Proteins must fold into their correct three-dimensional conformation in order to attain their biological function. Conversely, protein aggregation and misfolding are primary contributors to many devastating human diseases, such as prion-mediated infections, Alzheimer’s disease, type II diabetes and cystic fibrosis. While the native conformation of a polypeptide is encoded within its primary amino acid sequence and is sufficient for protein folding in vitro (Anfinsen, 1973), the situation inside the cells is considerably more complex. Inside the cell, proteins are synthesized or folded continuously; a process that is greatly assisted by molecular chaperones. Molecular chaperones are a group of structurally diverse and mechanistically distinct proteins that either promote folding or prevent the aggregation of other proteins. With our increasing understanding of the proteome, it is becoming clear that the number of proteins that can be classified as molecular chaperones is increasing steadily. Many of these proteins have novel but essential cellular functions that differ from that of more “conventional” chaperones, such as Hsp70 and the GroE system. This review focuses on the emerging role of molecular chaperones in protein quality control, i.e. the mechanism that rids the cell of misfolded or incompletely synthesized polypeptides that otherwise would interfere with normal cellular function.

Keywords: Clp/Hsp100, Molecular chaperones, Proteasome, Protein folding, Protein quality control

Introduction

Molecular chaperones are a group of structurally diverse proteins that share the ability to interact with the non-native conformation of other proteins. While the native conformation of a polypeptide is encoded within its primary amino acid sequence and is sufficient for protein folding in vitro (Anfinsen, 1973), the situation inside the cells is considerably more complex. It is now firmly established that molecular chaperones play a crucial role in facilitating the correct folding of proteins in vivo by either preventing protein aggregation or facilitating the forward folding and assembly of proteins into higher order structures (Hartl and Hayer-Hartl, 2002; Young et al., 2004). This may be achieved either via an active, ATP-dependent mechanism (e.g. folding machines) or via a passive, ATP-independent mechanism (e.g. holders and folding catalysts). Many of the proteins that function as molecular chaperones were originally identified as “heat shock proteins” (Hsp), since their abundance increases in cells upon thermal stress. Heat shock proteins are categorized by their molecular weight and many molecular chaperones are better known by this nomenclature (Fig. 1).

With our increasing understanding of the proteome of several microbial and other higher-order organisms, it is becoming clear that the number of proteins that fulfill the “classical” definition of a molecular chaperone is increasing rapidly. While some of these proteins are not required for thermotolerance, others are essential to the stress-response but neither promote the forward folding nor prevent the aggregation of proteins (Sauer et al., 2004; Wang et al., 2004). The latter group may even recognize native, folded proteins as substrates and, unlike more conventional chaperones, mediate the disassembly or unfolding of substrate proteins (Sauer et al., 2004). Many of these “unfoldases” belong to the superfamily of AAA+ ATPases (ATPases associated with a variety of functional activities) that function as molecular machines by converting metabolic energy in the form of ATP into mechanical work (Ogura and Wilkinson, 2001).

The role of molecular chaperones in protein folding has been reviewed extensively in the literature (Hartl and Hayer-Hartl, 2002; Walter and Buchner, 2002; Deuerling and Bukau, 2004; Young et al., 2004). Hence, this review will focus on the emerging role of molecular chaperones in “protein triage”, i.e. the function of molecular chaperones in protein quality control.
Molecular chaperones in protein quality control

Protein quality control is an essential cellular process underlying all stages of life. While the individual components and the nature of their interactions differ, the fundamental mechanism underlying this process is remarkably conserved in all three kingdoms.

At the core of the protein quality control system are energy-dependent proteases that are required for the degradation of substrate proteins (Horwich et al., 1999; Groll et al., 2005). Notable examples include the 20S proteasome and the proteolytic core of the ATP-hydrolyzing protease Ti (ClpP). At the molecular level, the 20S proteasome is composed of its 11S RP from Trypanosoma brucei (Whitby et al., 2000). The 20S proteasome is colored in orange and the 11S RP in green. The structure of the 20S proteasome resembles that of a long, cylinder-shaped barrel. Unlike the 19S RP, the 11S regulator adopts a seven-fold symmetry as seen with the 20S proteasome. (B) Structural model for the BAP-ClpP-BAP, disaggregating-degrading machine (Weibezahn et al., 2004). The BAP-ClpP complex is assembled from the crystal structures of Escherichia coli ClpP (Wang et al., 1997) and Thermus thermophilus ClpB (Lee et al., 2003b). ClpP is colored in orange and BAP in green. It is worth noting that the quaternary structure of the ClpP peptidase shows a remarkable structural resemblance to the inner core of the 20S proteasome.

Molecular chaperones in protein quality control

Protein quality control is an essential cellular process underlying all stages of life. While the individual components and the nature of their interactions differ, the fundamental mechanism underlying this process is remarkably conserved in all three kingdoms.

At the core of the protein quality control system are energy-dependent proteases that are required for the degradation of
Molecular Chaperones 261

Many RP components belong to the AAA+ ATPase superfamily. While RPs by themselves can self-assemble and function as molecular chaperones in vitro and in vivo, their primary function is to facilitate the unfolding of substrates for subsequent degradation by the associated proteases (Horwich et al., 1999).

**AAA+ ATPases: A superfamily of ATPases associated with diverse cellular activities**

AAA ATPases are found in all living organisms (Beyer, 1997; Neuwald et al., 1999) and, as their name suggests, are involved in a variety of cellular processes including DNA replication, membrane fusion and protein degradation (Vale, 2000; Ogura and Wilkinson, 2001). Advances in bioinformatics methods have revealed that the AAA ATPase family is in fact much broader than originally anticipated (Neuwald et al., 1999). In addition to classical AAA proteins, the extended AAA or AAA+ family also includes chaperone-like ATPases that function in the assembly, operation and disassembly of protein complexes.

AAA+ ATPases possess at least one Walker-type nucleotide-binding domain (NBD) and have a strong propensity to form hexameric ring structures as first determined by electron microscopy (Vale, 2000; Ogura and Wilkinson, 2001). High-resolution structural information of several AAA+ proteins is now available (Vale, 2000; Ogura and Wilkinson, 2001). These structures reveal that the NBD or α/β domain has a RecA-like mononucleotide-binding fold. The NBD is immediately followed by an α-helical domain, which together form the functional AAA+ module required for stable ATP binding.

In yeast and other higher Eukarya, the RP of the 26S proteosome (19S) consists of at least 18 polypeptides. The base of the 19S RP is composed of six distinct AAA+ ATPases (Rpt1-Rpt6). These six AAA+ ATPases form a ring-like assembly that abuts the axial pore of the 20S complex and exhibits chaperone activity in vitro (Glickman et al., 1998; Braun et al., 1999). Since the 20S proteosome forms a heptameric assembly, the association of the 19S RP with the 20S proteasome creates a symmetry mismatch. It is conceivable that this mismatch facilitates the unraveling of folded substrates. However, the underlying mechanism and the stereo-specific interactions between the base of the 19S RP and the 20S proteasome remain unclear.

In eubacteria, the ClpP peptidase is the structural homolog to the 20S proteosome (19S). Like its eukaryotic/archaeal counterpart, ClpP associates with AAA+ ATPases, such as ClpA and ClpX, which facilitate the unfolding of substrate proteins in an ATP-dependent manner. ClpA and ClpX belong to the Clp/Hsp100 family of AAA+ ATPase, which, like the base of the 19S RP, form a hexameric ring-structure and exhibit chaperone activity (Wickner et al., 1994; Levchenko et al., 1995; Wawrzynow et al., 1995).

The Clp/Hsp100 family of molecular chaperones

Members of the Clp/Hsp100 chaperone family are found in Prokarya, yeast and higher Eukarya including man. As their name suggests, Clp/Hsp100 proteins have a molecular weight of about 100 kDa per monomer. In principle, two types of Clp/Hsp100 proteins are distinguished (Fig. 3): Class 1 members, such as ClpA and ClpB, posses two distinct but

---

**Fig. 3.** Three-dimensional structure of class 1 and class 2 Clp/Hsp100 AAA+ ATPases. The ribbon diagrams depict the monomer structure of representative members of each family. (A) *Escherichia coli* ClpA (Guo et al., 2002), (B) *Thermus thermophilus* ClpB (Lee et al., 2003b), (C) *Helicobacter pylori* ClpX (Kim and Kim, 2003) and (D) *Haemophilus influenzae* ClpY (HslU) (Sousa et al., 2000). The N-terminal domain is colored in gold, the α/β domain of the first AAA+ module (NBD1) in cyan, the α-helical domain of the first AAA+ module in pink, the α/β domain of the second AAA+ module (NBD2) in green, and the α-helical domain of the second AAA+ module in purple. Bound nucleotides are shown as CPK models and are colored in blue (NBD1) and red (NBD2). The Clp/Hsp100 structures were superimposed and aligned through the second AAA+ module that is common to all Clp/Hsp100 proteins. ClpB and HslU possess additional structural features, known as the M-domain in ClpB and the I-domain in HslU, which are specific to their respective chaperone function. The location of the ClpA IGF/L tripeptide motif is shown and labeled accordingly.
conserved NBDs, whereas class 2 members, such as ClpX and ClpY (HslU), have a single NBD that is sufficient for function. It is believed that members of the class 1 family may have arisen by gene fusion since the sequence homology between equivalent NBDs of different class 1 members is significantly greater than between the first and second NBD within the same protein (Schirmer et al., 1996).

Structural and biochemical studies have shown that Clp/Hsp100 proteins form a homo-hexameric ring structure and function as molecular chaperones by themselves (Sauer et al., 2004). Inside the cell, however, Clp/Hsp100 proteins are usually associated with the ClpP protease to form fully-competent degrading machines (Fig. 2). Binding to ClpP requires an IGF/L tripeptide motif that is found in most Clp/Hsp100 proteins (Kim et al., 2001; Joshi et al., 2004), including ClpA, ClpX and HslU, but with the exception of ClpB. Unlike ClpA and ClpX, members of the ClpB family neither associate with energy-dependent proteases nor mediate the degradation of their substrate proteins. Instead, ClpB is a bona fide molecular chaperone that has the remarkable ability to rescue stress-damaged proteins from an aggregated state (Parsell et al., 1994). Hence, Clp/Hsp100 proteins play a pivotal role in deciding whether proteins are disaggregated and refolded or unfolded and degraded, and thus are central to protein triage in eubacteria.

The high-resolution crystal structure of ClpA (Guo et al., 2002), ClpB (Lee et al., 2003b), ClpX (Kim and Kim, 2003), and HslU (Bochtler et al., 2000; Sousa et al., 2000) have been determined. These structures reveal that Clp/Hsp100 proteins share a similar domain structure (Fig. 3). The relative orientation of these domains, however, can differ dramatically. This is particularly noticeable by comparing the structure of ClpA with ClpB, which both possess two AAA+ domains (Fig. 3). The functional implication of this is unclear. It is conceivable that the different domain arrangements are caused by different crystal lattice interactions. Nevertheless, the structural plasticity suggests that Clp/Hsp100 proteins are flexible and can undergo structural changes, consistent with the notion that Clp/Hsp100 proteins are molecular machines that perform mechanical work.

**ClpB-A protein disaggregating machine**

ClpB and its eukaryotic orthologs, yeast Hsp104 and plant Hsp101, are heat-shock induced and are essential for survival at high temperatures (Sanchez and Lindquist, 1990; Squires et al., 1991; Squires and Squires, 1992; Schirmer et al., 1996). In addition, Hsp104 has also been implicated more broadly in other protective functions such as providing tolerance to ethanol, arsenate, and long-term storage in the cold (Sanchez et al., 1992). Members of the ClpB family differ from other Clp/Hsp100 proteins by possessing a longer middle region (M-domain) that is inserted into the α-helical domain of the first AAA+ module (Fig. 3). This M-domain forms a flexible, 85 Å long coiled-coil that resembles in structure the shape of a two-bladed propeller of an airplane (Lee et al., 2003b, Lee et al., 2004). The mobility and orientation of this coiled-coil is essential for chaperone activity (Cashikar et al., 2002; Mogk et al., 2003) as immobilizing the coiled-coil by sulfhydryl crosslinking abolishes chaperone activity (Lee et al., 2003b).

However, while ClpB is a bona fide ATP-dependent molecular chaperone, it differs from other chaperones by neither promoting the forward folding nor preventing the aggregation of proteins. Instead, ClpB has the ability to rescue stress-damaged proteins from an aggregated state. The full recovery of these proteins requires the assistance of the cognate DnaK chaperone system (Glover and Lindquist, 1998; Goloubinoff et al., 1999; Mogk et al., 1999; Motohashi et al., 1999; Zolkiewski, 1999), which together with ClpB form a bi-chaperone network.

We have recently crystallized and determined the three-dimensional crystal structure of the *Thermus thermophilus* ClpB monomer in the AMPPNP-bound state (Lee et al., 2003a; Lee et al., 2003b). The crystals belong to the orthorhombic space group P2,2,2 and lack the 6-fold screw axis that is present in the crystal structure of many other AAA+ ATPases, including ClpA (Guo et al., 2002) and ClpX (Kim and Kim, 2003). Instead, in our crystal, there are three independent representations of full-length ClpB in the crystallographic asymmetric unit. The three molecules are related by a ~60° rotation and a 1/6 translation along the crystallographic two-fold screw axis parallel to α, which gives rise to a helical assembly similar to that as observed with other AAA+ ATPases. Most remarkably, each ClpB molecule adopts a different three-dimensional conformation, even though they are in the same nucleotide-bound state. The presence of three distinct conformations of full-length ClpB is very unusual, as most crystal structures are by definition static. The observed structural plasticity suggests that ClpB is a dynamic molecule, consistent with the notion that ClpB is a molecular machine that undergoes large, *en bloc* conformational changes to perform its chaperone function.

To examine the structure of the functional ClpB assembly, we have determined the structure of the ClpB-AMPPNP complex using electron cryomicroscopy (cryo-EM) (Lee et al., 2003a; Lee et al., 2003b). Our cryo-EM reconstruction shows that ClpB is a hexamer that forms a two-tiered ring structure. There is an ~16 Å wide hole in the top ring and six smaller openings on the lateral surface of the molecule. This quaternary structure gives rise to an internal cavity of about 60,000 Å³, which is less than half of that reported for ClpA (~140,000 Å³) (Beurun et al., 1998). However, since ClpB is a molecular machine that undergoes large ATP-driven conformational changes (S. Lee and F.T.F. Tsai, unpublished data), it is likely that the volume of the internal cavity changes during the ATP binding and hydrolysis cycle.

While the structure-function relationship of ClpB is beginning to be understood, the underlying biochemistry and cellular function of ClpB has remained unclear. To address
this question, we have engineered a ClpB variant, known as BAP, by replacing a helix-loop-helix motif in ClpB with the analogous motif of ClpA, which contains the conserved IGF/L tripeptide motif required for ClpP binding (Weibezahn et al., 2004). Unlike ClpB, BAP can associate with the ClpP protease in a stable, ATP-dependent manner. While ClpB and BAP share the ability to disaggregate proteins, in the presence of ClpP, BAP (but not ClpB) functions as a novel disaggregating-degrading machine (Fig. 4). This BAP-ClpP complex resembles both structurally and functionally that of the ClpA-ClpP complex, but with somewhat different specificities. Moreover, using BAP, we have demonstrated that substrates must translocate through the central pore of the ClpB hexamer and, more importantly, that thermostolerance requires the refolding of aggregated proteins, i.e. it is not the aggregate itself that causes cell death.

Conclusion

While the three-dimensional structure of representative members from each major chaperone family has been determined, there is still much to be learned before we can fully understand and appreciate the biochemistry and function of molecular chaperones in vitro and in vivo. Also, it is clear that the number of novel chaperones with unique and highly specialized function is increasing, supporting the notion that chaperones play a key role in many if not all cellular processes that are essential to life. This includes the emerging role of molecular chaperones in remodeling transcriptional regulatory complexes, in the dislocation of misfolded membrane proteins, and many other vital cellular functions. The availability of high-resolution, three-dimensional structure information opens up new avenues of exciting research that focuses on the mechanism and mechanical function of molecular chaperones and may be exploited in biotechnology and nanomedicine. Clearly, “This is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning.” (Winston Churchill).

Acknowledgments

We wish to thank members of the Tsai lab for their various contributions. S.L. is supported by the US Department of the Army (W81XWH-04-1-0033) and F.T.F.T. by the National Institute of Health (GM67672) and the Robert A. Welch Foundation (Q-1530).

References


