a high-resolution structure of *E. coli* ClpP (147, 148), and homology models for ClpX and ClpA based upon the structure of the HslU ATPase (Figure 7). The crystal structure of ClpP reveals a subunit fold that is quite distinct from that of HslV; however, the overall architecture of the ClpP and HslV complexes is strikingly similar. In common with the HslUV protease, association of ClpP with either of its specificity ATPases requires ATP binding but not hydrolysis, to form complexes that have ClpA/C or ClpX oligomers stacked at each end of the central ClpP oligomer (64, 149-152). Interestingly, this complex is rotationally asymmetric since both ClpA/C and ClpX are hexamers while ClpP is comprised of two heptameric rings producing a 6-7 mismatch (64, 147, 151, 153). At present, the functional significance of the asymmetry in the Clp protease is not clear; however, a similar 6-7 asymmetry is thought to occur in the eukaryotic 26S proteasome.

Conserved Mechanism of ATP-Dependent Protein Degradation

In the last several years, the basic features of the proteolytic and chaperone mechanism of these ATP-dependent proteases have come into focus. These studies have largely exploited native substrates

that are recognized and either remodeled or degraded in their roles as regulators of specific biochemical pathways rather than in their role in protein QC (68-71, 116, 117, 154-161). However, at the level of mechanistic detail currently available, the consensus mechanism provides a suitable starting point for understanding how these proteins function in QC. This model involves five main steps, as shown in Figure 8. Initially, a substrate targeted for degradation binds at one end of the ATPase. The bound substrates are unfolded in a reaction that requires ATP hydrolysis. Additional rounds of ATP hydrolysis are then used to translocate the unfolded substrate through the barrel-like ATPase complex into a chamber formed by the protease component, where it is degraded. Finally, the peptide products are discharged, either passively through diffusion, or actively, out of the chamber to complete the cycle. The penultimate step of peptide-bond cleavage does not require energy input and is accomplished entirely by the active sites in the proteolytic component (162, 163).

At least one or more specific domains in the ATPases recognize substrates destined for degradation. The sensor and substrate discrimination (SSD) domain is one such domain that has been implicated in substrate recognition by the Lon, Clp and HslUV proteases and it is located at the C-terminus of their respective ATPase domains (164, 165). Interestingly, a homologous domain is found at the C-terminus of the ClpB chaperone and thus this domain may act as a link between the chaperone activity of ClpB and the energy-dependent degradation activity of the Lon, Clp, and HslU ATPases. In the structure of the HslUV complex, the SSD domain lies at the periphery of the ATPase near the HslUV interface (139, 166). It is expected to be in a similar position in the structures of the ClpA/CP and ClpXP complexes. This is somewhat surprising since, in the low-resolution electron microscopic model for substrate-bound ClpXP, the substrate is initially bound at the distal end of the ClpX in the ClpXP complex. A similar location was found for a substrate bound to ClpA in the ClpAP complex: in this case, the substrate was over 100 Å away from the SSD domain of ClpA. These results imply that

substrate targeting involves at least one additional domain (157, 159). Candidates for this function are the amino-terminal domains in ClpA and ClpX, and the central coiled-coil (I-domain) in HslU. These three domains are thought to occupy analogous positions in the three-dimensional structure of the ATPases, although they are found in different positions in the linear sequence. In the structural model of HslU, the I-domain is located at the distal end of HslU, opposite the HslV-binding face (139, 161, 167, 168). Based upon modeling studies, the amino-terminal domain of ClpA occupies a similar position with respect to its ClpP-binding surface, as does the nonhomologous amino-terminal domain of ClpX. These three putative substrate-binding domains do not share any sequence homology and thus appear to be good candidates for defining, at least in part, the substrate specificity of the individual proteases. However, several recent studies suggest that although these domains are required for the binding of some substrates, they are dispensable in the degradation of others, and thus the details of how these proteases target substrates remains an open question (161, 167, 168). By contrast, evidence is accumulating that one of the features in the substrate that both the Clp and Lon proteases recognize are hydrophobic patches adjacent to basic residues (154, 162, 165, 169). The SSD domain is probably a key determinant in recognizing this motif (164, 165, 170). Interestingly, this motif resembles that recognized by the DnaK/DnaJ chaperones, described above, and may represent another link between the chaperones and ATP-dependent proteases. One plausible role for substrate recognition by the SSD domain is in discriminating between unfolded polypeptides that may still be capable of folding and those that are no longer folding competent. In this model, the SSD domain senses the degree of unfolding; only highly extended substrates will be capable of interacting with the SSD domain, and thus triggering degradation.

Once stably bound to the ATPases, the substrates are completely unfolded, preparing them for translocation and entry into the proteolytic chamber. Unfolding is also likely to be important even with

damaged or misfolded proteins, since the substrate-binding chamber of the ATPase subcomplex and the proteolytic chamber of the protease complex are linked by channels that are between 5-13 Å in diameter at their narrowest points, placing severe constraints on the conformation of substrate proteins that can be translocated into the proteolytic core (68, 70, 71, 116, 171, 172). Mechanistic studies following the degradation of folded substrates by ClpXP and ClpAP indicate that folding requires ATP hydrolysis and is the rate-limiting step in their degradation (158). At present, little is known about how these ATPases couple ATP-hydrolysis to substrate unfolding for two main reasons: there is a high background rate of ATP hydrolysis and coupling between ATP hydrolysis and their chaperone-like activities is weak at best (144, 149, 173, 174). Thus, it is not clear whether unfolding requires the concerted hydrolysis of ATP by multiple subunits, as in the GroEL reaction cycle, or involves sequential or random hydrolysis. In either case, recent evidence suggests that unfolding involves the peeling away of segments of the polypeptide from the native structure of the bound substrate (175, 176). For substrates that have degradation tags at either the N- or C-terminal ends, segments of the polypeptide chain near the tag are peeled first and are then funneled into the central proteolytic chamber for degradation (176). One interesting aspect of this mechanism is that the susceptibility of a substrate to unfolding is determined not by its global stability but rather by the stability of structural elements near the degradation tag. Global unfolding occurs after this local structural element is peeled away. In essence, this mechanism parallels that proposed to explain the ability of DnaK to maintain bound substrates in an unfolded conformation, as discussed above.

Translocation of substrates from their initial binding sites at the ends of the ATPase subunits or domains into the proteolytic core also involves ATP hydrolysis. In the ClpAP protease, substrates are translocated ~150 Å before entering ClpP by a process that does not involve any major structural rearrangement in the ATPase component, as judged by image reconstruction of electron micrographs

(70, 159). This latter fact imposes extreme topological constraints upon the mechanism of this process. In addition, it is also unclear whether translocation requires the entire protein to be unfolded, or can begin even as the protein is unfolding. Furthermore, it is not obvious how binding of the ATPase to the degradation tag promotes substrate insertion into the translocation channel. However, recent biochemical experiments have shed some light on the vectorial nature of this process. For multidomain proteins, Lee et al. have shown that degradation appears to be sequential, using multidomain fusion proteins tagged with a degradation signal (176). In the case of an N-terminal tag fused to a DHFR (dihydrofolate reductase) module followed by a barnase module, degradation by ClpAP could be prevented by methotrexate, a small molecule that binds and stabilizes DHFR. When the positions of DHFR and barnase modules were reversed, the barnase module was selectively degraded even when methotrexate was present. By contrast, the DHFR module in this construct was not degraded; instead it was released as a stable DHFR-containing fragment. The vectorial translocation of a bound substrate starting from a degradation tag has also recently been demonstrated directly (69). In these experiments a fluorophore was covalently attached at either end of a substrate for the ClpAP protease. Translocation was monitored by FRET (Fluorescence Resonance Energy Transfer) from a donor fluorophore covalently attached in the catalytic chamber of ClpP to the acceptor fluorophore attached to the translocating substrate. Energy transfer occurs when the two fluorescent probes are brought in close proximity. In these experiments, energy transfer occurs 2-4 sec sooner with the substrate labeled near the degradation tag compared with that labeled at the opposite end of the polypeptide sequence (69). As with the experiments examining protein unfolding, translocation experiments exploited native substrates targeted for degradation by specific sequence tags. For most of the damaged or misfolded protein targets of these ATP-dependent proteases, there are no specific sequence tags and their unfolding and initiation of translocation may proceed by subtly different pathways.

ClpX and ClpA in their role as molecular chaperones are likely to exploit similar, if not identical, mechanisms for recognizing, unfolding and translocating substrates (68, 71, 176, 177). Based upon the structure of HslU, and the constraints imposed on substrate entry into ClpP (147, 166), the substrates for the chaperone function of these proteins are likely to be ejected from the ATPase in an extended state. A major question remains: are the proposed chaperone functions an artifact due to the absence of ClpP in these *in vitro* studies? A partial answer can be reasoned in the following way. Under normal bacterial growth conditions, the *clpA, clpP* and *clpX* genes are all highly expressed. In a significant fraction of known bacterial species the genes for *clpP* and *clpX* are stress inducible and adjacent to one another on the chromosome. In contrast, the *clpA* gene is normally found in a distinct chromosomal locus and is not regulated by stress (178). This implies that the cellular concentrations of ClpA and ClpX are not coordinated and may exceed the total cellular ClpP concentration. Since both ClpA and ClpX bind to ClpP with similar affinities (152), both free ClpA and ClpX may be available under certain cellular conditions to perform a chaperone role *in vivo*.

Degradation of the translocated substrate occurs entirely within the central proteolytic chamber. Substrates enter this chamber in a largely extended conformation (147, 148) in which the peptide bonds to be cleaved are freely accessible and are degraded in a processive manner: once the substrate is committed to turnover by these systems, degradation proceeds to completion producing short 3-15 residue peptide products (144). Due to the architecture of the proteolytic chamber, the active sites exist at very high concentrations (~500 mM), as are the peptide bonds in actively-translocated polypeptide substrates (147, 148). In fact, substrates as large as ~30 kDa can be completely translocated into the proteolytic chamber of a catalytically-defective ClpP oligomer (70). This high concentration ensures the complete degradation of polypeptide substrates. The tightly-defined size distribution of products was initially thought to result from the concerted cleavage of peptide by adjacent active sites, a mechanism

termed the "molecular ruler" model (179). More recently, it has become clear that the size distribution of products during protein degradation is actually defined by two competing factors: the overall rate of peptide bond hydrolysis that generates the peptide fragments, and the rate that these products can diffuse out of the access pores (147, 148, 180, 181). This mechanism has been termed the "molecular sieve" model since the access pores act like molecular sieves to control the sizes of peptides that can diffuse freely into or out of the proteolytic chamber (147, 148).

Evidence of an *in vivo* Role of ATP-Dependent Proteases in QC

Over the past ~15 years, a substantial body of evidence has accumulated implicating ATP-dependent proteases in protein QC *in vivo* (90, 114, 182, 183). Initial experiments with puromycin-treated *E. coli* cells indicated a role for the Lon and Clp proteases in degrading damaged proteins (105, 184, 185). In addition, numerous studies have defined a role for these proteins in degrading misfolded proteins under normal and heat-shock growth conditions (10, 186). For example, in *clpC* and *clpP* mutant *B. subtilis* strains, large electron-dense protein aggregates accumulate under both stressed and non-stressed conditions, suggesting a direct role for the Clp protease in removing misfolded proteins (186). Moreover, in a wild-type strain subjected to heat-shock antibodies raised against ClpC, ClpP and ClpX were localized in these aggregates. These results show that in *B. subtilis* the Clp proteins play an important role in degrading misfolded proteins *in vivo*. However, this should not be taken to mean that the Clp proteins are the major ATP-dependent proteases for degrading misfolded proteins in all eubacteria, since in other bacteria Lon appears to play an important role in this process (111). More recently, the ClpXP protease has been implicated in a novel ribosome rescue pathway that frees stalled ribosomes on a damaged mRNA (187). mRNAs that lack stop codons interfere with the ribosomal

termination and reinitiation cycle, giving rise to truncated polypeptides. ClpXP participates with a novel tRNA-mRNA hybrid, known alternatively as SsrA, tmRNA or 10Sa RNA, that acts both as a tRNA and an mRNA to direct the non-template encoded addition of an 11-residue degradation tag (187-190). Proteins synthesized in this way are subsequently degraded by the ClpXP protease. This QC mechanism ensures that ribosomes are not sequestered on terminator-less mRNAs and that prematurely truncated proteins do not accumulate. It may also contribute to recycling of ribosomes stalled for other reasons.

Synergistic Action of Chaperones and Proteases in Degrading Damaged Polypeptides

Several studies over the last few years have suggested a direct relationship between the chaperones DnaK/DnaJ and GroEL and ATP-dependent protein degradation (Figure 9) (10, 191-197). Among the first examples of the involvement of DnaJ/DnaK in the degradation of specific abnormal proteins was a variant of alkaline phosphatase (PhoA61), which is not secreted from the cytosol due to a missense mutation in its signal sequence (196, 198). Its rapid degradation under normal growth conditions is mediated in part by Lon and requires DnaK and DnaJ. In addition, the dnaK756 mutation, which slows the release of substrates bound to DnaK, enhanced PhoA61 degradation, whereas a *dnaJ* mutant that reduces PhoA61 association with DnaK slowed its turnover. These results suggest that prolonged association with the DnaK/DnaJ chaperone system promotes degradation. More recently, studies have been conducted to characterize the role of DnaK depletion on global ATP-dependent protein turnover in bacteria (10, 17). The results of this study suggest that DnaK is not an essential factor for ATP-dependent degradation of misfolded proteins under normal growth conditions, but under heat-shock conditions DnaK and the ATP-dependent proteases act synergistically to remove damaged proteins.

HOMOLOGOUS SYSTEMS FOR PROTEIN QUALITY CONTROL IN EUKARYOTES

The problem of protein QC is not unique to bacteria; in fact, eukaryotic cells are posed with the even more daunting task of managing protein folding in multiple compartments. For the most part, the protein QC systems in eukaryotic cells are analogous to the bacterial systems. For example, in the eukaryotic cytoplasm, there are multiple Hsp70 systems involved in *de novo* folding, recovery of misfolded aggregated proteins, and in the ATP-dependent degradation of misfolded proteins. Similarly, the Group II chaperonin TRiC plays an equivalent role to GroEL in the folding of some proteins (94), and the 26S proteasome plays the corresponding role of the bacterial ATP-dependent proteases in protein degradation (199). Moreover, protein QC in mitochondria is highly homologous to that in bacteria (200-207). The one major exception is the protein QC system found in the endoplasmic reticulum. This organelle is responsible for folding proteins in the secretory pathway. Protein QC in this organelle involves N-linked oligosaccharides, which are used by lectin-specific chaperones and modifying enzymes in *de novo* protein folding or in targeting misfolded proteins for degradation (208, 209).

CLOSING REMARKS

Presently, our understanding of the mechanisms regulating the expression, cellular content and intracellular distribution of protein refolding and degradation machinery is somewhat limited. The exquisite synergy between refolding and recycling is becoming clearer. However, the observations reviewed here indicate that although significant progress has been made recently in elucidating their mechanisms, many more questions still remain unanswered. The dramatic recent progress in elucidating

the structures and functions of many QC machines should facilitate a deeper understanding of these regulatory mechanisms in the next decade.

ACKNOWLEDGMENTS

We wish to thank Vito Graziano and Kathleen Griffin for their assistance in the preparation of this manuscript. In addition, we would like to thank the Editor for her unwavering adherence to the rules of English grammar, such as they are, and in particular always to unsplitting the infinitive. This work was supported by grants GM57390 and GM53181 from the NIH, and the Office of Biological and Environmental Research of the US Department of Energy under Prime Contract DE-AC0298-CH10886 with Brookhaven National Laboratory.

REFERENCES

1 Zimmerman, S. B. and Trach, S. O. (1991) J. Mol. Biol. 222, 599-620.

2 Ellis, R. J. and Hartl, F. U. (1999) Curr. Opin. Struct. Biol. 9, 102-110.

3 Ellis, R. J. (1997) Curr. Biol. 7, R531-533.

4 Feldman, D. E. and Frydman, J. (2000) Curr. Opin. Struct. Biol. 10, 26-33.

5 Slavotinek, A. M. and Biesecker, L. G. (2001) Trends Genet. 17, 528-535.

6 Macario, A. J. and Conway de Macario, E. (2000) Int. J. Clin. Lab. Res. 30, 49-66.

7 Carmichael, J., Chatellier, J., Woolfson, A., Milstein, C., Fersht, A. R. and Rubinsztein, D. C. (2000) Proc. Nat. Acad. Sci. U.S.A. 97, 9701-9705.

8 Henderson, B., Nair, S. P. and Coates, A. R. (1996) Inflamm. Res. 45, 155-158.

9 Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A. and Bukau, B. (1999) Nature 400, 693-696.

10 Tomoyasu, T., Mogk, A., Langen, H., Goloubinoff, P. and Bukau, B. (2001) Mol. Microbiol. 40, 397-413.

11 Teter, S. A., Houry, W. A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C. and Hartl, F. U. (1999) Cell 97, 755-765.

12 Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, I. N. and Furtak, K. (1993) Cell 74, 909-917.

13 Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F. and Hartl, F. U. (1999) Nature 402, 147-154.

14 Bukau, B. and Horwich, A. L. (1998) Cell 92, 351-366.

15 Diamant, S., Ben-Zvi, A. P., Bukau, B. and Goloubinoff, P. (2000) J. Biol. Chem. 275, 21107-21113.

16 Goloubinoff, P., Mogk, A., Zvi, A. P., Tomoyasu, T. and Bukau, B. (1999) Proc. Nat. Acad. Sci. U.S.A. 96, 13732-13737.

17 Mogk, A., Tomoyasu, T., Goloubinoff, P., Rudiger, S., Roder, D., Langen, H. and Bukau, B. (1999) EMBO J. 18, 6934-6949.

18 Zolkiewski, M., Kessel, M., Ginsburg, A. and Maurizi, M. R. (1999) Protein Sci. 8, 1899-1903.

19 Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C. and Zylicz, M. (1991) Proc. Nat. Acad. Sci. U.S.A. 88, 2874-2878.

20 Seaton, B. L. and Vickery, L. E. (1994) Proc. Nat. Acad. Sci. U.S.A. 91, 2066-2070.

21 Itoh, T., Matsuda, H. and Mori, H. (1999) DNA Res. 6, 299-305.

22 Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M. K. and Hartl, F. U. (1992) Nature 356, 683-689.

23 Schroder, H., Langer, T., Hartl, F. U. and Bukau, B. (1993) EMBO J. 12, 4137-4144.

24 Szabo, A., Langer, T., Schroder, H., Flanagan, J., Bukau, B. and Hartl, F. U. (1994) Proc. Nat. Acad. Sci. U.S.A. 91, 10345-10349.

25 Schonfeld, H. J.; Schmidt, D., Schroder, H. and Bukau, B. (1995) J. Biol. Chem. 270, 2183-2189.

26 Ueguchi, C., Shiozawa, T., Kakeda, M., Yamada, H. and Mizuno, T. (1995) J. Bacteriol. 177, 3894-3896.

27 Wegrzyn, A., Taylor, K. and Wegrzyn, G. (1996) J. Bacteriol. 178, 5847-5849.

28 Genevaux, P., Wawrzynow, A., Zylicz, M., Georgopoulos, C. and Kelley, W. L. (2001) J. Biol. Chem. 276, 7906-7912.

29 Szabo, A., Korszun, R., Hartl, F. U. and Flanagan, J. (1996) EMBO J. 15, 408-417.

30 Brehmer, D., Rudiger, S., Gassler, C. S., Klostermeier, D., Packschies, L., Reinstein, J., Mayer, M. P. and Bukau, B. (2001) Nature Struct. Biol. 8, 427-432.

31 Choglay, A. A., Chapple, J. P., Blatch, G. L. and Cheetham, M. E. (2001) Gene 267, 125-134.

32 Naylor, D. J., Stines, A. P., Hoogenraad, N. J. and Hoj, P. B. (1998) J. Biol. Chem. 273, 21169-21177.

33 Schlicher, T. and Soll, J. (1997) Plant Mol. Biol. 33, 181-185.

34 Macario, A. J., Simon, V. H. and Conway de Macario, E. (1995) Biochim. Biophys. Acta 1264, 173-177.

35 Voos, W., Gambill, B. D., Laloraya, S., Ang, D., Craig, E. A. and Pfanner, N. (1994) Mol Cell Biol 14, 6627-6634.

36 Ikeda, E., Yoshida, S., Mitsuzawa, H., Uno, I. and Toh-e, A. (1994) FEBS Lett. 339, 265-268.

37 Gragerov, A., Zeng, L., Zhao, X., Burkholder, W. and Gottesman, M. E. (1994) J. Mol. Biol. 235, 848-854.

38 Rudiger, S., Germeroth, L., Schneider-Mergener, J. and Bukau, B. (1997) EMBO J. 16, 1501-1507.

39 Flanagan, J. M., Kataoka, M., Shortle, D. and Engelman, D. M. (1992) Proc. Nat. Acad. Sci. U.S.A. 89, 748-752.

40 Shortle, D. and Abeygunawardana, C. (1993) Structure 1, 121-134.

41 Laufen, T., Mayer, M. P., Beisel, C., Klostermeier, D., Mogk, A., Reinstein, J. and Bukau, B. (1999) Proc. Nat. Acad. Sci. U.S.A. 96, 5452-5457.

42 Suh, W. C., Lu, C. Z. and Gross, C. A. (1999) J. Biol. Chem. 274, 30534-30539.

43 Karzai, A. W. and McMacken, R. (1996) J. Biol. Chem. 271, 11236-11246.

44 Suh, W. C., Burkholder, W. F., Lu, C. Z., Zhao, X., Gottesman, M. E. and Gross, C. A. (1998) Proc. Nat. Acad. Sci. U.S.A. 95, 15223-15228.

45 Rudiger, S., Schneider-Mergener, J. and Bukau, B. (2001) EMBO J. 20, 1042-1050.

46 Martinez-Yamout, M., Legge, G. B., Zhang, O., Wright, P. E. and Dyson, H. J. (2000) J. Mol. Biol. 300, 805-818.

47 Sha, B., Lee, S. and Cyr, D. M. (2000) Structure Fold Des. 8, 799-807.

48 McCarty, J. S., Buchberger, A., Reinstein, J. and Bukau, B. (1995) J. Mol. Biol. 249, 126-137.

49 Theyssen, H., Schuster, H. P., Packschies, L., Bukau, B. and Reinstein, J. (1996) J. Mol. Biol. 263, 657-670.

50 Packschies, L., Theyssen, H., Buchberger, A., Bukau, B., Goody, R. S. and Reinstein, J. (1997) Biochemistry 36, 3417-3422.

51 Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. and Sigler, P. B. (1994) Nature 371, 578-586.

52 Sigler, P. B. and Horwich, A. L. (1995) Philos. Trans. Roy. Soc. Lond. B Biol. Sci. 348, 113-119.

53 Boisvert, D. C., Wang, J., Otwinowski, Z., Horwich, A. L. and Sigler, P. B. (1996) Nature Struct. Biol. 3, 170-177.

54 Mayhew, M. and Hartl, F. U. (1996) Science 271, 161-162.

55 Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L. and Deisenhofer, J. (1996) Nature 379, 37-45.

56 Xu, Z., Horwich, A. L. and Sigler, P. B. (1997) Nature 388, 741-750.

57 Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A. and Horwich, A. L. (1998) Annu. Rev. Biochem. 67, 581-608.

58 Fenton, W. A., Weissman, J. S. and Horwich, A. L. (1996) Chem. Biol. 3, 157-161.

59 Farr, G. W., Furtak, K., Rowland, M. B., Ranson, N. A., Saibil, H. R., Kirchhausen, T. and Horwich, A. L. (2000) Cell 100, 561-573.

60 Weissman, J. S., Hohl, C. M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H. R., Fenton, W. A. and Horwich, A. L. (1995) Cell 83, 577-587.

61 Fenton, W. A., Kashi, Y., Furtak, K. and Horwich, A. L. (1994) Nature 371, 614-619.

62 Parsell, D. A., Kowal, A. S. and Lindquist, S. (1994) J. Biol. Chem. 269, 4480-4487.

63 Kim, K. I., Cheong, G. W., Park, S. C., Ha, J. S., Woo, K. M., Choi, S. J. and Chung, C. H. (2000) J. Mol. Biol. 303, 655-666.

64 Beuron, F., Maurizi, M. R., Belnap, D. M., Kocsis, E., Booy, F. P., Kessel, M. and Steven, A. C. (1998) J. Struct. Biol. 123, 248-259.

65 Schirmer, E. C., Glover, J. R., Singer, M. A. and Lindquist, S. (1996) Trends Biochem. Sci. 21, 289-296.

66 Woo, K. M., Kim, K. I., Goldberg, A. L., Ha, D. B. and Chung, C. H. (1992) J. Biol. Chem. 267, 20429-20434.

67 Horwich, A. L. (1995) Curr. Biol. 5, 455-458.

68 Weber-Ban, E. U., Reid, B. G., Miranker, A. D. and Horwich, A. L. (1999) Nature 401, 90-93.

69 Reid, B. G., Fenton, W. A., Horwich, A. L. and Weber-Ban, E. U. (2001) Proc. Nat. Acad. Sci. U.S.A. 98, 3768-3772.

70 Singh, S. K., Grimaud, R., Hoskins, J. R., Wickner, S. and Maurizi, M. R. (2000) Proc. Nat. Acad. Sci. U.S.A. 97, 8898-8903.

71 Hoskins, J. R., Singh, S. K., Maurizi, M. R. and Wickner, S. (2000) Proc. Nat. Acad. Sci. U.S.A. 97, 8892-8897.

72 Hesterkamp, T. and Bukau, B. (1996) FEBS Lett. 389, 32-34.

73 Hesterkamp, T., Hauser, S., Lutcke, H. and Bukau, B. (1996) Proc. Nat. Acad. Sci. U.S.A. 93, 4437-4441.

74 Hesterkamp, T., Deuerling, E. and Bukau, B. (1997) J. Biol. Chem. 272, 21865-21871.

75 Thomas, J. G. and Baneyx, F. (2000) Mol. Microbiol. 36, 1360-1370.

76 Nemoto, T. K., Ono, T. and Tanaka, K. (2001) Biochem. J. 354, 663-670.

77 Spence, J. and Georgopoulos, C. (1989) J. Biol. Chem. 264, 4398-4403.

78 Young, J. C., Schneider, C. and Hartl, F. U. (1997) FEBS Lett. 418, 139-143.

79 Dittmar, K. D., Banach, M., Galigniana, M. D. and Pratt, W. B. (1998) J. Biol. Chem. 273, 7358-7366.

80 Allen, S. P., Polazzi, J. O., Gierse, J. K. and Easton, A. M. (1992) J. Bacteriol. 174, 6938-6947.

81 Laskowska, E., Wawrzynow, A. and Taylor, A. (1996) Biochimie 78, 117-122.

82 Lund, P. A. (2001) Adv. Microbial Physiol. 44, 93-140.

83 Thomas, J. G. and Baneyx, F. (1998) J. Bacteriol. 180, 5165-5172.

84 Kitagawa, M., Matsumura, Y. and Tsuchido, T. (2000) FEMS Microbiol. Lett. 184, 165-171.

85 Kim, R., Kim, K. K., Yokota, H. and Kim, S. H. (1998) Proc. Nat. Acad. Sci. U.S.A. 95, 9129-9133.

86 Kim, K. K., Kim, R. and Kim, S. H. (1998) Nature 394, 595-599.

87 Yang, H., Huang, S., Dai, H., Gong, Y., Zheng, C. and Chang, Z. (1999) Protein Sci. 8, 174-179.

88 Fenton, W. A. and Horwich, A. L. (1997) Protein Sci. 6, 743-760.

89 Frydman, J. and Hartl, F. U. (1996) Science 272, 1497-1502.

90 Wickner, S., Maurizi, M. R. and Gottesman, S. (1999) Science 286, 1888-1893.

91 Hesterkamp, T. and Bukau, B. (1998) EMBO J. 17, 4818-4828.

92 Netzer, W. J. and Hartl, F. U. (1998) Trends Biochem. Sci. 23, 68-73.

93 Frydman, J., Erdjument-Bromage, H., Tempst, P. and Hartl, F. U. (1999) Nature Struct. Biol. 6, 697-705.

94 Frydman, J. (2001) Annu. Rev. Biochem. 70, 603-647.

95 Ewalt, K. L., Hendrick, J. P., Houry, W. A. and Hartl, F. U. (1997) Cell 90, 491-500.

96 Petit, M. A., Bedale, W., Osipiuk, J., Lu, C., Rajagopalan, M., McInerney, P., Goodman, M. F. and Echols, H. (1994) J. Biol. Chem. 269, 23824-23829.

97 Buchberger, A., Schroder, H., Hesterkamp, T., Schonfeld, H. J. and Bukau, B. (1996) J. Mol. Biol. 261, 328-333.

98 Thomas, J. G. and Baneyx, F. (1996) Mol. Microbiol. 21, 1185-1196.

99 Fink, A. L. (1999) Physiol. Rev. 79, 425-449.

100 Jaenicke, R. (1995) Philos. Trans. Roy. Soc. Lond. B Biol. Sci. 348, 97-105.

101 Ranson, N. A., Dunster, N. J., Burston, S. G. and Clarke, A. R. (1995) J. Mol. Biol. 250, 581-586.

102 Jaenicke, R. and Seckler, R. (1997) Adv. Protein Chem. 50, 1-59.

103 Kopito, R. R. and Sitia, R. (2000) EMBO Rep. 1, 225-231.

104 Ben-Zvi, A. P. and Goloubinoff, P. (2001) J. Struct. Biol. 135, 84-93.

105 Goldberg, A. L. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 422-426.

106 Yen, C., Green, L. and Miller, C. G. (1980) J. Mol. Biol. 143, 21-33.

107 Schubert, U., Anton, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W. and Bennink, J. R. (2000) Nature 404, 770-774.

108 Turner, G. C. and Varshavsky, A. (2000) Science 289, 2117-2120.

109 Yarus, M. (1979) Prog. Nucl. Acid Res. Mol. Biol. 23, 195-225.

110 Zhu, J. and Winans, S. C. (2001) Proc. Nat. Acad. Sci. U.S.A. 98, 1507-1512.

111 Maurizi, M. R. (1992) Experientia 48, 178-201.

112 Goldberg, A. L. (1990) Semin. Cell Biol. 1, 423-432.

113 Vu, P. K. and Sakamoto, K. M. (2000) Mol. Genet. Metab. 71, 261-266.

114 Gottesman, S., Wickner, S. and Maurizi, M. R. (1997) Genes Dev. 11, 815-823.

115 Horwich, A. L., Weber-Ban, E. U. and Finley, D. (1999) Proc. Nat. Acad. Sci. U.S.A. 96, 11033-11040.

116 Wickner, S., Gottesman, S., Skowyra, D., Hoskins, J., McKenney, K. and Maurizi, M. R. (1994) Proc. Nat. Acad. Sci. U.S.A. 91, 12218-12222.

117 Levchenko, I., Luo, L. and Baker, T. A. (1995) Genes Dev. 9, 2399-2408.

118 Gottesman, S. (1996) Annu. Rev. Genet. 30, 465-506.

119 Katayama, Y., Gottesman, S., Pumphrey, J., Rudikoff, S., Clark, W. P. and Maurizi, M. R. (1988) J. Biol. Chem. 263, 15226-15236.

120 Rohrwild, M., Coux, O., Huang, H. C., Moerschell, R. P., Yoo, S. J., Seol, J. H., Chung, C. H. and Goldberg, A. L. (1996) Proc. Nat. Acad. Sci. U.S.A. 93, 5808-5813.

121 Hwang, B. J., Park, W. J., Chung, C. H. and Goldberg, A. L. (1987) Proc. Nat. Acad. Sci. U.S.A. 84, 5550-5554.

122 Zwickl, P., Baumeister, W. and Steven, A. (2000) Curr. Opin. Struct. Biol. 10, 242-250.

123 Ogura, T. and Wilkinson, A. J. (2001) Genes Cells 6, 575-597.

124 Zwickl, P. and Baumeister, W. (1999) Nature Cell Biol. 1, E97-98.

125 Neuwald, A. F., Aravind, L., Spouge, J. L. and Koonin, E. V. (1999) Genome Res. 9, 27-43.

126 Waxman, L. and Goldberg, A. L. (1982) Proc. Nat. Acad. Sci. U.S.A. 79, 4883-4887.

127 Menon, A. S., Waxman, L. and Goldberg, A. L. (1987) J. Biol. Chem. 262, 722-726.

128 Rasulova, F. S., Dergousova, N. I., Starkova, N. N., Melnikov, E. E., Rumsh, L. D., Ginodman, L. M. and Rotanova, T. V. (1998) FEBS Lett. 432, 179-181.

129 Ebel, W., Skinner, M. M., Dierksen, K. P., Scott, J. M. and Trempy, J. E. (1999) J. Bacteriol. 181, 2236-2243.

130 Tomoyasu, T., Gamer, J., Bukau, B., Kanemori, M., Mori, H., Rutman, A. J., Oppenheim, A. B., Yura, T., Yamanaka, K., Niki, H., Hiraga, S. and Ogura, T. (1995) EMBO J. 14, 2551-2560.

131 Karata, K., Verma, C. S., Wilkinson, A. J. and Ogura, T. (2001) Mol. Microbiol. 39, 890-903.

132 Guelin, E., Rep, M. and Grivell, L. A. (1994) Yeast 10, 1389-1394.

133 Lindahl, M., Tabak, S., Cseke, L., Pichersky, E., Andersson, B. and Adam, Z. (1996) J. Biol. Chem. 271, 29329-29334.

134 Yoo, S. J., Seol, J. H., Shin, D. H., Rohrwild, M., Kang, M. S., Tanaka, K., Goldberg, A. L. and Chung, C. H. (1996) J. Biol. Chem. 271, 14035-14040.

135 Seol, J. H., Yoo, S. J., Shin, D. H., Shim, Y. K., Kang, M. S., Goldberg, A. L. and Chung, C. H. (1997) Eur. J. Biochem. 247, 1143-1150.

136 Rohrwild, M., Pfeifer, G., Santarius, U., Muller, S. A., Huang, H. C., Engel, A., Baumeister, W. and Goldberg, A. L. (1997) Nature Struct. Biol. 4, 133-139.

137 Wang, J. (2001) J. Struct. Biol. 134, 15-24.

138 Wang, J., Song, J. J., Franklin, M. C., Kamtekar, S., Im, Y. J., Rho, S. H., Seong, I. S., Lee, C. S., Chung, C. H. and Eom, S. H. (2001) Structure 9, 177-184.

139 Sousa, M. C., Trame, C. B., Tsuruta, H., Wilbanks, S. M., Reddy, V. S. and McKay, D. B. (2000) Cell 103, 633-643.

140 Bochtler, M., Ditzel, L., Groll, M. and Huber, R. (1997) Proc. Nat. Acad. Sci. U.S.A. 94, 6070-6074.

141 Kang, M. S., Lim, B. K., Seong, I. S., Seol, J. H., Tanahashi, N., Tanaka, K. and Chung, C. H. (2001) EMBO J. 20, 734-742.

142 Lupas, A., Flanagan, J. M., Tamura, T. and Baumeister, W. (1997) Trends Biochem. Sci. 22, 399-404.

143 Huang, H. and Goldberg, A. L. (1997) J. Biol. Chem. 272, 21364-21372.

144 Maurizi, M. R., Thompson, M. W., Singh, S. K. and Kim, S. H. (1994) Meth. Enzymol. 244, 314-331.

145 Shanklin, J., DeWitt, N. D. and Flanagan, J. M. (1995) Plant Cell 7, 1713-1722.

146 Wojtkowiak, D., Georgopoulos, C. and Zylicz, M. (1993) J. Biol. Chem. 268, 22609-22617.

147 Wang, J., Hartling, J. A. and Flanagan, J. M. (1997) Cell 91, 447-456.

148 Wang, J., Hartling, J. A. and Flanagan, J. M. (1998) J. Struct. Biol. 124, 151-163.

149 Singh, S. K., Guo, F. and Maurizi, M. R. (1999) Biochemistry 38, 14906-14915.

150 Maurizi, M. R., Singh, S. K., Thompson, M. W., Kessel, M. and Ginsburg, A. (1998) Biochemistry 37, 7778-7786.

151 Kessel, M., Maurizi, M. R., Kim, B., Kocsis, E., Trus, B. L., Singh, S. K. and Steven, A. C. (1995) J. Mol. Biol. 250, 587-594.

152 Grimaud, R., Kessel, M., Beuron, F., Steven, A. C. and Maurizi, M. R. (1998) J. Biol. Chem. 273, 12476-12481.

153 Flanagan, J. M., Wall, J. S., Capel, M. S., Schneider, D. K. and Shanklin, J. (1995) Biochemistry 34, 10910-10917.

154 Gonciarz-Swiatek, M., Wawrzynow, A., Um, S. J., Learn, B. A., McMacken, R., Kelley, W. L., Georgopoulos, C., Sliekers, O. and Zylicz, M. (1999) J. Biol. Chem. 274, 13999-14005.

155 Gottesman, S., Roche, E., Zhou, Y. and Sauer, R. T. (1998) Genes Dev. 12, 1338-1347.

156 Hoskins, J. R., Pak, M., Maurizi, M. R. and Wickner, S. (1998) Proc. Nat. Acad. Sci. U.S.A. 95, 12135-12140.

157 Ishikawa, T., Beuron, F., Kessel, M., Wickner, S., Maurizi, M. R. and Steven, A. C. (2001) Proc. Nat. Acad. Sci. U.S.A. 98, 4328-4333.

158 Kim, Y. I., Burton, R. E., Burton, B. M., Sauer, R. T. and Baker, T. A. (2000) Mol. Cell 5, 639-648.

159 Ortega, J., Singh, S. K., Ishikawa, T., Maurizi, M. R. and Steven, A. C. (2000) Mol. Cell 6, 1515-1521.

160 Pak, M., Hoskins, J. R., Singh, S. K., Maurizi, M. R. and Wickner, S. (1999) J. Biol. Chem. 274, 19316-19322.

161 Singh, S. K., Rozycki, J., Ortega, J., Ishikawa, T., Lo, J., Steven, A. C. and Maurizi, M. R. (2001) J. Biol. Chem. 276, 29420-29429.

162 Thompson, M. W. and Maurizi, M. R. (1994) J. Biol. Chem. 269, 18201-18208.

163 Thompson, M. W., Singh, S. K. and Maurizi, M. R. (1994) J. Biol. Chem. 269, 18209-18215.

164 Levchenko, I., Smith, C. K., Walsh, N. P., Sauer, R. T. and Baker, T. A. (1997) Cell 91, 939-947.

165 Smith, C. K., Baker, T. A. and Sauer, R. T. (1999) Proc. Nat. Acad. Sci. U.S.A. 96, 6678-6682.

166 Trame, C. B. and McKay, D. B. (2001) Acta Crystallogr. D Biol. Crystallogr. 57, 1079-1090.

167 Banecki, B., Wawrzynow, A., Puzewicz, J., Georgopoulos, C. and Zylicz, M. (2001) J. Biol. Chem. 276, 18843-18848.

168 Lo, J. H., Baker, T. A. and Sauer, R. T. (2001) Protein Sci. 10, 551-559.

169 Gonzalez, M., Frank, E. G., McDonald, J. P., Levine, A. S. and Woodgate, R. (1998) Acta Biochim. Pol. 45, 163-172.

170 Wickner, S. and Maurizi, M. R. (1999) Proc. Nat. Acad. Sci. U.S.A. 96, 8318-8320.

171 Baker, T. A. (1999) Nature 401, 29-30.

172 Burton, R. E., Siddiqui, S. M., Kim, Y. I., Baker, T. A. and Sauer, R. T. (2001) EMBO J. 20, 3092-3100.

173 Gottesman, S., Clark, W. P., de Crecy-Lagard, V. and Maurizi, M. R. (1993) J. Biol. Chem. 268, 22618-22626.

174 Singh, S. K. and Maurizi, M. R. (1994) J. Biol. Chem. 269, 29537-29545.

175 Hochstrasser, M. and Wang, J. (2001) Nature Struct. Biol. 8, 294-296.

176 Lee, C., Schwartz, M. P., Prakash, S., Iwakura, M. and Matouschek, A. (2001) Mol. Cell 7, 627-637.

177 Burton, B. M., Williams, T. L. and Baker, T. A. (2001) Mol. Cell 8, 449-454.

178 Chuang, S. E. and Blattner, F. R. (1993) J. Bacteriol. 175, 5242-5252.

179 Wenzel, T., Eckerskorn, C., Lottspeich, F. and Baumeister, W. (1994) FEBS Lett. 349, 205-209.

180 Kisselev, A. F., Akopian, T. N., Woo, K. M. and Goldberg, A. L. (1999) J. Biol. Chem. 274, 3363-3371.

181 Kisselev, A. F., Akopian, T. N. and Goldberg, A. L. (1998) J. Biol. Chem. 273, 1982-1989.

182 Suzuki, C. K., Rep, M., van Dijl, J. M., Suda, K., Grivell, L. A. and Schatz, G. (1997) Trends Biochem. Sci. 22, 118-123.

183 Schumann, W. (1999) FEMS Microbiol. Rev. 23, 1-11.

184 Goff, S. A. and Goldberg, A. L. (1987) J. Biol. Chem. 262, 4508-4515.

185 Goldberg, A. L. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 2640-2644.

186 Kruger, E., Witt, E., Ohlmeier, S., Hanschke, R. and Hecker, M. (2000) J. Bacteriol. 182, 3259-3265.

187 Keiler, K. C., Waller, P. R. and Sauer, R. T. (1996) Science 271, 990-993.

188 Karzai, A. W., Roche, E. D. and Sauer, R. T. (2000) Nature Struct. Biol. 7, 449-455.

189 Flynn, J. M., Levchenko, I., Seidel, M., Wickner, S. H., Sauer, R. T. and Baker, T. A. (2001) Proc. Nat. Acad. Sci. U.S.A. 98, 10584-10589.

190 Levchenko, I., Seidel, M., Sauer, R. T. and Baker, T. A. (2000) Science 289, 2354-2356.

191 Jubete, Y., Maurizi, M. R. and Gottesman, S. (1996) J. Biol. Chem. 271, 30798-30803.

192 Kandror, O., Sherman, M. and Goldberg, A. (1999) J. Biol. Chem. 274, 37743-37749.

193 Sherman, M. Y. and Goldberg, A. L. (1996) in Stress-inducible Cellular Responses. (Feige, U., Morimoto, R. I., Yahara, I. and Bolla, B., eds.) pp. 57-78, Birkhause Verlag, Basel.

194 Kandror, O., Sherman, M., Rhode, M. and Goldberg, A. L. (1995) EMBO J. 14, 6021-6027.

195 Kandror, O., Busconi, L., Sherman, M. and Goldberg, A. L. (1994) J. Biol. Chem. 269, 23575-23582.

196 Sherman, M. and Goldberg, A. L. (1992) EMBO J. 11, 71-77.

197 Goff, S. A., Casson, L. P. and Goldberg, A. L. (1984) Proc. Nat. Acad. Sci. U.S.A. 81, 6647-6651.

198 Huang, H. C., Sherman, M. Y., Kandror, O. and Goldberg, A. L. (2001) J. Biol. Chem. 276, 3920-3928.

199 Hershko, A. and Ciechanover, A. (1998) Annu. Rev. Biochem. 67, 425-479.

200 Truscott, K. N., Pfanner, N. and Voos, W. (2001) Rev. Physiol. Biochem. Phamacol. 143, 81-136.

201 Matouschek, A., Pfanner, N. and Voos, W. (2000) EMBO Rep. 1, 404-410.

202 Stahlberg, H., Kutejova, E., Suda, K., Wolpensinger, B., Lustig, A., Schatz, G., Engel, A. and Suzuki, C. K. (1999) Proc. Nat. Acad. Sci. U.S.A. 96, 6787-6790.

203 Sakuragi, S., Liu, Q. and Craig, E. (1999) J. Biol. Chem. 274, 11275-11282.

204 Leonhard, K., Stiegler, A., Neupert, W. and Langer, T. (1999) Nature 398, 348-351.

205 Terada, K., Kanazawa, M., Bukau, B. and Mori, M. (1997) J. Cell. Biol. 139, 1089-1095.

206 Krzewska, J., Langer, T. and Liberek, K. (2001) FEBS Lett. 489, 92-96.

207 Liu, Q., Krzewska, J., Liberek, K. and Craig, E. A. (2001) J. Biol. Chem. 276, 6112-6118.

208 Brodsky, J. L. and McCracken, A. A. (1999) Semin. Cell. Dev. Biol. 10, 507-513.

209 Chevet, E., Cameron, P. H., Pelletier, M. F., Thomas, D. Y. and Bergeron, J. J. (2001) Curr. Opin. Struct. Biol. 11, 120-124.

210 Kraulis, P. J. (1991) J. Appl. Crystallog. 24, 946-950.

211 Nicholls, A., Sharp, K. A. and Honig, B. (1991) Proteins 11, 281-296.

Figure 1: Possible fates of misfolded proteins in the cell.

Figure 2: The DnaK/DnaK/GrpE chaperone system. (A) The substrate binding (1DKZ) and ATPase domains (1DKG) of DnaK. Arrows represent β–strands and coils represent α-helices. For this, and other figures, the four letter codes in parenthesis represent the PDB accession codes used to create the figure. (B) Current structural model of DnaJ. From top to bottom: the N-terminal, or J-homology domain (1BQ0), the zinc finger domain (1EXK), the GF domain of unknown structure and C-terminal domain (1C3G). (C) The structure of GrpE (1DKG). (D) The basic details of the DnaK/DnaJ/GrpE reaction cycle. Figure 2a-c, 3a and 7 were generating using the program Molscript (210).

Figure 3: The GroELGroES chaperone system. (A) GroEL undergoes a conformation change in the hinge region to change the position of the apical domain, so that GroES can bind. (1AON, 1DER) A cut away of the oligomer followed by the monomer fold is shown in each case, GroEL and GroEL/GroES/ATP. Pictures of the GroEL and GroEL/GroES oligomers were rendered with the program GRASP (211). (B) Cartoon of the reaction cycle of GroEL/GroES.

Figure 4: Domain components of AAA-ATPases. This figure is modeled after Figure 1 in ref (165).

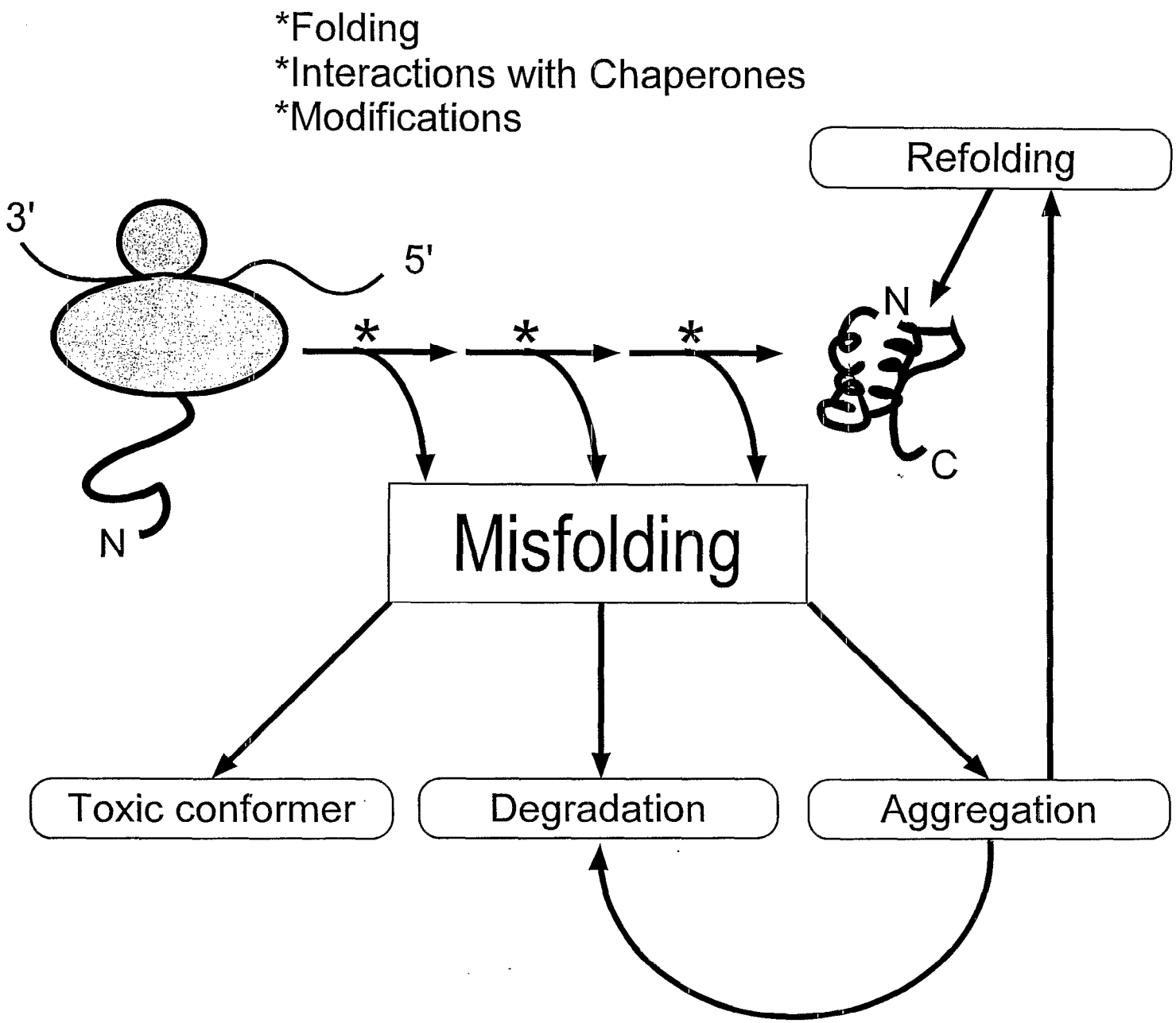
Figure 5: Interplay between Trigger factor and the DnaK and GroEL chaperone systems in *de novo* protein folding.

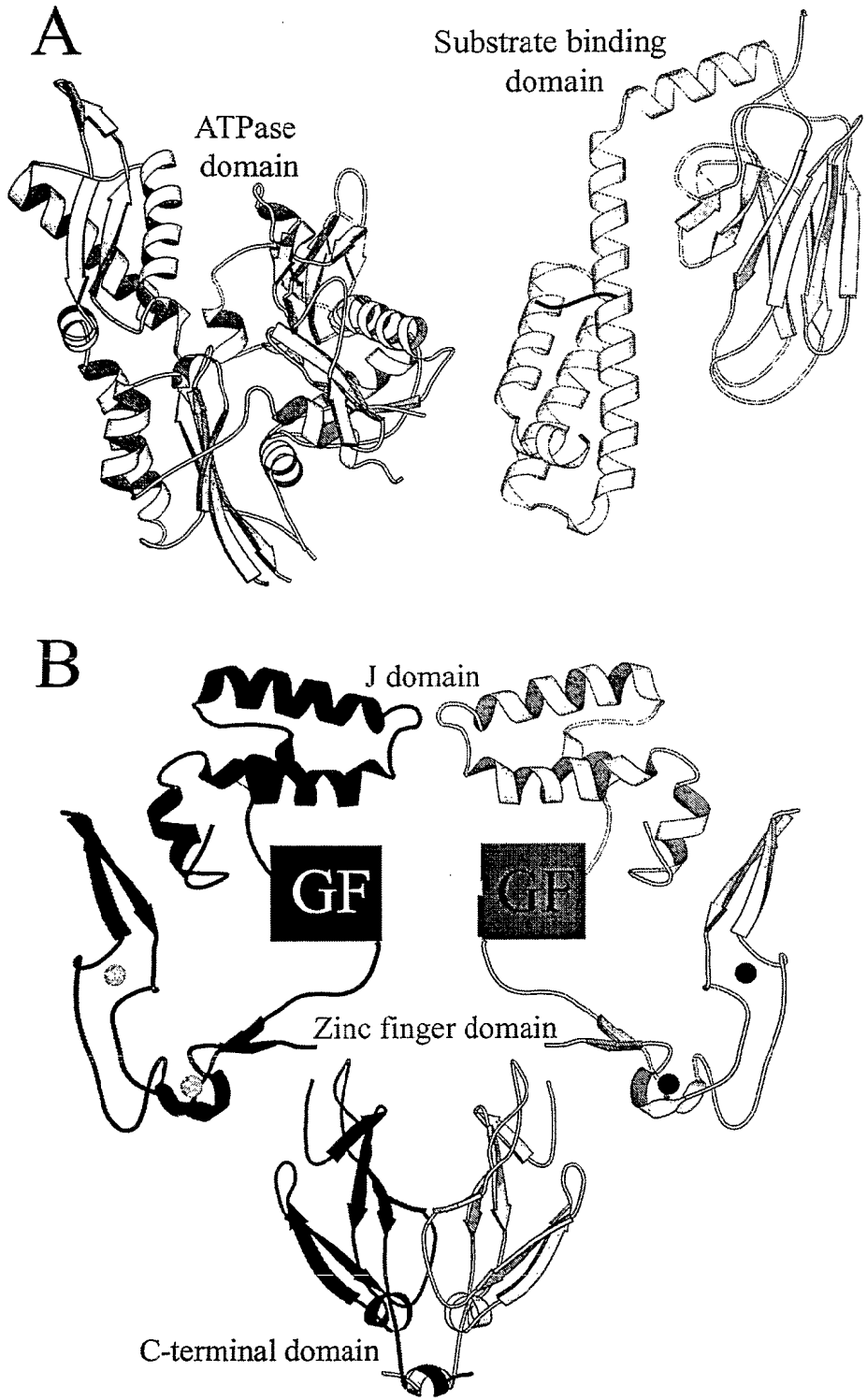
Figure 6: Interplay between the DnaK and ClpB chaperone systems.

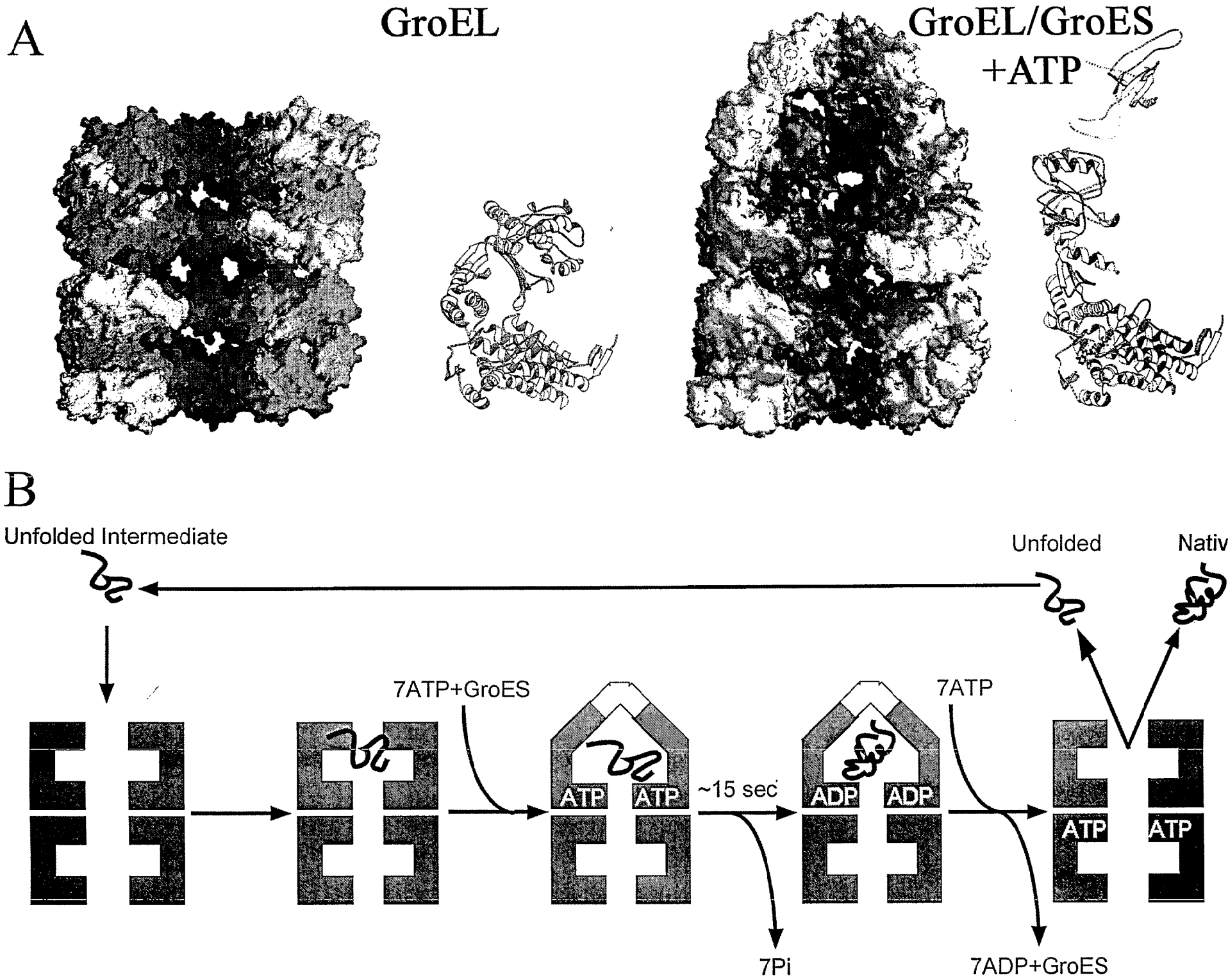
Figure 7: Crystal structures and molecular model of the ClpXP and HslUV protease machines. The ATPase components, ClpX and HslU, share a common fold, as shown in black. ClpX was obtained by homology modeling from the coordinates of HslU (1G41). The proteolytic components, ClpP (1TYF) and HslV (1G3I), are different, as shown in light gray.

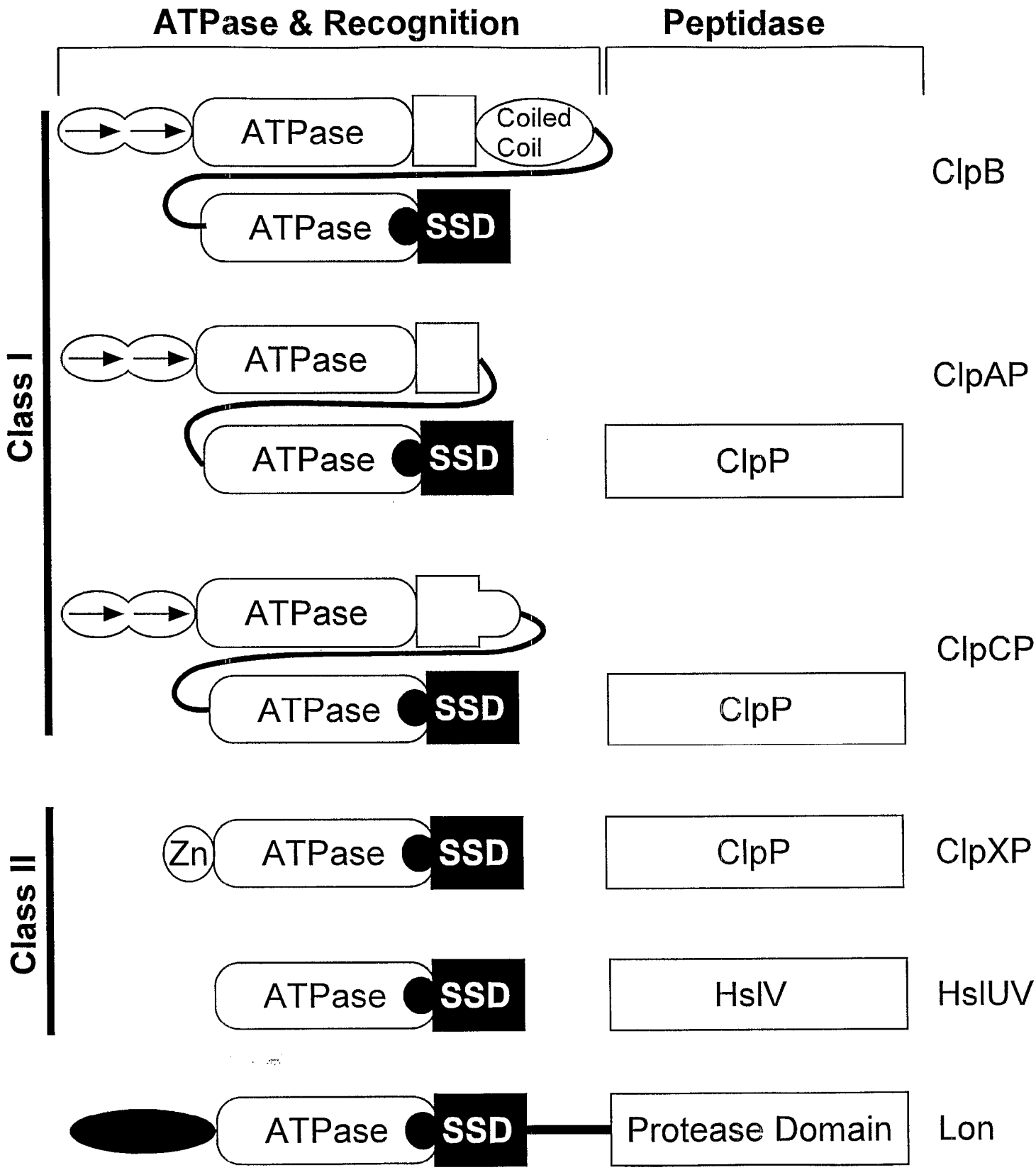
Figure 8: Basic features of ATP-dependent proteolysis.

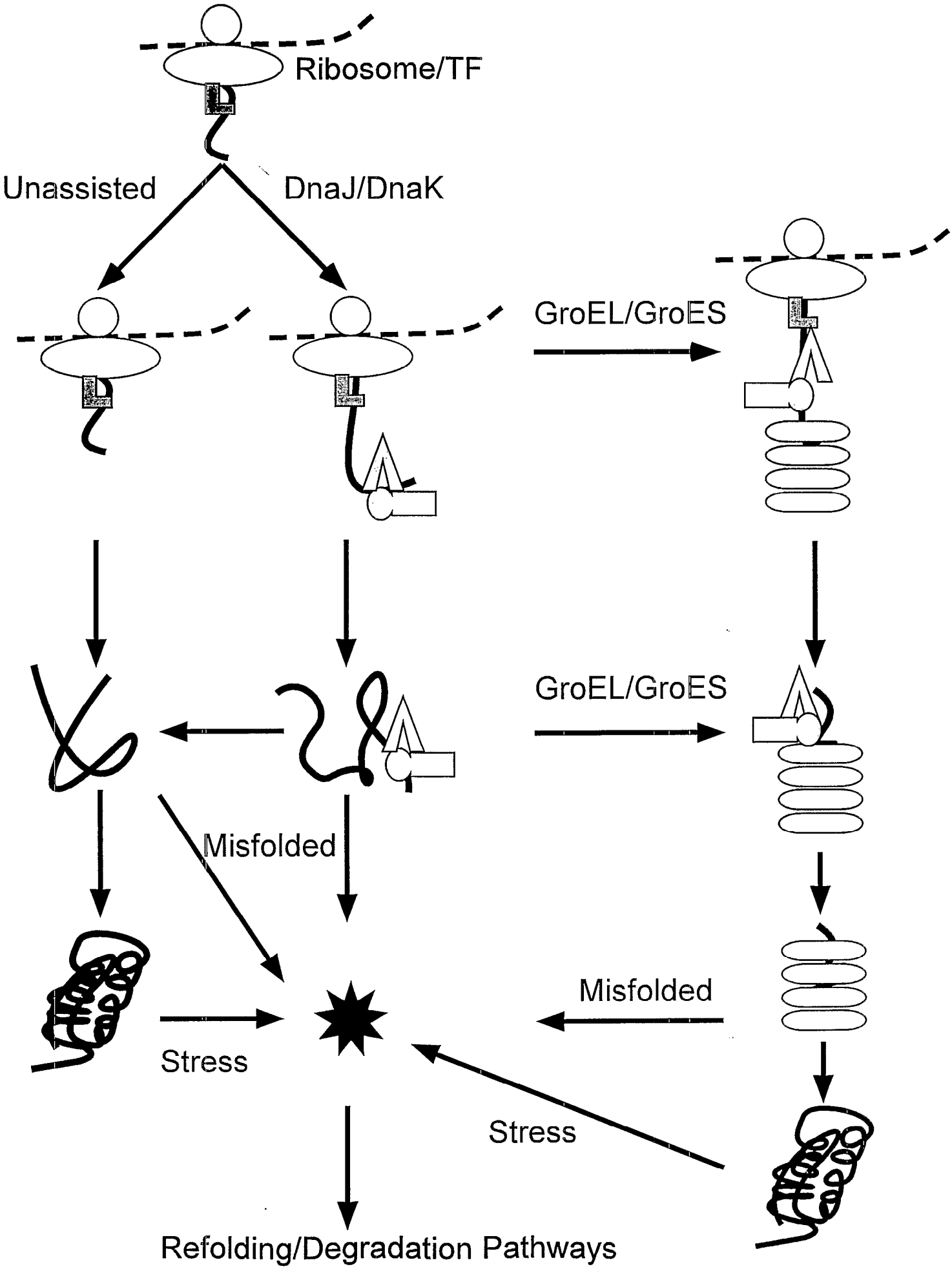
Figure 9: The recycling pathway gives rise to new protein synthesis.

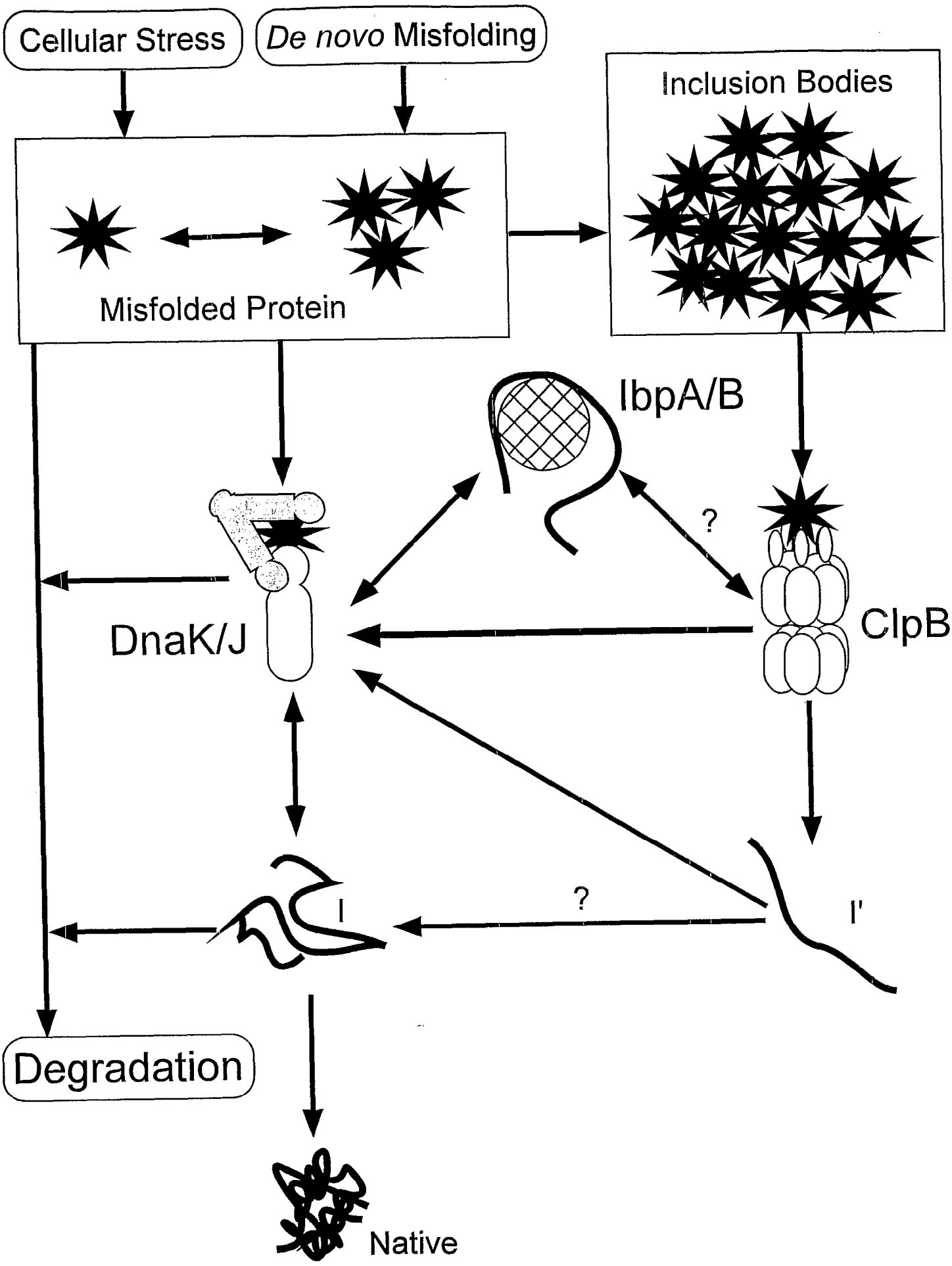

*Folding
*Interactions with Chaperones
*Modifications
Refolding
3'
5'
N
*
*
*
N
C
Misfolding
Toxic conformer
Degradation
Aggregation

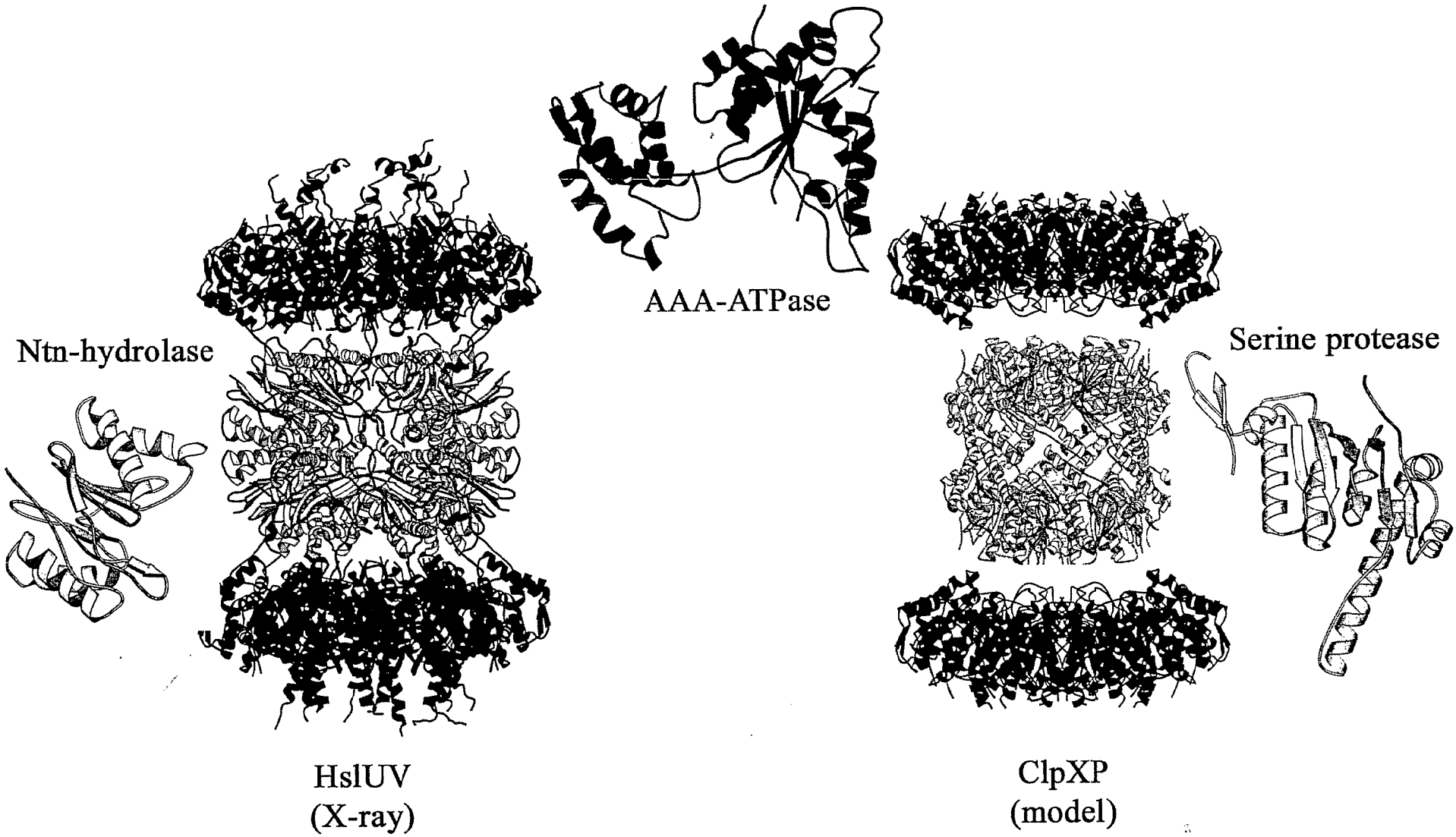

A
Substrate binding domain
ATPase domain
B
J domain
GF
GF
Zinc finger domain
C-terminal domain
C
D
DnaJ
I
Pi
ATP
DnaK
ADP
DnaK-System Reaction Cycle
GrpE
N
ATP
ADP,DnaJ

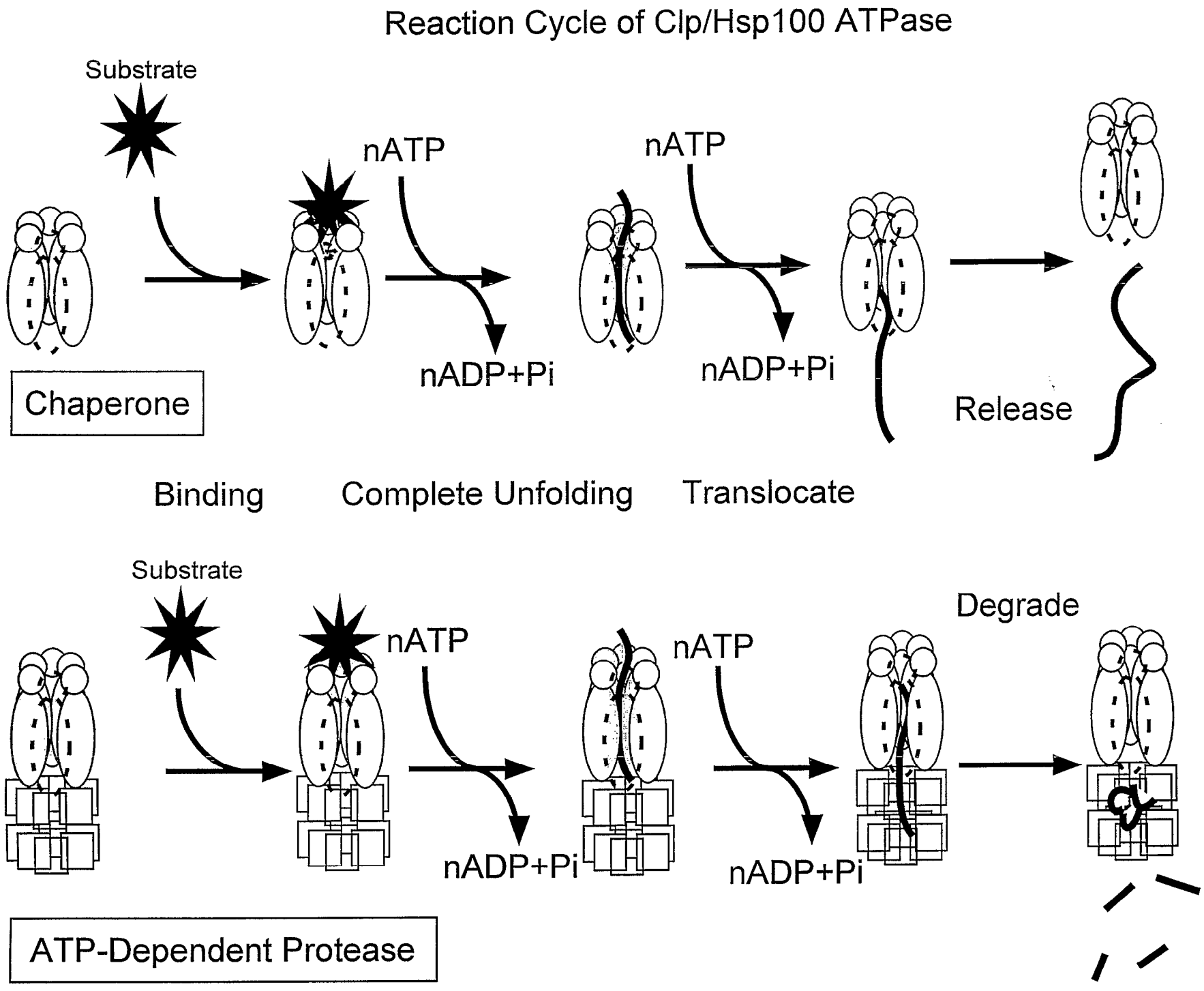

A
GroEL
GroEL/GroES
+ATP
B
Unfolded Intermediate
Unfolded
Native
7ATP+GroES
ATP
ATP
~15 sec
ADP
ADP
7ATP
ATP
ATP
7Pi
7ADP+GroES

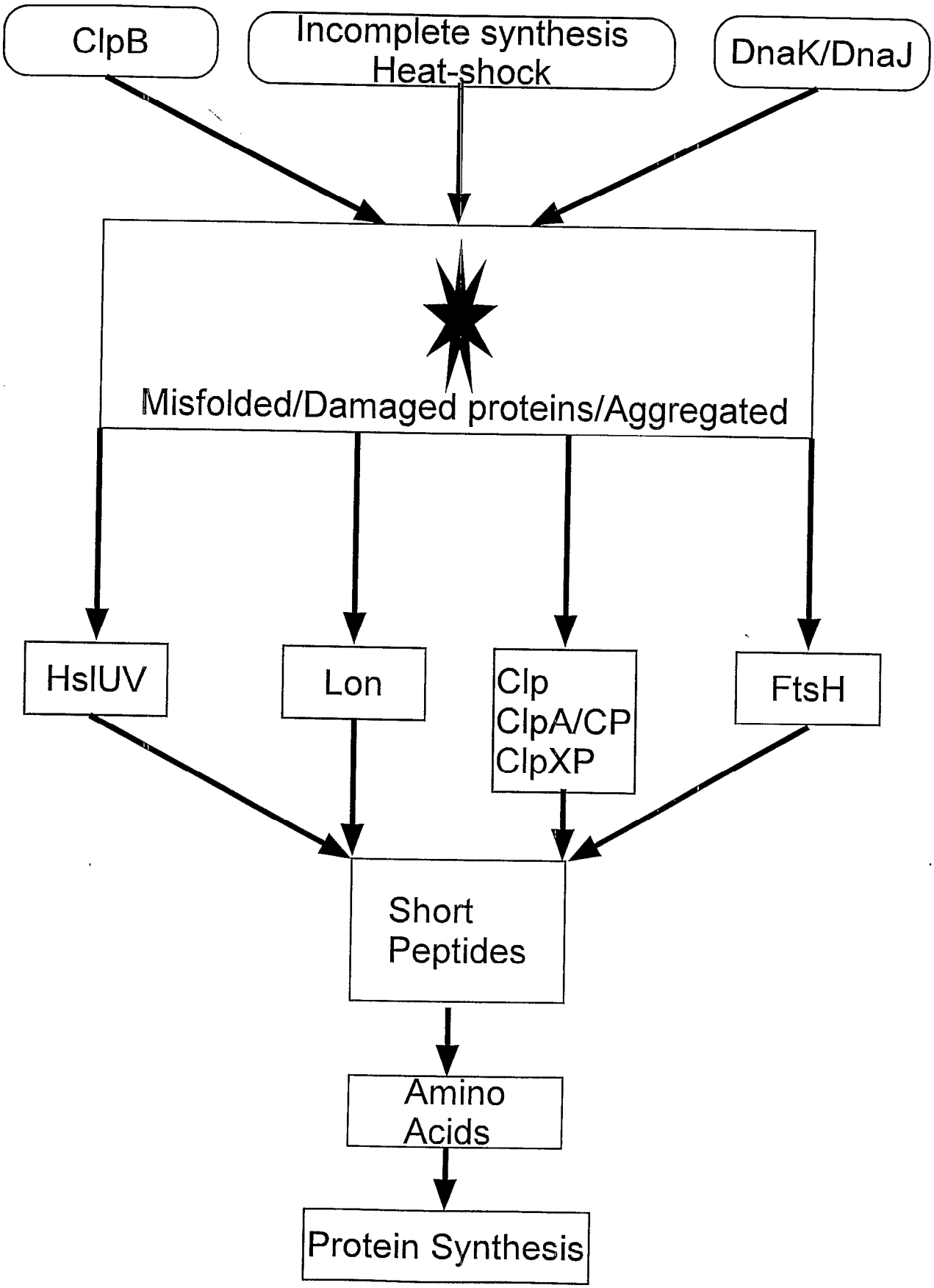

ATPase & Recognition
Peptidase
Class I
Class II
ATPase
Coiled Coil
SSD
ClpB
ClpAP
ClpP
ClpCP
Zn
ClpXP
HslV
HslUV
Protease Domain
Lon


Ribosome/TF
Unassisted
DnaJ/DnaK
GroEL/GroES
GroEL/GroES
Misfolded
Misfolded
Stress
Stress
Refolding/Degradation Pathways


Cellular Stress
De novo Misfolding
Inclusion Bodies
Misfolded Protein
IbpA/B
?
DnaK/J
ClpB
I
?
I'
Degradation
Native


Ntn-hydrolase
AAA-ATPase
Serine protease
HslUV
(X-ray)
ClpXP
(model)


Reaction Cycle of Clp/Hsp100 ATPase
Substrate
nATP
nATP
nADP+Pi
nADP+Pi
Chaperone
Release
Binding
Complete Unfolding
Translocate
Substrate
nATP
nATP
Degrade
nADP+Pi
nADP+Pi
ATP-Dependent Protease


ClpB
Incomplete synthesis
Heat-shock
DnaK/DnaJ
Misfolded/Damaged proteins/Aggregated
HslUV
Lon
Clp
ClpA/CP
ClpXP
FtsH
Short
Peptides
Amino
Acids
Protein Synthesis