Cell Stress & Chaperones (1996) 1 (3), 155–160

Discovery of molecular chaperones

R. John Ellis

Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK

When I wish to provoke my colleagues at the University of Warwick I invite them to challenge my contention that there have been no really major conceptual advances in biochemistry since 1963, when chemiosmosis and allostery illuminated the cellular scene. There have, of course, been huge increases in biochemical knowledge in the last 33 years, but my point is that these have been technical rather than conceptual in origin, and represent a consolidation phase in the development of the discipline rather than a wave of innovation. Occasionally I am offered the signal hypothesis in rebuttal, and there is much merit in this suggestion; not only does protein targeting maintain the ultrastructure of the cell from generation to generation, but its reliance on the correct membrane location of pre-existing signal receptors gives the lie to the oft-stated claim that all the information to specify a cell lies within the DNA of that cell. Of course what I am hoping they will suggest is that the general concept of molecular chaperones also provides an exception to my contention, but they never do—they recognize a fishing expedition when they see one! In this article I give a personal account of my involvement in the origin of the molecular chaperone concept to complement the article on the discovery of the heat shock response (Ritossa 1996).

CHLOROPLAST RESEARCH

In 1970 I joined the new Department of Biological Sciences at Warwick as the second appointment, and began the most productive part of my research career. SDS polyacrylamide gels had just been invented, and my plan was to form a Chloroplast Research Group who would use this technique, married to David Walker's methods for isolating pea leaf chloroplasts with their envelope membranes intact, to determine the function of the genetic system found inside these organelles.

Received 25 March 1996; Accepted 15 April 1996

Correspondence to: R. John Ellis, Tel: +44 1203 523509; Fax: +44 1203 523568; E-mail: je@dna.bio.warwick.ac.uk

Chloroplasts were known at that time to contain many copies of a small circular DNA genome, but more strikingly they also contain about 50% of all the ribosomes found inside leaf cells. This means that chloroplast ribosomes are the most abundant ribosomes in nature, so why are so many needed in this location? During the next 3 years my first PhD student at Warwick, Eric Blair, established a system in which intact isolated pea chloroplasts use light energy to incorporate labelled amino acids into proteins. The major soluble product of this light-driven protein synthesis was identified as the large subunit of the enyzme rubisco (ribulose bisphosphate carboxylase-oxygenase), the photosynthetic CO2-fixing enzyme of the Calvin cycle (Blair and Ellis 1973).

Plant rubisco has a miserably low turnover number, which probably accounts for this single enzyme comprising at least half of all the soluble protein in leaf cells, giving rise to the claim that this enzyme is the most abundant protein in the world (see Ellis 1979). The discovery that rubisco large subunits are made inside the chloroplast rather than in the cytosol offers a neat explanation for the high abundance of chloroplast ribosomes—so many are needed because one of their products is a component of the most abundant leaf protein. The other component of rubisco, called the small subunit, is by inference synthesized by cytosolic ribosomes and subsequently imported into the chloroplasts, a belief later substantiated by research by other members of the Chloroplast Research Group (see Ellis 1981). Meanwhile we turned our attention to the physical state of the rubisco large subunits made in isolated chloroplasts by analysing the products on native polyacrylamide gels containing no denaturing agent.

AN ERRONEOUS CONCLUSION

Much to our surprise the labelled rubisco large subunits were found not to have assembled into the holoenzyme of rubisco, but to migrate with another prominent staining band of protein, subsequently found to have an

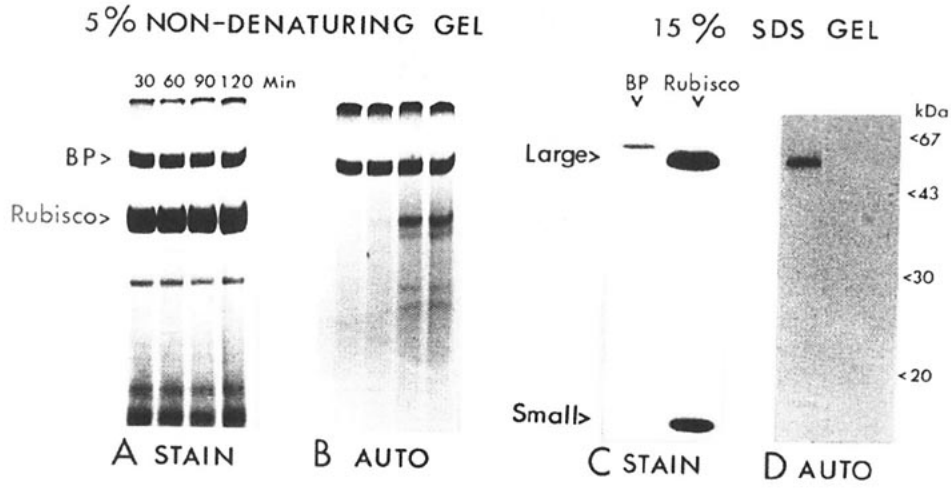


Figure Discovery of the chloroplast-binding protein (BP). Intact chloroplasts were isolated from young seedlings of *Pisum sativum* and illuminated at 20°C in a medium containing sorbitol as osmoticum and ^{35}S-methionine as labelled precursor; the illumination serves as an energy source for protein synthesis by ribosomes inside the intact chloroplasts. Samples were removed at intervals, the chloroplasts centrifuged down, lysed in hypotonic buffer, and the soluble fraction electrophoresed on a 5% non-denaturing polyacrylamide gel (A and B). The gel was stained in Coomassie blue (A) and an autoradiograph made (B). The stained bands of rubisco holoenzyme and BP were excised from the 30-min track and analysed separately on a 15% SDS polyacrylamide gel (C and D). The SDS gel was stained (C) and an autoradiograph made (D). Note that labelled rubisco large subunits comigrate exactly with the staining band of the binding protein (compare the precise shapes of the bands in A and B, especially at the 120-min time point). These large subunits can be visualized by their radioactivity but not by staining, since the chemical amount made in this system is very small (compare C and D). The binding protein oligomer (BP) is visible as a prominent stained band, as are its 60 kDa subunits, but these are not radioactive since they are made in the cytosol (compare C and D). BP, rubisco subunit binding protein; rubisco, holoenzyme of ribulose bisphosphate carboxylase-oxygenase; large and small, large and small subunits of rubisco, respectively. Reproduced from Johnson (1987) with kind permission.

apparent mass of about 700 kDa (Ellis 1977). Examination of many samples showed that the migration of the radioactively labelled large subunits with the staining protein was always exact. The conclusion, therefore, seemed obvious—the staining band of protein represents some oligomeric form of rubisco large subunit which the leaf cells accumulate before imported small subunits are added to assemble the rubisco holoenzyme. This conclusion was reached in 1973, and since it seemed an eminently plausible and thus rather dull conclusion, we turned our attention to other problems, with some success. Peter Highfield provided the first evidence that isolated chloroplasts can import newly synthesized rubisco small subunits post-translationally (Highfield and Ellis 1978), Annabel Wheeler reported the first in vitro translation of a mRNA for a plant enzyme—the rubisco large subunit (Hartley et al 1975)—while Martin Hartley characterized the precursor ribosomal RNA molecules synthesized by isolated chloroplasts (Hartley and Ellis 1973).

It was not until 1980 that a postdoc, Roger Barraclough, discovered the error in our earlier conclusion—the staining band we had noted is not an oligomeric form of rubisco large subunit, but a different protein that binds newly synthesized rubisco large subunits. The stoichiometry of binding is so low that it did not allow the complex of binding protein with labelled rubisco large subunits to be resolved from the bulk of the unlabelled binding protein, and so the complex migrates on native gels and sucrose density gradients with a mobility indistinguishable from the bulk of the staining band (Barraclough and Ellis 1980). With the hindsight provided by recent research on chaperonin 60, it is likely that in this complex one molecule of labelled rubisco large subunit (Mr 52 000) is bound to one molecule of binding protein (Mr 840 000). The Figure shows a more recent repetition by Richard Johnson of the experiment that revealed the existence of the binding protein.

PROPOSED ROLE OF THE BINDING PROTEIN

The identification of the rubisco subunit binding protein was the first report of a protein that binds the newly synthesized form of another protein. Analysis of the staining band on SDS gels revealed an apparent subunit mass of about 60 kDa, so the native binding protein was assumed to be composed of about 12 subunits. It was also noted that labelled rubisco large subunits bound to the binding protein are not precipitable by antiserum to rubisco, unlike labelled large subunits migrating with the rubisco holoenzyme. We concluded that the antigenic groups of the large subunits are masked in the complex, an interpretation that anticipated the current view that polypeptides bind within the central cavity of chaperonin 60. What could be the biological role of this binding protein?

Time-course experiments by Roger Barraclough showed that, when the incubation medium for the isolated chloroplasts is altered slightly and the incubation time prolonged, the amount of labelled large subunits bound to the binding protein declines, while the amount of labelled large subunits migrating with the rubisco holoenzyme increases (Barraclough and Ellis 1980). In addition, the chloroplasts are isolated from young rapidly growing leaves, and rubisco was known to be very stable in vivo until leaves senesce, so it did not seem likely that the binding protein formed part of a rubisco degradation mechanism. Another line of evidence available at the time was the report that plant rubisco large subunits isolated from the holoenzyme by the use of urea or SDS have a very strong tendency to form insoluble aggregates on removal of the denaturant. To this day it has proved impossible to reconstitute enzymically active plant rubisco by removing the denaturing agent from a denatured sample; this problem is peculiar to plant rubisco, since successful renaturation has been achieved with the rubisco from prokaryotic cells. For all these reasons it seemed plausible to propose that the complex of rubisco large subunit with the binding protein is an intermediate step in the assembly of plant rubisco, perhaps by keeping the large subunits in a soluble state suitable for binding to imported small subunits, which suffer from no such aggregation problem. The closing sentence of the paper of Barraclough and Ellis (1980) suggests the possibility that 'the aggregate of polypeptide 60 and large subunit is an obligatory intermediate in the assembly of ribulose bisphosphate carboxylase'.

A SPELL IN LIMBO

It is my belief that scientists should resist the natural tendency to ignore unexpected observations that do not fit the existing paradigm, but take the risk of pursuing them in the hope that they lead to new ideas and discoveries. I decided to concentrate on the binding protein. In the next 8 years, a number of postgraduates and postdocs worked on this protein from several plant species.

We never managed to obtain really convincing evidence that the binding protein is required for rubisco assembly in chloroplasts, and that is still the case today, but the protein was purified and characterized. Sean Hemmingsen joined my laboratory in 1981, and proved to be an enthusiastic and productive researcher. In the course of his work it was discovered that the protein consists of equal amounts of two closely migrating subunits called α and β (Hemmingsen and Ellis 1986) with different aminoterminal sequences (Musgrove et al 1987). The binding protein was also found to occur in all plastids examined, including the colourless plastids from tissues such as seed endosperm, which are not photosynthetic but neverthless contain large amounts of rubisco. An important observation made by Sean Hemmingsen was that antisera to the binding protein detect a band of about 60 kDa in extracts of not just photosynthetic bacteria, but also of bacteria such as *Escherichia coli.*

I recall that I was initially sceptical of the significance of this finding, since I was then wedded to the idea that the binding protein was nature's way of overcoming the aggregation properties of rubisco.

During the period 1980–1987, I presented our studies on the binding protein at many conferences of plant molecular biology and photosynthesis. Few animal or microbial biochemists attend such conferences so these studies did not percolate into these fields. The response to the idea that the binding protein is required for rubisco assembly was generally polite scepticism, which I attribute to the acceptance of the principle of protein self-assembly. Several commentators pointed out that the folding of newly synthesized polypeptide chains, and any subsequent association into oligomers, is a self-assembly process, requiring no macromolecules other than those in the assembled product—and so it said in the textbooks. Only one other laboratory pursued the binding protein in this time—that of Harry Roy. We were greatly encouraged by his confirmation of our observations, and by the extension to the demonstration that the transfer of labelled large subunit from the binding protein to rubisco holoenzyme in isolated chloroplasts requires ATP (Bloom et al 1983).

John Gray pointed out that what we were observing might be a binding artefact of no significance. This was a valid criticism, since in the complex a very small amount of a highly labelled protein is bound to a very large amount of unlabelled protein—ideal stoichiometry for an artefact! So I was desperate for precedent and was referred by a staff colleague familiar with the literature in animal biochemistry, Alan Colman, to a paper on the assembly of nucleosomes in extracts of Xenopus eggs (Laskey et al 1978). This paper used the term 'molecular chaperone' to describe the properties of a protein, nucleoplasmin, required for the assembly of nucleosome cores in egg extracts.

ORIGIN OF THE TERM MOLECULAR CHAPERONE

Ron Laskey and his colleagues were interested in the very rapid formation of nucleosomes that occurs when amphibian eggs are fertilized, so it was natural to ask whether isolated nucleosomes, after dissociation into their histone and DNA components by exposure to high salt concentrations, can be reassociated by lowering the salt concentration to intracellular levels. This experiment was a spectacular failure; addition of monomeric histones

to DNA at physiological ionic strength results in the rapid appearance of non-specific aggregates rather than nucleosomes. However, addition of small amounts of Xenopus egg homogenate prevents this aggregation and results in nucleosome formation. The active factor was purified from homogenates, characterized as an abundant acidic nuclear protein, and called nucleoplasmin. This protein binds to histones and thereby reduces their strong positive charge; addition of negatively charged DNA at physiological ionic strength then results by an unknown mechanism in the formation of nucleosomes.

Two important characteristics of the action of nucleoplasmin were important in the development of subsequent ideas. Firstly, nucleoplasmin is required only for nucleosome assembly—it is not a component of the assembled nucleosomes themselves. Secondly, nucleosomes can be assembled from histones and DNA in the absence of nucleoplasmin if the high salt concentration is reduced slowly by dialysis. Thus the role of nucleoplasmin is not to provide steric information for nucleosome assembly, but to reduce the positive charge of the strongly charged histone monomers, and so allow correct interactions with DNA to predominate over incorrect interactions. In the words of Laskey et al (1978): 'We suggest that the role of the protein we have purified is that of a "molecular chaperone" which prevents incorrect ionic interactions between histones and DNA.'

EXTENSION TO OTHER PROTEINS

The work on nucleoplasmin suggested to me that, if unassembled rubisco large subunits have a strong tendency to undergo incorrect interactions, perhaps the role of the binding protein is to prevent this from happening by masking the interactive surfaces involved. The suggestion that the binding protein could be regarded as a second example of a molecular chaperone was made at a Royal Society Discussion meeting on rubisco that I organized in 1985, and was subsequently published in the proceedings (Musgrove and Ellis 1986).

I initially thought that nucleoplasmin and the chloroplast-binding protein were special cases evolved to deal with certain oligomeric proteins whose assembly presents particular difficulty because of the propensity of their subunits to aggregate incorrectly. What prompted me to extend the chaperone idea further was the paper of Pelham (1986). This paper does not discuss either nucleoplasmin or the chloroplast-binding protein; instead it speculates that the heat shock 70 and 90 proteins are involved in a variety of protein assembly and disassembly processes occurring in cells under non-stress conditions. It occurred to me that all these ideas could be gathered together under the chaperone umbrella. Perhaps many different types of molecular chaperones exist, in which case the problem of incorrect interactions is not confined to particular proteins but is more widespread.

I presented this idea at the NATO Advanced Study Institute meeting on plant molecular biology organized by Diter von Wettstein in Copenhagen in June 1987. A representative from *Nature* was at that meeting and he encouraged me to write a News and Views article describing this more generalized concept of molecular chaperones. This article appeared in July 1987 with the following opening sentence: 'At a recent meeting I proposed the term "molecular chaperone" to describe a class of cellular proteins whose function is to ensure that the folding of certain other polypeptide chains and their assembly into oligomeric structures occur correctly' (Ellis 1987). Thus, the general concept of molecular chaperones was borne from my realization that several unrelated discoveries in biochemical research could be regarded as particular examples of a widespread, but hitherto unrecognized, cellular phenomenon.

DISCOVERY OF THE CHAPERONINS

Before returning to Canada in 1984, Sean Hemmingsen started to learn cloning methods so that the sequences of the two subunits of the chloroplast-binding protein could be determined. This work reached fruition in 1987, when he determined the sequence of the α subunit of the binding protein found in the colourless plastids of castor bean endosperm. Checking with the sequence databases revealed a high amino acid sequence similarity to the ams gene of *Escherichia coli*, a gene implicated in the control of stability of mRNA (Chanda et al 1985). Sean contacted H. F. Kung, the senior author on the Chanda paper, who told him about a related protein sequenced by Rick Young; this protein was the 65 kDa common bacterial antigen of *Mycobacterium leprae*. The common bacterial antigen, as the name implies, is an antigen found in all bacteria examined, and is the dominant antigen in human bacterial diseases such as tuberculosis and leprosy. This link was consistent with Sean Hemmingsen's earlier finding that antisera against the binding protein detect a 60 kDa protein in bacterial extracts. However, there was no information available about the function of the common bacterial antigen, so this similarity was puzzling for a time.

The breakthrough came near the end of 1987, when Rick Young informed Sean about a protein sequence determined by Roger Hendrix; this protein was GroEL from *Escherichia coli*. The resulting conversation between Sean and Roger left them both excited by the similarity (about 50% identity) of two proteins from different sources, both implicated in the assembly of other proteins. It emerged that Roger had sequenced the GroEL

protein of *Escherichia coli* some time before, but had not yet put the sequence into the database. Indeed, he had the draft of a manuscript describing the GroE operon on his desk. Sean relayed this news to me in the autumn of 1987, and I started to look up the literature on GroEL, a protein I was unaware of until then.

I still recall the growing excitement I felt as I realized that the reported properties of GroEL were very similar to those of the chloroplast-binding protein. Not only was GroEL a large oligomer of 14 65–kDa subunits (Hendrix 1979), but amazingly it was also implicated in a protein assembly process. GroEL was identified in several laboratories in the early 1970s as a bacterially-encoded protein required for several phages to replicate inside *Escherichia coli*, including phages lambda and T4 (Georgopoulos et al 1983). There was evidence that the GroEL oligomer binds transiently and non-covalently to subunits of phage lambda protein B; the complex is stable and can be detected on density gradients (Kochan and Murialdo 1983). This complex was believed to be a necessary intermediate in the formation of an oligomeric structure called the preconnector, made of 12 phage protein B subunits. Moreover a mutation in GroEL results in the head proteins of phage T4 forming insoluble aggregates that associate with the bacterial cell membrane (Takano and Kakefunda 1972). However, in the 1970s, most attention was paid to the role of GroEL in phage assembly rather than to its role in the uninfected cell. This was the period when research on the assembly of phages was in its heyday and there was little attention paid to what the normal role of GroEL might be.

Meanwhile, Saskia van der Vies, a PhD student in my laboratory, had sequenced a cDNA clone for the α subunit of the binding protein of wheat chloroplasts: the derived amino acid sequence is about 46% identical with that of GroEL and about 80% identical to that of the α subunit from castor bean. Sean and I realized that there was now evidence from two diverse sources (chloroplast and bacterial) linking the involvement of highly similar pre-existing proteins in the assembly of other protein structures in a manner that fitted the general concept of molecular chaperones. Sean had the inspiration to call these proteins 'chaperonins' and was able to convince Roger Hendrix and his colleagues of the merits of pooling their sequence data with ours and present the whole package to *Nature*. This paper appeared in May 1988 (Hemmingsen et al 1988), and sparked the continuing wave of research on the structure and function of the chaperonins.

One action I took before we had submitted the manuscript to *Nature* reflects the idealistic nature of the times, and is inconceivable in today's highly competitive climate. George Lorimer had long shown interest in my chloroplast work because of his contributions to the rubisco field. So I phoned to tell him of the connection between the chloroplast-binding protein and GroEL. I can still hear his Scottish brogue in my ear: 'GroEL—what is that?' He very rapidly realized the potential significance of this connection, and at the end of 1989 the first paper describing the effects of GroEL on the refolding of denatured bacterial rubisco was published (Goloubinoff et al 1989). The rest, as they say, is history. Some of this history is recounted in a recently published book about the chaperonins, to which the reader is referred for more recent information about chaperonin research (Ellis 1996).

Note added in proof

Tom Creighton has pointed out to me that the word 'chaperone' was used in a biochemical context prior to its use in 'molecular chaperone' by Laskey et al (1978). Purification of a toxin from the venom of the taipan snake revealed three equimolar and aminoterminal sequence-related subunits, only one of which displays potent neurotoxicity, but still less than that of the holotoxin (Fohlman et al 1976). The roles of the other subunits were not established, but these authors suggest that 'Probably they function as "chaperones" sharpening the specificity and increasing the stability of the toxic protein so as to minimize distraction and destruction en route to the proper site of action.' These authors did not extend the term to other systems.

REFERENCES

Barraclough R and Ellis RJ (1980) Protein synthesis in chloroplasts IX. Assembly of newly synthesized large subunits into ribulose bisphosphate carboxylase in isolated intact pea chloroplasts. Biochim. Biophys. Acta **608**, 19–31.

Blair GE and Ellis RJ (1973) Protein synthesis in chloroplasts I. Light-driven synthesis of the large subunit of fraction I protein by isolated pea chloroplasts. Biochim. Biophys. Acta **319**, 223–234.

Bloom MV, Milos P and Roy H (1983) Light-dependent assembly of ribulose 1, 5-bisphosphate carboxylase. Proc. Natl Acad. Sci. USA **80**, 1013–1017.

Chanda PK, Ono M, Kuwano M and Kung HF (1985) Cloning, sequence, analysis and expression of alteration of the mRNA stability gene (ams+) of *Escherichia coli*. J. Bacteriol. **161**, 446–449.

Ellis RJ (1977) Protein synthesis by isolated chloroplasts. Biochim. Biophys. Acta **463**, 185–215.

Ellis RJ (1979) The most abundant protein in the world. Trends Biochem. Sci. **4**, 241–244.

Ellis RJ (1981) Chloroplast proteins: synthesis, transport and assembly. Annu. Rev. Plant Physiol. **32**, 111–137.

Ellis RJ (1987) Proteins as molecular chaperones. Nature **328**, 378–379.

Ellis RJ (1996) Chaperonins: introductory perspective. In: The Chaperonins, ed RJ Ellis, Academic Press, Orlando, 1–25.

Fohlman J, Eaker D, Karlsson E and Thesleff S (1976) Taipoxin, an extremely potent presynaptic neurotoxin from the venom of the Australian snake taipan. Eur. J. Biochem. **68**, 457–469.

Georgopolous CP, Tilly K and Casgens SR (1983) In: Lambda II, eds RW Hendrix, JW Roberts, FW Stahl and RA Weisberg, Cold Spring Harbor Laboratory Press, New York, 279–304.

Goloubinoff P, Christeller JT, Gatenby AA and Lorimer GC (1989) Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. Nature **342**, 884–889.

Hartley MR and Ellis RJ (1973) Ribonucleic acid synthesis in chloroplasts. Biochem. J. **134**, 249–262.

Hartley MR, Wheeler A and Ellis RJ (1975) Protein synthesis in chloroplasts. V. Translation of the messenger RNA for the large subunit of fraction I protein in a heterologous cell-free system. J. Mol. Biol. **91**, 67–77.

Hemmingsen SM and Ellis RJ (1986) Purification and properties of ribulose bisphosphate carboxylase large subunit binding protein. Plant Physiol. **80**, 269–276.

Hemmingsen SM, Woolford C, van der Vies SM et al (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. Nature **333**, 330–334.

Hendrix RW (1979) Purification and properties of GroEL, a host protein involved in bacteriophage assembly. J. Mol. Biol. **129**, 375–392.

Highfield PE and Ellis RJ (1978) Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase. Nature **271**, 420–424.

Johnson RA (1987) PhD thesis, University of Warwick UK.

Kochan JA and Murialdo H (1983) Early intermediates in bacteriophage lambda prohead assembly. Virology **131**, 100–115.

Laskey RA, Honda BM, Mills AD and Finch JT (1978) Nucleosomes are assembled by an acidic protein that binds histones and transfers them to DNA. Nature **275**, 416–420.

Musgrove JE and Ellis RJ (1986) The rubisco large subunit binding protein. Phi. Trans. Roy. Soc. Lond. B **313**, 419–428.

Musgrove JE, Johnson RA and Ellis RJ (1987) Dissociation of the ribulose bisphosphate carboxylase large subunit binding protein into dissimilar subunits. Eur. J. Biochem. **163**, 529–534.

Pelham HRB (1986) Speculations on the functions of the major heat shock and glucose-regulated proteins. Cell **46**, 956–961.

Ritossa F (1996) Reflections: discovery of the heat shock response. Cell Stress Chap. **1(2)**, 97–98.

Takano T and Kakefuda T (1972) Involvement of a bacterial factor in the morphogenesis of bacteriophage capsid. Nature New Biol. **239**, 34–37.

REVIEW: PROTEIN FOLDING

Molecular Chaperones in the Cytosol: from Nascent Chain to Folded Protein

F. Ulrich Hartl* and Manajit Hayer-Hartl

Efficient folding of many newly synthesized proteins depends on assistance from molecular chaperones, which serve to prevent protein misfolding and aggregation in the crowded environment of the cell. Nascent chain–binding chaperones, including trigger factor, Hsp70, and prefoldin, stabilize elongating chains on ribosomes in a nonaggregated state. Folding in the cytosol is achieved either on controlled chain release from these factors or after transfer of newly synthesized proteins to downstream chaperones, such as the chaperonins. These are large, cylindrical complexes that provide a central compartment for a single protein chain to fold unimpaired by aggregation. Understanding how the thousands of different proteins synthesized in a cell use this chaperone machinery has profound implications for biotechnology and medicine.

To become functionally active, newly synthesized protein chains must fold to unique three-dimensional structures. How this is accomplished remains a fundamental problem in biology. Although it is firmly established from refolding experiments in vitro that the native fold of a protein is encoded in its amino acid sequence (*1*), protein folding inside cells is not generally a spontaneous process. Evidence accumulated over the last decade indicates that many newly synthesized proteins require a complex cellular machinery of molecular chaperones and the input of metabolic energy to reach their native states efficiently (*2*–*5*). The various chaperone factors protect nonnative protein chains from misfolding and aggregation, but do not contribute conformational information to the folding process. Here we focus on recent advances in our mechanistic understanding of de novo protein folding in the cytosol and seek to provide a coherent view of the overall flux of newly synthesized proteins through the chaperone system.

Protein Aggregation

Spontaneous refolding in vitro is generally efficient for small, single-domain proteins that bury exposed hydrophobic amino acid residues rapidly (within milliseconds) upon initiation of folding (*1*). In contrast, larger proteins composed of multiple domains often refold inefficiently, owing to the formation of partially folded intermediates, including misfolded states, that tend to aggregate (Fig. 1). Misfolding originates from interactions between regions of the folding polypeptide chain that are separate in the native protein and that may be stable enough to prevent folding from proceeding at a biologically relevant time scale. These nonnative states, though compact in shape, often expose hydrophobic amino acid residues and segments of unstructured polypeptide backbone to the solvent. They readily self-associate into disordered complexes (Fig. 1), driven by hydrophobic forces and interchain hydrogen bonding (*1*, *6*). This aggregation process irreversibly removes proteins from their productive folding pathways, and must be prevented in vivo by molecular chaperones. A certain level of protein aggregation does occur in cells despite the presence of an exclusive chaperone machinery and, in special cases, can lead to the formation of structured, fibrillar aggregates, known as amyloid, that are associated with diseases such as Alzheimer's or Huntington's disease (*6*, *7*) (Fig. 1). Compared to refolding in dilute solution, the tendency of nonnative states to aggregate in the cell is expected to be sharply increased as a result of the high local concentration of nascent chains in polyribosomes and the added effect of macromolecular crowding.

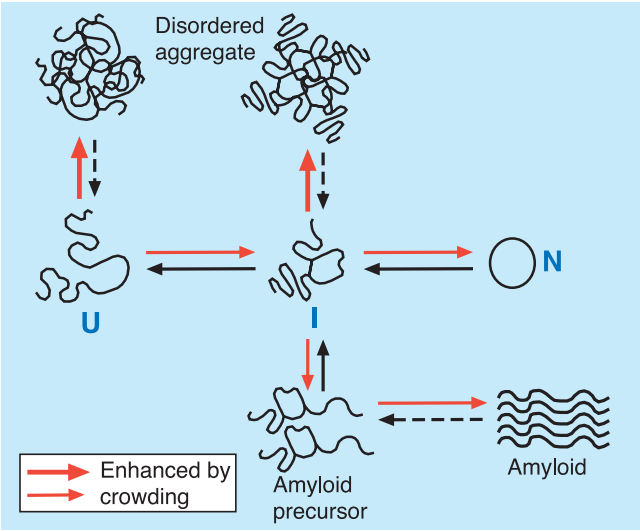


Fig. 1. Aggregation of nonnative protein chains as a side-reaction of productive folding in the crowded environment of the cell. Enhancement of aggregation and chain compaction by macromolecular crowding (red arrows). U, unfolded protein chain released from ribosome; I, partially folded intermediate; N, native, folded protein. Crowding is predicted to enhance the formation of amyloid fibrils, but this effect has not yet been demonstrated experimentally. [Adapted from (*1*)]

Nascent chains. During translation, the folding information encoded in the amino acid sequence becomes available in a vectorial fashion. The polypeptide exit channel in the large ribosomal subunit is 100 Å long, a distance spanned by an extended chain of ~30 amino acid residues or an α helix of 65 residues (*8*). The channel is on average only 15 Å wide and is expected to prohibit folding beyond helix formation inside the ribosome, unless the tunnel is conformationally dynamic. Because the formation of stable tertiary structure is a cooperative process at the level of protein domains (50 to 300 amino acid residues), an average domain can complete folding only when its entire sequence has emerged from the ribosome. It takes more than a minute to synthesize a 300-residue protein in eukaryotes. As a consequence, many nascent chains expose non-native features for a considerable length of time and are prone to aggregation. This tendency to aggregate is thought to be greatly increased by the close proximity of nascent chains of the same type

Department of Cellular Biochemistry, Max-Planck-Institut für Biochemie, Am Klopferspitz 18A, D-82152 Martinsried, Germany.

*To whom correspondence should be addressed. E-mail: uhartl@biochem.mpg.de

in polyribosome complexes (*5*), thus leading to the requirement for chaperones to maintain nascent chains in a nonaggregated, folding-competent conformation.

Macromolecular crowding. The excluded volume effects resulting from the highly crowded nature of the cytosol (300 to 400 g/liter of proteins and other macromolecules in *Escherichia coli*) (*9*) are predicted to enhance the aggregation of nonnative protein chains substantially by increasing their effective concentrations (*10*) (Fig. 1). Crowding generally provides a nonspecific force for macromolecular compaction and association (*11*), including the collapse of protein chains during folding (*9*) and the interaction of nonnative proteins with molecular chaperones (*12*).

How Chaperones Prevent Aggregation

The cellular chaperone machinery counteracts the aggregation of nonnative proteins, both during de novo folding and under conditions of stress, such as high temperature, when some native proteins unfold. Many chaperones, though constitutively expressed, are synthesized at greatly increased levels under stress conditions and are classified as stress proteins or heat-shock proteins (Hsps) (*3*). In general, all these chaperones recognize hydrophobic residues and/or unstructured backbone regions in their substrates, i.e., structural features typically exposed by nonnative proteins but normally buried upon completion of folding. Chaperones that participate broadly in de novo protein folding, such as the Hsp70s and the chaperonins, promote the folding process through cycles of substrate binding and release regulated by their adenosine triphosphatase (ATPase) activity and by cofactor proteins. Chaperone binding may not only block intermolecular aggregation directly by shielding the interactive surfaces of non-native polypeptides, including unassembled protein subunits, but may also prevent or reverse intramolecular misfolding. Certain chaperones of the Hsp100 or Clp family even have the ability to unfold proteins or to disrupt small-protein aggregates by an adenosine 5′-triphosphate (ATP)–dependent mechanism (*13*). For a growing number of proteins, chaperone function is combined with an additional activity, as is the case for certain protein disulfide isomerases and peptidyl-prolyl isomerases, enzymes that catalyze rate-limiting steps in the folding of some proteins (*14*).

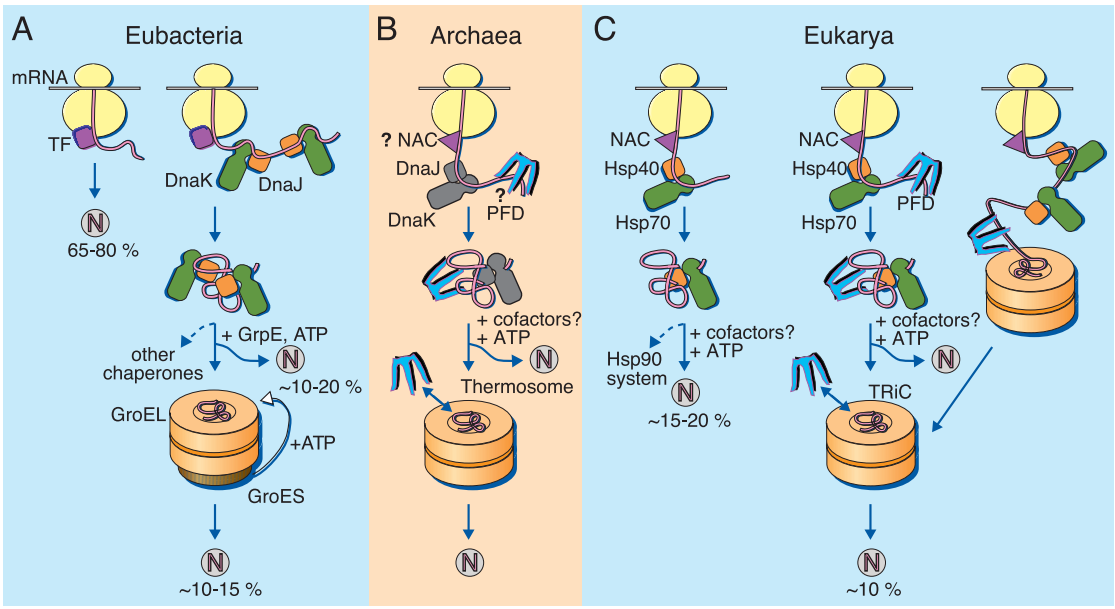


Fig. 2. Models for the chaperone-assisted folding of newly synthesized polypeptides in the cytosol. (**A**) Eubacteria. TF, trigger factor; N, native protein. Nascent chains probably interact generally with TF, and most small proteins (~65 to 80% of total) fold rapidly upon synthesis without further assistance. Longer chains (10 to 20% of total) interact subsequently with DnaK and DnaJ and fold upon one or several cycles of ATP-dependent binding and release. About 10 to 15% of chains transit the chaperonin system—GroEL and GroES—for folding. GroEL does not bind to nascent chains and is thus likely to receive an appreciable fraction of its substrates after their interaction with DnaK. (**B**) Archaea. PFD, prefoldin; NAC, nascent chain–associated complex. Only some archaeal species contain DnaK/DnaJ. The existence of a ribosome-bound NAC homolog, as well as the interaction of PFD with nascent chains, has not yet been confirmed experimentally. (**C**) Eukarya—the example of the mammalian cytosol. Like TF, NAC probably interacts generally with nascent chains. The majority of small chains may fold upon ribosome release without further assistance. About 15 to 20% of chains reach their native states in a reaction assisted by Hsp70 and Hsp40, and a fraction of these must be transferred to Hsp90 for folding. About 10% of chains are co- or posttranslationally passed on to the chaperonin TRiC in a reaction mediated by PFD.

Protein Flux Through the Chaperone System

Cytosolic chaperones participate in de novo folding mainly through two distinct mechanisms. Chaperones, such as trigger factor and the Hsp70s, act by holding nascent and newly synthesized chains in a state competent for folding upon release into the medium. In contrast, the large, cylindrical chaperonin complexes provide physically defined compartments inside which a complete protein or a protein domain can fold while being sequestered from the cytosol. These two classes of chaperone are conserved in all three domains of life and can cooperate in a topologically and timely ordered manner (*15*–*17*) (Fig. 2, A to C).

Although the essential nature of the chaperonins has long been recognized (*18*, *19*), it has proved more difficult to establish the essential role of nascent chain-binding chaperones in protein folding, because of considerable functional redundancy among components (*20*, *21*). Some of these chaperones, including trigger factor and specialized Hsp70 proteins, bind directly to the ribosome near the polypeptide exit site and are positioned to interact generally with nascent chains (Fig. 2). The majority of small proteins are thought to fold rapidly and without further assistance upon completion of synthesis and release from this first set of components (Fig. 2A). Longer chains interact subsequently with members of a second class of nascent chain-binding chaperones, including the classical Hsp70s and prefoldin, which do not associate directly with the ribosome (*20*–*22*). In addition to stabilizing elongating chains, these chaperones also assist in co- or posttranslational folding, or facilitate chain transfer to downstream chaperones (Fig. 2, A and C) (*17*, *20*, *21*). A subset of slow-folding and aggregation-sensitive proteins (10 to 15% of total) interact with a chaperonin for folding in both prokaryotes and eukaryotes (*22*–*24*). Many eukaryotic kinases and other signal-transduction proteins use an additional chaperone pathway from Hsp70 to Hsp90 (Fig. 2C), a specialized ATP-dependent chaperone that cooperates with ancillary factors in protein folding and regulation. [For a detailed discussion of the Hsp90 system, see (*25*, *26*).]

Ribosome-Binding Chaperones

Trigger factor (TF), a eubacterial protein of 48 kD, binds to ribosomes at a 1:1 stochiometry and interacts with nascent chains as short as 57 residues (*27*). The nascent chain–TF complex dissociates, in an ATP-independent manner, after chain release from the ribosome (*27*). Although TF exhibits peptidyl-prolyl cis/trans isomerase (PPIase) activity in vitro, recognition of target polypeptides by TF is independent of proline residues (*28*) and is mediated by short sequences enriched in hydrophobic (aromatic) amino acids (*28*). TF has an overlapping chaperone function with the main bacterial Hsp70 system, DnaK and DnaJ, in stabilizing nascent chains in a state competent for subsequent folding (*20*, *21*). *E. coli* cells lacking TF (Δ*tig*) or DnaK (Δ*dnak*) exhibit no apparent folding defects at 37°C; however, deletion of *dnaK* in a Δ*tig* strain is lethal. In light of this functional redundancy, the biological significance of the PPIase activity of TF remained unclear, but a recent study suggests that DnaK has a related activity in accelerating the cis/trans isomerization of nonprolyl peptide bonds (*29*). These isomerase activities may allow TF and the Hsp70s to maintain nascent and newly synthesized chains in a flexible state, poised for rapid folding upon release. In contrast to DnaK, a role of TF in mediating folding posttranslationally has not yet been demonstrated, but would be consistent with the finding that only half of total TF is ribosome bound (*30*)

The eukaryotic cytosol lacks TF but contains a ribosome-associated heterodimeric complex of α (33 kD) and β (22 kD) subunits, termed NAC (nascent chain–associated complex) (Fig. 2C) (*31*). A homolog of α-NAC appears to be present in some archaea (*32*). Although NAC lacks a PPIase domain, it has properties that suggest a functional similarity to TF. NAC associates with short nascent chains and dissociates upon chain release from the ribosome (*4*, *33*). However, a direct role for NAC in protein folding remains to be established.

Whereas the Hsp70 proteins in bacteria and higher eukaryotes act both co- and posttranslationally (see below), yeast and other fungi have cytosolic Hsp70 homologs that are specialized in nascent chain binding. The Ssb1 and Ssb2 proteins in the yeast *Saccharomyces cerevisiae* interact with the ribosome and with short nascent chains (*34*). Interestingly, this function of the Ssb proteins appears to be mediated by yet another Hsp70, Ssz1, which forms a stable ribosome-associated complex (RAC) with zuotin (*35*, *36*), the Hsp40 partner of Ssb1 and Ssb2 (*30*). RAC and the Ssb proteins are thought to act in concert in stabilizing nascent chains.

The Hsp70 System

The classic, nonribosome-binding members of the Hsp70 family exist in the cytosol of eubacteria, eukarya, and some archaea, as well as within eukaryotic organelles, such as mitochondria and endoplasmic reticulum. *S. cerevisiae* has four nonribosome-binding Hsp70 proteins in the cytosol, namely, Ssa1 to Ssa4. The cytosol of higher eukaryotes contains both constitutively expressed Hsp70 homologs (Hsc70) and stress-inducible forms (Hsp70). Together with cochaperones of the Hsp40 (DnaJ) family, these Hsp70s function by binding and releasing, in an ATP-dependent manner, extended polypeptide segments that are exposed by proteins in their non-native states.

Structure and reaction cycle. The structural and mechanistic aspects of the Hsp70 system are best understood for the eubacterial Hsp70, termed DnaK, its Hsp40 cochaperone, DnaJ, and the nucleotide exchange factor GrpE. DnaK consists of a ~44-kD NH_2-terminal ATPase domain and a ~27-kD COOH-terminal peptide-binding domain (*37*) (Fig. 3A). The latter is divided into a β-sandwhich subdomain with a peptide-binding cleft and an α-helical latchlike segment (*38*). Target peptides are ~seven residues long and are typically hydrophobic in their central region, with leucine and isoleucine residues being preferred by DnaK (*4*, *39*) (Fig. 3A). These binding sites occur statistically every ~40 residues in proteins and are recognized with affinities of 5 nM to 5 μM (*37*). The peptides are bound to DnaK in an extended state through hydrophobic side-chain interactions and hydrogen bonds with

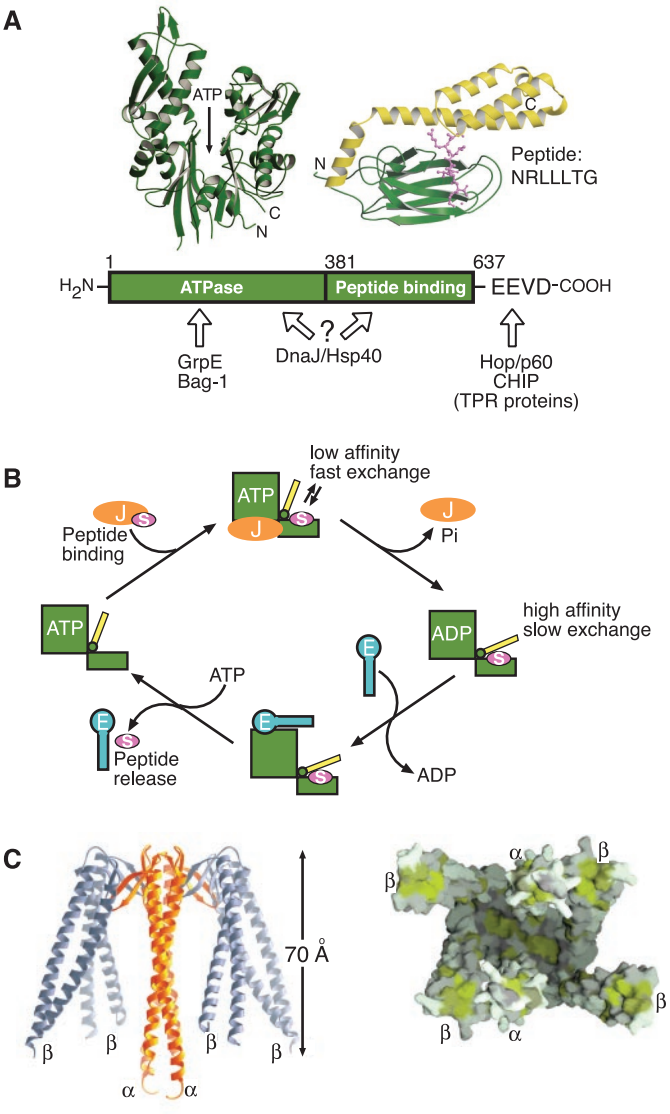


Fig. 3. Structure and function of chaperones with the ability to bind nascent chains. (**A**) (Top) Structures of the ATPase domain (*40*) and the peptide-binding domain (*38*) of Hsp70 shown representatively for *E. coli* DnaK, generated with the program MOLSCRIPT (*87*). The α-helical latch of the peptide binding domain is shown in yellow and a ball-and-stick model of the extended peptide substrate in pink. ATP indicates the position of the nucleotide binding site. The amino acid sequence of the peptide is indicated in single-letter code (D, Asp; E, Glu; G, Gly; L, Leu; N, Asn; R, Arg; T, Thr; and V, Val). (Bottom) The interaction of prokaryotic and eukaryotic cofactors with Hsp70 is shown schematically. Residue numbers refer to human Hsp70. Only the Hsp70 proteins of the eukaryotic cytosol have the COOH-terminal sequence EEVD that is involved in binding of tetratricopeptide repeat (TPR) cofactors. (**B**) Simplified reaction cycle of the DnaK system with DnaK colored as in (A). J, DnaJ; E, GrpE; S, substrate peptide. GrpE is drawn to reflect the extended shape of the protein. Not all substrates are presented to DnaK by DnaJ. The intermediate DnaK-DnaJ-substrate-ATP is probably very transient, as this is the fastest step of the cycle. (**C**) (Left) Side view and dimension of the structure of achaeal PFD with the two α subunits shown in gold and the four β subunits in gray. (Right) Bottom view of the PFD complex showing the central space enclosed by the six coiled-coil segments. Surface representation is shown with hydrophobic patches in yellow and hydrophilic regions in gray [reproduced from (*54*) with permission].

the peptide backbone (*38*). Thus, Hsp70 recognizes structural features common to most nascent chains: exposed hydrophobic amino acid side chains, in conjunction with an accessible polypeptide backbone.

Rapid peptide binding occurs in the ATP-bound state of DnaK in which the α-helical latch over the peptide-binding cleft is in an open conformation (Fig. 3B). Stable holding of peptide involves closing of the latch, a conformational change that is achieved by hydrolysis of bound ATP to adenosine 5′-diphosphate (ADP). The cycling of DnaK between these states is regulated by DnaJ (41 kD) and by GrpE, a homodimer of 20-kD subunits (*37*, *40*). The NH_2-terminal J domain of DnaJ binds to DnaK and accelerates hydrolysis of ATP by DnaK, thus facilitating peptide capture (*41*, *42*). The COOH-terminal domain of DnaJ (and of other Hsp40s) functions as a chaperone in recognizing hydrophobic peptides and can thus recruit DnaK to nascent chains (*15*, *43*, *44*). GrpE induces the release of ADP from DnaK (*40*), and upon rebinding of ATP the DnaK-peptide complex dissociates, completing the reaction cycle (Fig. 3B).

Some eukaryotic Hsp70 homologs, such as BiP in the endoplasmic reticulum, cooperate with J-domain proteins that lack a separate affinity for hydrophobic sequences. These Hsp70s may be able to bind extended polypeptide chains more generally, independent of exposed hydrophobic features (*45*). Whereas all Hsp70s seem to cooperate with J proteins, most eukaryotic Hsp70 proteins may be independent, for their general function, of a GrpE-like nucleotide exchange factor. Such a factor appears to be dispensable because the rate-limiting step in the ATPase cycle of eukaryotic Hsp70 is normally not the dissociation of bound ADP, but ATP hydrolysis itself. On the other hand, a small protein, Bag-1, acts as a nucleotide exchange factor and specific regulator of Hsp70 in the eukaryotic cytosol (*46*). The Hsp70-interacting Bag domain is structurally unrelated to GrpE (*40*, *47*) and in Bag-1 is linked to an NH_2-terminal ubiquitin-homology domain (see below).

Substrates and mechanism of folding. The cellular concentration of DnaK (~50 μM) exceeds that of ribosomes (~30 μM) (*4*), assuming an even cytosolic distribution. DnaK preferentially associates with elongating polypeptides larger than 20 to 30 kD and thus acts on nascent chains subsequent to TF (*21*) (Fig. 2A). Upon deletion of TF, the fraction of nascent and newly synthesized polypeptides interacting with DnaK increases from ~15 to ~40% (*20*, *21*). Whereas some chains transit DnaK with half-lives of less than 1 min, consistent with rapid folding upon completion of synthesis, other newly synthesized proteins are released from DnaK slowly, with half-lives of 10 min or more. Large proteins >60 kD, which do not fit into the central cavity of the chaperonin GroEL (see below), constitute an appreciable fraction of these substrates, suggesting that DnaK facilitates the posttranslational folding of multidomain proteins through cycles of binding and release (*21*). Consistent with this conclusion, depletion of DnaK in TF-deleted cells causes the aggregation of many large, newly synthesized polypeptides (*20*). Similar to DnaK, mammalian Hsc70 also binds a wide range of nascent and newly synthesized chains (>15 to 20% of total) (Fig. 2C), including many multidomain proteins >50 kD (*22*).

How do cycles of Hsp70 binding and release promote protein folding? Generally, on release from Hsp70, an unfolded chain is free to partition to its native state. Rebinding of slow-folding intermediates to Hsp70 follows this release and prevents aggregation. Assuming that for long protein chains cycling is mediated by multiple Hsp70 molecules at the level of individual domains, the Hsp70 system could promote the folding of multidomain proteins by preventing (and perhaps reversing) intramolecular misfolding. The recently discovered isomerase activity of Hsp70 for nonprolyl peptide bonds may support this function (*29*). Consistent with this model, the Hsp70 system strongly accelerates the slow, spontaneous refolding of chemically denatured firefly luciferase (~60 kD) in vitro (*48*, *49*). The enzyme ornithine transcarbamylase accumulates in a misfolded but soluble form in vivo when expressed in Hsp70 (Ssa)-deficient yeast (*50*).

Surprisingly, the components of the Hsp70 system are missing in certain species of archaea (*32*). How these cells protect nascent and newly synthesized polypeptides from aggregating is not yet clear, but a candidate chaperone for nascent chains in archaea is prefoldin.

Prefoldin

Prefoldin (PFD) (*51*), also known as the Gim complex (genes involved in microtubule biogenesis) (*52*), is a ~90-kD complex of two α and four β subunits in the archaeal and eukaryotic cytosol. The eukaryotic α and β subunits are not identical but orthologous (*53*). The structure of PFD resembles that of a jellyfish, with six α-helical coiled-coil tentacles emanating from a β-barrel body (Fig. 3C). At the tips these ~65 Å long coiled coils are partially unwound, exposing hydrophobic amino acid residues for the binding of nonnative protein (*54*) (Fig. 3C). Substrate binding and release by PFD is ATP independent, and in vitro, mammalian and archaeal PFD can stabilize nonnative proteins for subsequent transfer to a chaperonin (*51*, *53*). PFD binds to nascent chains (*55*, *56*) and cooperates in the folding of actin and tubulin with the eukaryotic chaperonin (*17*). Interestingly, a combined deletion of the Ssb-class Hsp70s and of PFD in yeast results in a pronounced synthetic growth defect (*56*), resembling the synthetically lethal phenotype of the TF and DnaK deletions in *E. coli* (*20*, *21*). These findings underscore the functional redundancy among nascent chain–binding chaperones and suggest that PFD may have a DnaK or TF-like role in the archaeal cytosol.

The Chaperonins

The chaperonins are a conserved class of large double-ring complexes of ~800 kD enclosing a central cavity. They occur in two subgroups that are similar in architecture but distantly related in sequence. Group I chaperonins, also known as Hsp60s, are generally found only in eubacteria and in organelles of endosymbiotic origin—mitochondria and chloroplasts. They cooperate with cofactors of the GroES or Hsp10 family. Group II chaperonins exist in the archaeal and the eukaryotic cytosol and are GroES independent. The chaperonin mechanism differs fundamentally from that of the Hsp70 system, although in both cases protein binding and release is ATP regulated. Nonnative substrate protein is first captured through hydrophobic contacts with multiple chaperonin subunits and is then displaced into the central ring cavity where it folds, protected from aggregating with other nonnative proteins.

Group I chaperonins—structure and reaction cycle. E. coli GroEL and its cofactor GroES represent the paradigmatic Group I chaperonin system. In GroEL, two heptameric rings of identical 57-kD subunits are stacked back-to-back. Each subunit consists of three domains: The equatorial domain harbors the ATP binding site and is connected through an intermediate, hingelike domain to the apical domain (Fig. 4A). The latter makes up the opening of the cylinder and exposes a number of hydrophobic residues toward the ring cavity for substrate binding. GroES is a homoheptameric ring of ~10-kD subunits that cycles on and off the ends of the GroEL cylinder, in a manner regulated by the GroEL ATPase (*4*, *37*, *57*) (Fig. 4A).

The hydrophobic surfaces exposed by the apical domains (Fig. 4A) interact with hydrophobic amino acid residues on compact folding intermediates (*4*, *37*, *57*). Hydrophobic sequences bind to a flexible groove between two amphiphilic helices in the apical domain. This region can accommodate a peptide either as a β hairpin or an amphiphilic α-helical conformation (*58*, *59*). Stable substrate binding with nanomolar affinity relies on the interaction of a nonnative polypeptide with multiple apical domains (*60*). The GroES subunits have mobile sequence loops that contact the substrate-binding regions in the apical domains of GroEL and

mediate substrate dissociation (*37*, *57*, *61*, *62*).

GroEL is functionally asymmetrical; the two rings are coupled by negative allostery and do not occur in the same nucleotide-bound state. The chaperonin reaction begins by the binding of substrate polypeptide to the free end (i.e., the trans ring) of a GroEL-GroES complex (Fig. 4B). This step is closely followed by the binding of seven ATP molecules and GroES, resulting in the displacement of substrate into a GroES-capped cavity and causing the dissociation of the seven ADP molecules and GroES from the former cis complex. Upon binding to GroES, the apical domains undergo a massive rotation and upward movement (*61*, *63*), resulting in an enlargement of the cavity and a shift in its surface properties from hydrophobic to hydrophilic (Fig. 4A). Non-native proteins up to ~60 kD can be encapsulated and are free to fold in the resulting GroEL-GroES cage (also termed "Anfinsen cage") (*64*–*67*). Folding is allowed to proceed for ~10 s, timed by the hydrolysis of the seven ATP molecules in the cis ring. Upon completion of hydrolysis, binding of seven ATP molecules to the trans ring triggers the opening of the cage. Both folded and nonnative protein exit at this point (Fig. 4B), but folding intermediates that still expose extensive hydrophobic surfaces are rapidly recaptured and folding cycles are repeated until the protein reaches its native state. Oligomeric assembly occurs in solution after subunit folding inside the cage.

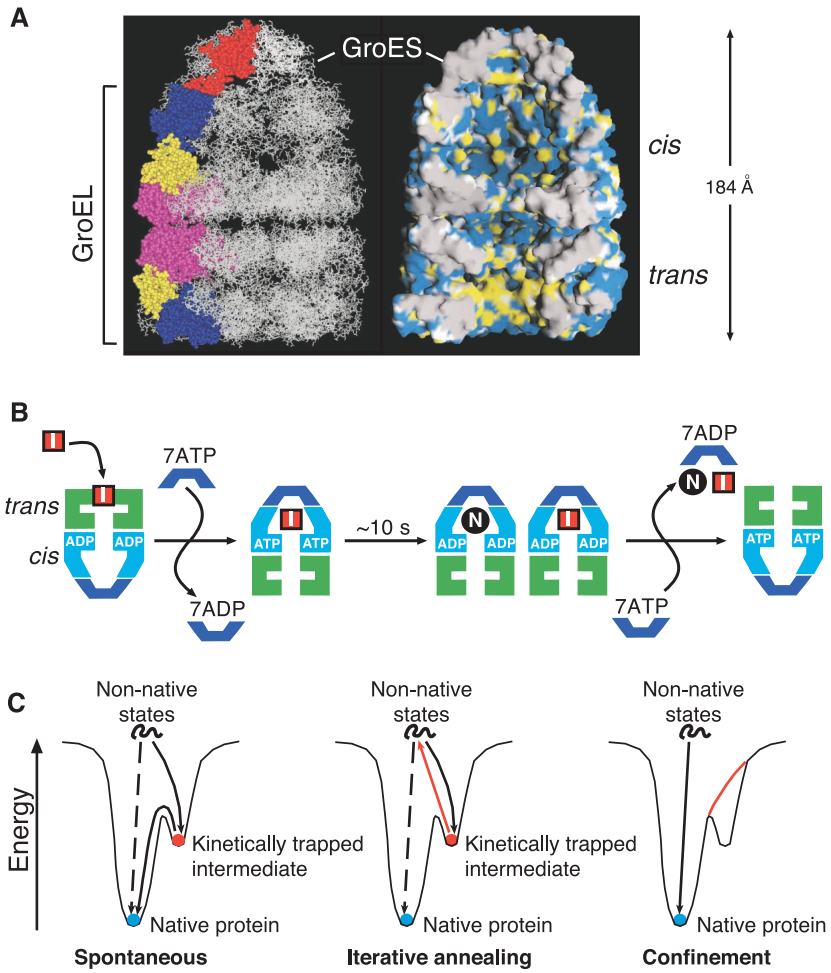


Fig. 4. The GroEL-GroES chaperonin system. (**A**) (Left) View of the asymmetric GroEL-GroES-$(ADP)_7$ complex generated with the coordinates 1AON (*61*) and program Weblab ViewLite 4.0 (Molecular Simulations). The equatorial, intermediate, and apical domains of one subunit each in the cis and trans ring of GroEL are colored in pink, yellow, and dark blue, respectively, and one subunit of GroES is colored red. (Right) The accessible surface of the central cavity of the GroEL-GroES complex. Polar and charged side-chain atoms, blue; hydrophobic side-chain atoms, yellow; backbone atoms, white; and solvent-excluded surfaces at subunit interfaces, gray. [Reprinted from (*61*) with permission] (**B**) Simplified reaction of protein folding in the GroEL-GroES cage. I, folding intermediate bound by the apical domains of GroEL; N, native protein folded inside the cage. For a typical GroEL substrate, multiple rounds of chaperonin action are required for folding; both I and N accumulate after a single reaction cycle and exit the cage upon GroES dissociation. I is then rapidly re-bound by GroEL. (**C**) Mechanisms of accelerated folding. Simple energy diagrams are shown for a protein that forms a kinetically trapped intermediate during spontaneous folding (left). In the iterative annealing model, this intermediate is thought to be actively unfolded by GroEL/GroES (*69*) and allowed to repartition (middle), whereas confinement of nonnative protein in the narrow, hydrophilic environment of the GroEL-GroES cage is suggested to result in a smoothing of the energy landscape (right), such that formation of certain trapped intermediates is avoided (*67*). Both proposed mechanisms would result in accelerated folding.

Substrates and folding mechanisms. About 10% of newly synthesized polypeptides normally transit GroEL posttranslationally (*23*, *24*), consistent with the cytosolic concentration of GroEL (~3 μM) relative to that of ribosomes (~30 μM). Most of these proteins are between 20 and 60 kD in size and leave GroEL with half-lives between 15 s and several minutes (*23*). The essential nature of GroEL and GroES (*18*) may be explained in principle by the existence of at least one essential *E. coli* protein with an absolute chaperonin dependence. Alternatively, loss of chaperonin could be lethal because it results in a reduced efficiency, rather than a complete loss, of folding for many proteins. Analysis of the highly streamlined proteome of *Ureaplasma urealyticum* (*68*), the first eubacterium that lacks GroEL and GroES, may now offer an opportunity to test these hypotheses.

About 50 proteins interacting with GroEL in the *E. coli* cytosol have been identified, and many of them contain two or more domains with α/β folds (*24*). Proteins with such complex topologies often fold slowly and are aggregation prone, owing to the exposure of extensive hydrophobic surfaces in their non-native states. Stringent model substrates of GroEL, such as bacterial RuBisCo (ribulose-1,5-bisphosphate carboxylase-oxygenase), share this domain topology and fold efficiently only when in the GroEL-GroES cage (*67*).

In addition to preventing aggregation during folding, encapsulation of nonnative RuBisCo (50 kD) in the hydrophilic cage speeds up the folding reaction substantially (*67*). Confinement in the cage may smooth the energy landscape of folding for some larger proteins, either by preventing the formation of certain kinetically trapped intermediates or by facilitating their progression toward the compact, native state (Fig. 4C). This acceleration of RuBisCo folding had previously been attributed to a mechanism of "iterative annealing" (*57*, *69*), rather than to an effect of confinement. In this alternative model the chaperonin is suggested to facilitate folding by cycles of unfolding kinetically trapped states, followed by repartitioning of the unfolded protein between productive and nonproductive folding pathways (*57*) (Fig. 4C). Active unfolding of RuBisCo was suggested to result from GroES-induced movements of the apical GroEL domains, exerting a stretching force on the bound polypeptide (*69*), but this effect has not yet been confirmed with any other GroEL-dependent protein (*69*, *70*).

As shown recently, GroEL also interacts

with, and assists in the folding of, certain proteins too large to be encapsulated by GroES (*23*, *24*, *71*). Mitochondrial aconitase (82 kD), for example, can fold through ATP-regulated cycles of GroEL binding and release of nonnative states, with protein release being triggered by the binding of GroES to the opposite (trans) ring of GroEL (*71*).

Group II chaperonins. The Group II chaperonin of the eukaryotic cytosol, TRiC (TCP-1 ring complex, also called CCT for chaperonin-containing TCP-1), contains eight orthologous subunits per ring that differ primarily in their apical domains. The simpler archaeal chaperonin, referred to as thermosome, consists of up to three different subunits, which are arranged in eight- or nine-membered rings. The backbone trace of the chaperonin II apical domain is virtually identical to that of GroEL, with the exception of an α-helical insertion that protrudes from the ring opening (*72*, *73*) and, in the absence of a separate GroES-like cofactor, is thought to function as a built-in lid of the central cavity.

The mechanism of Group II chaperonins is not yet well understood, and the nature and exact location of the substrate binding site(s) on the apical domains are still undefined. The most abundant substrates of TRiC are the cytoskeletal proteins actin and tubulin. Strikingly, folding of these proteins cannot be mediated by GroEL and GroES, suggesting a more specific role for TRiC in folding beyond prevention of aggregation. Actin binds to TRiC through at least two distinct regions and interacts with specific TRiC subunits (*74*, *75*). ATP binding induces encapsulation of the protein by the apical-domain protrusions and initiates folding (*75*, *76*). Through its built-in lid mechanism, TRiC may act cotranslationally in the folding of discrete domains of proteins that are too large to be encapsulated as a whole (*16*) (Fig. 2C).

The subunit heterogeneity of TRiC suggested that the cytosolic chaperonin may be adapted to assisting the folding of a small set of specific proteins, including actins and tubulins. However, as determined by a recent pulse-chase analysis in mammalian cells, TRiC interacts transiently with a wider range of newly synthesized proteins of 30 to 120 kD in size, constituting ~12% of total synthesized chains (*22*). The list of model substrates includes, among others, firefly luciferase, α-transducin, and the von Hippel–Lindau tumor suppressor protein (*5*).

Coordination of Translation and Chaperone Activities

To ensure an efficient use of the cytosolic folding machinery, protein synthesis on ribosomes must be coordinated with the activities of the various chaperone systems in stabilizing nascent chains and in promoting folding. The mechanistic principles underlying this functional cooperation are not yet well understood, but plausible models have been developed based on a combination of in vitro and in vivo studies.

Cotranslational domain folding. It has been suggested that translation itself can have a "chaperone-like" role in the folding of larger proteins composed of multiple domains (*77*). Attempts to refold such proteins in vitro often result in intramolecular misfolding and aggregation (*5*). Cotranslational and sequential domain folding, i.e., the folding of one domain well before another is synthesized, avoids this problem, as demonstrated with artificial two-domain fusion proteins combining the ~20-kD proteins H-ras and dihydrofolate reductase (DHFR) (*77*).

Domain folding during translation occurs in the prokaryotic and eukaryotic cytosol (*78*), but a difference in the efficiencies of folding was noted for certain multidomain proteins when comparing both systems (*77*). The bacterial two-domain protein OmpR and the ras-DHFR fusion proteins fold cotranslationally in a mammalian cell lysate or in intact cells but posttranslationally upon synthesis in the *E. coli* system. Although the individual domains fold efficiently in *E. coli*, bacterial expression of ras-DHFR does not result in an active protein. Eukaryotic cells contain a much greater number of proteins with multiple domains than bacteria (*77*). Although the sizes of protein domains are uniformly distributed in all domains of life, in *E. coli* only 13% of all 4300 proteins exceed a length of 500 residues (~55 kD), compared with 38% of the 5800 proteins in *S. cerevisiae*. Thus, the eukaryotic translation and folding machineries may have been optimized in evolution to facilitate cotranslational domain folding. This optimization may be reflected in the 5- to 10-times slower speed of translation in eukaryotes compared with bacteria and in a functional adaptation of the eukaryotic chaperone machinery. Eukaryotic TRiC, for example, may mediate cotranslational domain folding for some proteins (*16*, *56*, *79*), whereas folding in the bacterial GroEL-GroES cage is strictly posttranslational (*23*) (Fig. 2).

Processivity of chaperone action. The notion of cooperation between mechanistically distinct chaperones in protein folding is now firmly established (*15*, *16*, *20*, *21*, *26*), but how the different components of the folding machinery are functionally integrated is not yet well understood. In principle, substrate transfer between chaperones could be accomplished by free partitioning of nonnative states through the solution. However, considering the highly crowded nature of the cytosol, it is difficult to envisage how aggregation is avoided in this model. Alternatively, ordered pathways of cellular folding may exist in which different chaperones function in a processive manner to minimize the exposure of nonnative proteins to the bulk cytosol.

Whereas in both prokaryotes and eukaryotes specific chaperones are recruited to nascent chains by virtue of their affinity for the ribosome (Fig. 2), the existence of processive chaperone pathways has so far been demonstrated only in the eukaryotic system. As shown in yeast and mammalian cells, folding intermediates generated during biosynthesis are not freely exposed to the bulk cytosol, but rather are functionally compartmentalized (*17*, *22*). In these experiments, certain nascent and newly synthesized protein chains do not bind to a heterologously expressed, noncycling mutant of GroEL (Trap-GroEL), but instead interact productively with the endogenous eukaryotic chaperones. In the specific case of actin, an obligatory substrate of TRiC, protection from exposure to the bulk cytosol during folding is mediated by PFD. The speed and efficiency of actin folding is markedly reduced in PFD-deficient yeast, with nonnative chains being released into the cytosol (*17*). PFD may deliver substrate proteins to TRiC by binding both to nascent chains (*55*, *56*) and to TRiC itself (*51*). In addition, PFD and TRiC seem to cooperate functionally in actin folding, such that nonnative chains are not released into the cytosol during folding cycles (*17*) (Fig. 2C).

Another example of chaperone coupling in the eukaryotic cytosol is the cooperation between Hsc70 and Hsp90 in the folding of signal-transduction proteins (*26*). Substrate transfer from Hsc70 to Hsp90 is mediated by Hop (Hsp organizing protein; also known as p60), an adaptor protein that physically links both chaperones. Hop contains two tetratricopeptide repeat (TPR) domains, which bind the extended COOH-terminal sequences of Hsc70 and Hsp90, respectively (*80*) (Fig. 3A). As shown recently, similar mechanisms are involved in regulating the transfer of nonnative or irreversibly misfolded proteins from these chaperones to the ubiquitin-proteasome machinery. The protein CHIP associates with Hsp90 through an NH_2-terminal TPR domain and targets certain Hsp90 substrates for degradation through a COOH-terminal ubiquitin ligase domain (*81*, *82*). CHIP cooperates functionally with Bag-1 (see above), which binds to Hsc70 and to the proteasome (*83*). These findings provide the first insight into the mechanisms that integrate chaperone-assisted folding and proteolytic degradation, the two main components of protein quality control in the cytosol.

Perspectives

Recent years have seen major advances in our understanding of the basic mechanisms of chaperone-assisted protein folding. Future efforts will define more comprehensively the rules for how the thousands of different pro-

teins in a cell use the chaperone machinery. Global analyses of chaperone usage and folding properties in eukaryotes and prokaryotes will address these questions through a combination of proteomics and high-throughput protein expression (*84*). These studies may eventually offer a rational basis to optimize recombinant protein production in organisms that have been genetically modified to provide the appropriate folding machineries.

In addition to its biotechnological interest, understanding the complex functions of the chaperone arsenal will likely prove useful in dissecting the mechanisms by which protein misfolding and aggregation cause disease. Are chaperones capable of preventing the deposition of amyloid aggregates, and if so, why do these defense mechanisms fail in the millions of patients suffering from neurodegenerative maladies such as Alzheimer's or Huntington's disease? Recent reports that an up-regulation of the Hsp70 system can suppress the neurotoxicity of certain amyloidogenic proteins (*85*, *86*) point toward molecular chaperones as promising targets in the quest for treatment of protein-misfolding diseases.

References and Notes

1. C. Dobson, M. Karplus, *Curr. Opin. Struct. Biol.* **9**, 92 (1999).
2. R. J. Ellis, S. M. Hemmingsen, *Trends Biochem. Sci.* **14**, 339 (1989).
3. M.-J. Gething, J. Sambrook, *Nature* **355**, 33 (1992).
4. F. U. Hartl, *Nature* **381**, 571 (1996).
5. J. Frydman, *Annu. Rev. Biochem.* **70**, 603 (2001).
6. S. E. Radford, *Trends Biochem. Sci.* **25**, 611 (2000).
7. C. M. Dobson, *Trends Biochem. Sci.* **24**, 329 (1999).
8. P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, *Science* **289**, 920 (2000).
9. R. J. Ellis, *Curr. Opin. Struct. Biol.* **11**, 114 (2001).
10. B. van den Berg, R. J. Ellis, C. M. Dobson, *EMBO J.* **18**, 6927 (1999).
11. A. P. Minton, *Curr. Opin. Struct. Biol.* **10**, 34 (2000).
12. J. Martin, F. U. Hartl, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1107 (1997).
13. A. P. Ben-Zvi, P. Goloubinoff, *J. Struct. Biol.* **135**, 84 (2001).
14. C. Schiene, G. Fischer, *Curr. Opin. Struct. Biol.* **10**, 40 (2000).
15. T. Langer *et al.*, *Nature* **356**, 683 (1992).
16. J. Frydman, E. Nimmesgern, K. Ohtsuka, F. U. Hartl, *Nature* **370**, 111 (1994).
17. K. Siegers *et al.*, *EMBO J.* **18**, 75 (1999).
18. O. Fayet, T. Ziegelhoffer, C. Georgopoulos, *J. Bacteriol.* **171**, 1379 (1989).
19. A. L. Horwich, K. B. Low, W. A. Fenton, I. N. Hirshfield, K. Furtak, *Cell* **74**, 909 (1993).
20. E. Deuerling, A. Schulze-Specking, T. Tomoyasu, A. Mogk, B. Bukau, *Nature* **400**, 693 (1999).
21. S. A. Teter *et al.*, *Cell* **97**, 755 (1999).
22. V. Thulasiraman, C. F. Yang, J. Frydman, *EMBO J.* **18**, 85 (1999).
23. K. L. Ewalt, J. P. Hendrick, W. A. Houry, F. U. Hartl, *Cell* **90**, 491 (1997).
24. W. A. Houry, D. Frishman, C. Eckerskorn, F. Lottspeich, F. U. Hartl, *Nature* **402**, 147 (1999).
25. J. Buchner, *Trends Biochem. Sci.* **24**, 136 (1999).
26. J. C. Young, I. Moarefi, F. U. Hartl, *J. Cell Biol.* **154**, 267 (2001).
27. T. Hesterkamp, S. Hauser, H. Lutcke, B. Bukau, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4437 (1996).
28. H. Patzelt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14244 (2001).
29. C. Schiene-Fischer, J. Habazettl, G. Fischer, personal communication.
30. B. Bukau, E. Deuerling, C. Pfund, E. A. Craig, *Cell* **101**, 119 (2000).
31. B. Wiedmann, H. Sakai, T. A. Davis, M. Wiedmann, *Nature* **370**, 434 (1994).
32. M. R. Leroux, *Adv. Appl. Microbiol.* **50**, 219 (2001).
33. B. Beatrix, H. Sakai, M. Wiedmann, *J. Biol. Chem.* **275**, 37838 (2000).
34. C. Pfund *et al.*, *EMBO J.* **17**, 3981 (1998).
35. M. Gautschi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3762 (2001).
36. M. Gautschi, A. Mun, S. Ross, S. Rospert, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
37. B. Bukau, A. L. Horwich, *Cell* **92**, 351 (1998).
38. X. T. Zhu *et al.*, *Science* **272**, 1606 (1996).
39. S. Rüdiger, L. Germeroth, J. Schneider-Mergener, B. Bukau, *EMBO J.* **16**, 1501 (1997).
40. C. J. Harrison, M. Hayer-Hartl, M. Di Liberto, F. U. Hartl, J. Kuriyan, *Science* **276**, 431 (1997).
41. M. P. Mayer *et al.*, *Nature Struct. Biol.* **7**, 586 (2000).
42. M. Pellecchia *et al.*, *Nature Struct. Biol.* **7**, 298 (2000).
43. S. Rüdiger, J. Schneider-Mergener, B. Bukau, *EMBO J.* **20**, 1042 (2001).
44. B. D. Sha, S. Lee, D. M. Cyr, *Struct. Fold. Des.* **8**, 799 (2000).
45. B. Misselwitz, O. Staeck, T. A. Rapoport, *Mol. Cell* **2**, 593 (1998).
46. J. Höhfeld, S. Jentsch, *EMBO J.* **16**, 6209 (1997).
47. H. Sondermann *et al.*, *Science* **291**, 1553 (2001).
48. A. Szabo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10345 (1994).
49. R. Herbst, K. Gast, R. Seckler, *Biochemistry* **37**, 6586 (1998).
50. S. Kim, B. Schilke, E. A. Craig, A. L. Horwich, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12860 (1998).
51. I. E. Vainberg *et al.*, *Cell* **93**, 863 (1998).
52. S. Geissler, K. Siegers, E. Schiebel, *EMBO J.* **17**, 952 (1998).
53. M. R. Leroux *et al.*, *EMBO J.* **18**, 6730 (1999).
54. R. Siegert, M. R. Leroux, C. Scheufler, F. U. Hartl, I. Moarefi, *Cell* **103**, 621 (2000).
55. W. J. Hansen, N. J. Cowan, W. J. Welch, *J. Cell Biol.* **145**, 265 (1999).
56. K. Siegers, F. U. Hartl., personal communication.
57. P. B. Sigler *et al.*, *Annu. Rev. Biochem.* **67**, 581 (1998).
58. J. Chatellier, A. M. Buckle, A. R. Fersht, *J. Mol. Biol.* **292**, 163 (1999).
59. L. L. Chen, P. B. Sigler, *Cell* **99**, 757 (1999).
60. G. W. Farr *et al.*, *Cell* **100**, 561 (2000).
61. Z. H. Xu, A. L. Horwich, P. B. Sigler, *Nature* **388**, 741 (1997).
62. A. Richardson, S. J. Landry, C. Georgopoulos, *Trends Biochem. Sci.* **23**, 138 (1998).
63. A. M. Roseman, S. X. Chen, H. White, K. Braig, H. R. Saibil, *Cell* **87**, 241 (1996).
64. R. J. Ellis, *Curr. Biol.* **11**, R1038 (2001).
65. M. Mayhew *et al.*, *Nature* **379**, 420 (1996).
66. J. S. Weissman, H. S. Rye, W. A. Fenton, J. M. Beechem, A. L. Horwich, *Cell* **84**, 481 (1996).
67. A. Brinker *et al.*, *Cell* **107**, 223 (2001).
68. J. I. Glass *et al.*, *Nature* **407**, 757 (2000).
69. M. Shtilerman, G. H. Lorimer, S. W. Englander, *Science.* **284**, 822 (1999).
70. J. W. Chen, S. Walter, A. L. Horwich, D. L. Smith, *Nature Struct. Biol.* **8**, 721 (2001).
71. T. K. Chaudhuri, G. W. Farr, W. A. Fenton, S. Rospert, A. L. Horwich, *Cell* **107**, 235 (2001).
72. M. Klumpp, W. Baumeister, L. O. Essen, *Cell* **91**, 263 (1997).
73. L. Ditzel *et al.*, *Cell* **93**, 125 (1998).
74. O. Llorca *et al.*, *EMBO J.* **19**, 5971 (2000).
75. O. Llorca *et al.*, *EMBO J.* **20**, 4065 (2001).
76. I. Gutsche, J. Holzinger, N. Rauh, W. Baumeister, R. P. May, *J. Struct. Biol.* **135**, 139 (2001).
77. W. J. Netzer, F. U. Hartl, *Nature* **388**, 343 (1997).
78. A. V. Nicola, W. Chen, A. Helenius, *Nature Cell Biol.* **1**, 341 (1999).
79. C. D. McCallum, H. Do, A. E. Johnson, J. Frydman, *J. Cell Biol.* **149**, 591 (2000).
80. C. Scheufler *et al.*, *Cell* **101**, 199 (2000).
81. J. Demand, S. Alberti, C. Patterson, J. Höhfeld, *Curr. Biol.* **11**, 1569 (2001).
82. P. Connell *et al.*, *Nature Cell Biol.* **3**, 93 (2001).
83. J. Luders, J. Demand, J. Höhfeld, *J. Biol. Chem.* **275**, 4613 (2000).
84. S. A. Lesley, *Protein Expr. Purif.* **22**, 159 (2001).
85. P. Kazemi-Esfarjani, S. Benzer, *Science* **287**, 1837 (2000).
86. C. J. Cummings *et al.*, *Hum. Mol. Genet.* **10**, 1511 (2001).
87. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).

88. We thank our colleagues B. Bukau, E. Deuerling, R. J. Ellis, G. Fischer, A. Horwich, S. Rospert, and K. Siegers for providing access to unpublished information, and R. J. Ellis and members of the Hartl group for critically reading the manuscript.

CMLS, Cell. Mol. Life Sci. 59 (2002) 1589–1597
1420-682X/02/101589-09

CMLS Cellular and Molecular Life Sciences

Structure and function of the GroE chaperone

S. Walter

Institut für Organische Chemie and Biochemie, Technische Universität München, Lichtenbergstr. 4, 85747 Garching (Germany), Fax +49 89 289 13345, e-mail: stefan.walter@ch.tum.de

Abstract. The *Escherichia coli* proteins GroEL and GroES were the first chaperones to be studied in detail and have thus become a role model for assisted protein folding in general. A wealth of both structural and functional data on the GroE system has been accumulated over the past years, enabling us now to understand the basic principles of how this fascinating protein-folding machine accomplishes its task. According to the current model, GroE processes a nonnative polypeptide in a cycle consisting of three steps. First, the polypeptide substrate is captured by GroEL. Upon binding of the cochaperone GroES and ATP, the substrate is then discharged into a unique microenvironment inside of the chaperone, which promotes productive folding. After hydrolysis of ATP, the polypeptide is released into solution. Moreover, GroE may actively increase the folding efficiency, e.g. by unfolding of misfolded protein molecules. The mechanisms underlying these features, however, are yet not well characterized.

Key words. Molecular chaperones; protein folding; GroEL; aggregation; protein structure.

Introduction

Historically, the GroE proteins of *Escherichia coli* were the first chaperone proteins to be studied on a molecular level [1, 2]. In the early 1970s, temperature-sensitive, mutant *E. coli* strains were isolated that were unable to support the growth of bacteriophage λ [1]. Further analyses revealed that apparently two host proteins, GroEL (57 kDa) and GroES (10 kDa), were required for the correct assembly of the phage capsids. Both proteins were found to be essential for the growth of *E. coli* [3]. At that time, the cellular function of GroEL was unknown. The turning point came in the late 1980s, when George Lorimer and co-workers began to investigate whether the GroE proteins could assist in the biogenesis of Rubisco expressed in *E. coli* [4]. They observed that in wild-type cells the formation of active Rubisco was severely compromised. Upon overexpression of both GroEL and GroES, however, active Rubisco was produced. It was suggested that GroE's primary function was to prevent the aggregation of Rubisco during its folding [5], as could be demonstrated later [6, 7].

In the past decade, a wealth of both biochemical and structural data on the GroE chaperone has been accumulated, making it the most thoroughly investigated chaperone system so far [8–12]. According to the model emerging from this data, GroE-assisted protein folding is a three-step process. An aggregation-prone folding intermediate is first captured by GroEL and thereby becomes protected from aggregation. Upon binding of ATP and GroES to the GroEL/polypeptide complex, the polypeptide is ejected into a closed compartment formed by the GroE chaperone, where folding is initiated. After hydrolysis of ATP, both GroES and the polypeptide are released. While this basic mechanism of GroE action is now widely accepted, there are still a number of details that remain controversial and require further experimental investigation.

Architecture of the GroEL protein

The mechanism of GroE-mediated protein folding is intimately related with the oligomeric structure of the chaperone [13]. The GroEL molecule is a complex assembly comprising 14 identical 57-kDa subunits. The transitions between the different functional states of the chaperone are triggered by a set of domain movements which in turn are controlled by the binding of ATP and the cochaperone GroES.

The first images of the GroE chaperone were obtained by electron microscopy [14, 15]. They showed cylindrical particles containing a central channel, which could be occupied by a polypeptide substrate [16–18]. A more detailed picture became available with the X-ray structure of GroEL [19]. It confirmed that the GroEL molecule resembles a barrel with dimensions of 137 Å (diameter) and 146 Å (height). Its 14 subunits are arranged in two rings stacked back to back (fig. 1A). The two rings enclose two separate cavities (45 Å wide) that serve as folding compartments for polypeptide substrates (fig. 1B).

Each GroEL subunit can be dissected into three distinct domains (fig. 1A, C) [19]. The equatorial domains (residues 6–133 plus 409–523) constitute the central part of the cylinder and consist mainly of α helices. They serve as the foundations of the GroEL oligomer, since they mediate all interring contacts, and most of the intraring contacts. They also contain the binding pockets for ATP (fig. 1C), which are facing toward the inside of the central cavity.

In contrast to the equatorial domain, the apical domain (residues 191–376) is considerably less ordered. It is located at the opening of the GroEL cylinder (fig. 1) and contains the binding site for both GroES and the polypeptide substrate. Polypeptide binding occurs in a hydrophobic groove, which is formed by two helices facing the central channel (fig. 1C) [20, 21]. This is in agreement with results of an earlier analysis employing site-directed mutagenesis [22]. The bound substrate is stabilized by mainly hydrophobic interactions, but hydrogen bonds between its peptide backbone and the surrounding polar surface of the apical domain may contribute as well. Importantly, this hydrophobic groove is also responsible for the binding of GroES [23].

The intermediate domain (residues 134–190 plus 377–408) serves as a molecular hinge connecting the apical domain with the equatorial domain. Accordingly, its main function is the transmission of allosteric signals between both domains, thus establishing a tight coupling between nucleotide binding and GroES/polypeptide binding.

The interaction between GroES and GroEL

The cochaperone, GroES (fig. 2A), is a dome-shaped heptamer with diameter of 75 Å and a height of 30 Å [24]. It consists almost exclusively of β sheets. Residues 16–33 form the so-called mobile loops, flexible extensions that dangle from the GroES molecule like the tentacles of a jellyfish [25, 26]. Binding of GroES occurs at the apical domains of the GroEL tetradecamer (fig. 2B) and requires that the nucleotide binding sites of the respective GroEL ring are occupied with either ATP or ADP [15, 27, 28]. Upon association, the mobile loops of GroES bind to the hydrophobic peptide binding groove of GroEL and become immobilized [23, 25]. Because of the common seven-fold symmetry of both proteins, binding is thought to be highly cooperative.

Upon binding of its cochaperone, the GroEL molecule undergoes major structural rearrangements that are cen-

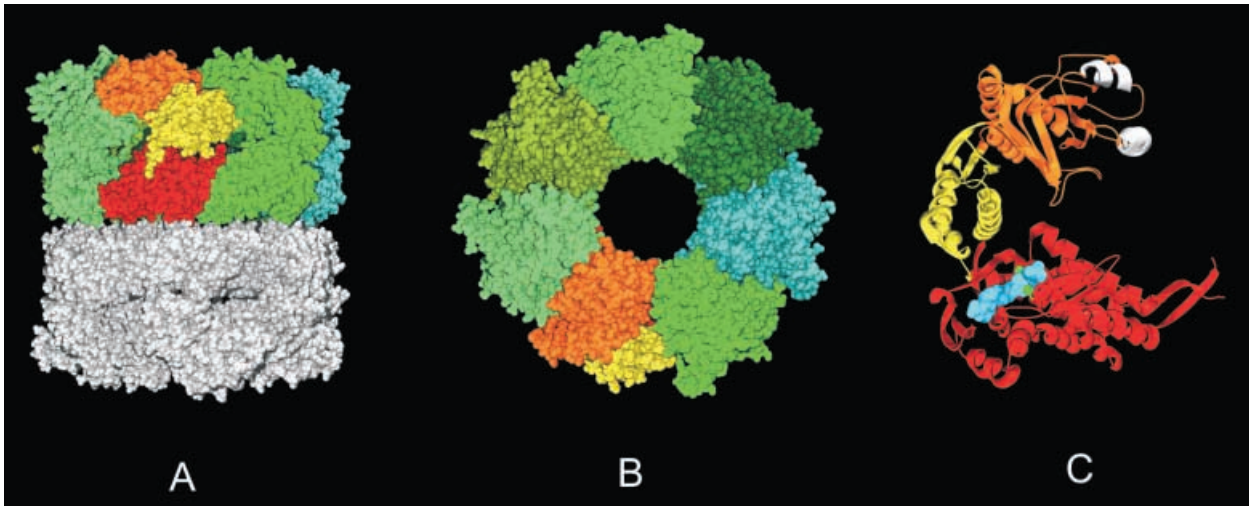


Figure 1. Structure of the GroEL chaperone from *E. coli* [19, 85]. (*A*) Side view of the GroEL tetradecamer. The particle is 137 Å wide and 146 Å high. Subunits comprising the top ring are shown in color, subunits of the bottom ring are shown in gray. Each subunit can be dissected into three domains: apical (orange), intermediate (yellow) and equatorial (red). (*B*) Top view of the GroEL tetradecamer. The diameter of the central cavity is 45 Å. The seven subunits of the ring are shown in shades of green. For one subunit, the apical and the intermediate domains are highlighted in orange and yellow, respectively. (*C*) Ribbon representation of a GroEL subunit. The equatorial domain (red) consists almost exclusively of α helices and contains the nucleotide binding site, which is occupied by ATPγS (blue). The intermediate domain (yellow) serves as a molecular hinge that connects the equatorial domain with the apical domain (orange). Binding of GroES and polypeptides occurs in a hydrophobic groove formed by the two helices (white) facing the central cavity.

tral to its functional cycle (fig. 2B, C) [23, 29]. First, the apical domains of the *cis* ring, i.e. the ring to which GroES binds, swing upward by ~60° and rotate outward by ~90°. As a result, the diameter of the central cavity almost doubles, and its volume increases from 85,000 Å^3 to 175,000 Å^3. Second, the hydrophobic residues, which form the peptide binding site of GroEL, are moved away from the cavity surface and become buried within the wall (fig. 2C). Thus, the surface of the *cis* cavity becomes largely hydrophilic. Third, GroES now blocks the exit of the cavity. As a result, the *cis* cavity is converted from an acceptor site for hydrophobic polypeptides into a closed microenvironment for protein folding.

Depending on the experimental conditions, two types of complexes between GroES and GroEL have been detected by electron microscopy. In the presence of ADP or micromolar concentrations of ATP, GroES binds to only one end of the GroEL cylinder forming asymmetric 'bullets' (see fig. 2B) [15, 16]. At ATP concentrations in the millimolar range, symmetrical 'footballs' have been observed in which both ends of the GroEL particle are capped with GroES [30–32]. It is assumed that the 'ADP bullet', i.e. the $GroES_7 \cdot ADP_7 \cdot GroEL_7/GroEL_7$ complex, represents the 'acceptor state' of GroE [33], which captures an unfolded polypeptide. The footballs presumably reflect a transient species that is formed during the functional cycle (section 6).

Polypeptide binding by GroEL

All molecular chaperones interact with unfolded or partially folded polypeptides. In the case of GroEL, a study using denatured proteins from cell extracts showed that ~40% of the *E. coli* proteins can bind to GroEL [34]. However, it is unlikely that GroEL participates in the folding of all these proteins, because its cellular concentration of ~ 1 μM is simply too small for that purpose [35]. A number of *E. coli* proteins that interact with GroEL in vivo have been identified [36], but it is not clear yet how many of them are stringently dependent on GroE in their folding.

Polypeptide binding to GroEL is primarily based on the hydrophobic effect, as was shown by a thermodynamic analysis of the binding reaction [37], although electrostatic interactions may play a role as well [38, 39]. Unlike correctly folded proteins, unfolded or partially folded polypeptides usually expose hydrophobic surfaces that can associate nonspecifically, forming higher-order aggregates [40, 41]. By binding to the hydrophobic groove in the apical do-

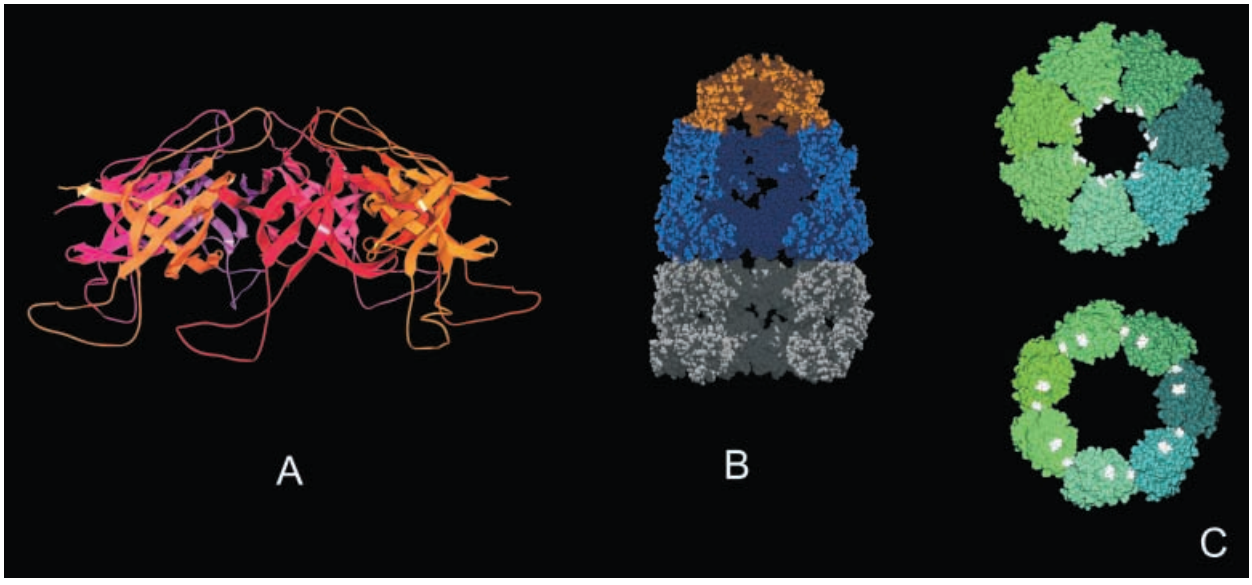


Figure 2. Structure of the GroE chaperone from *E. coli* [23, 24]. (*A*) Side view of the GroES heptamer. The individual subunits (in shades of red) consist mainly of β sheets and form a dome with a diameter of 75 Å. The flexible extensions on the bottom are the so-called mobile loops that mediate binding to GroEL. (*B*) Cross-section of a GroE 'bullet'. Each GroEL ring encloses a cavity that serves as a folding compartment for a polypeptide substrate. Some residues of the equatorial domains have not been resolved in the crystal structure, giving the wrong impression that the two cavities are contiguous. Binding of GroES (orange) to the top GroEL ring (blue) blocks the access to the upper cavity and concomitantly induces an en bloc movement of the apical domains. (*C*) Changes in the GroEL structure upon binding of GroES. In this top view, the seven subunits comprising one ring of GroEL are shown in shades of green and blue. The equatorial domains have been omitted for sake of clarity. The hydrophobic residues in the apical domains important for binding of polypeptide and GroES are shown in white. In the absence of GroES (top panel), these residues coat the inside of the central cavity and account for the high affinity for unfolded polypeptides of this state. Upon binding of GroES (lower panel) the apical domains rotate outwards by ~ 90°. The hydrophobic patches become buried in the subunit interfaces, rendering the inner surface of the cavity mainly hydrophilic and causing the release of a bound polypeptide. Concomitantly, the diameter of the cavity increases from 45 to 80 Å.

mains of GroEL (fig. 1C) these regions become shielded, and the molecule is protected from aggregation. Since most native proteins do not expose hydrophobic surfaces, GroEL will not recognize them as substrates.

The structure of various substrate proteins has been characterized while bound to GroEL. It appears that GroEL is capable of interacting with different conformations ranging from largely unfolded polypeptides to highly structured, stable folding intermediates [42–45]. It has long been a matter of debate whether GroEL recognizes certain structure motifs in its polypeptide substrates [36, 46], but most data suggest that this is not the case [47, 48]. Rather, it seems to be important whether structure formation in the polypeptide and binding to GroEL are synergistic or antagonistic processes. This is illustrated by a study using short peptides with identical amino acid composition but different sequence [49]. Only when the spacing of the hydrophobic residues was such that an amphipathic α helix could form, binding to GroEL was found to be tight. In this case, the formation of structure (= helix) and the formation of a high-affinity binding interface (= hydrophobic surface) were synergistic. Similar results were obtained for peptides mimicking β strands. It is likely that both the partially folded polypeptide and the binding site on GroEL undergo structural rearrangements upon association in order to optimize the binding interface [21, 50].

Allosteric interactions within the GroE chaperone

Though each GroEL ring consists of seven subunits, it represents a single operational unit [10, 51]. This behavior is the consequence of a framework of allosteric interactions that coordinates the binding properties of the individual subunits. There are two levels of cooperativity within the GroEL molecule. First, subunits of the same ring are subject to positive cooperativity. As an example, binding of ATP to one GroEL subunit promotes the binding of ATP to the other six subunits of the same ring [28, 52, 53]. Second, there is a negative cooperativity between the rings, i.e. binding of ATP to one ring reduces the affinity for ATP of the second ring [54, 55]. These homotropic effects can be described by a model of nested cooperativity (fig. 3) [56].

Each GroEL subunit can adopt one of two states: the (relaxed) R state, and the (tense) T state, which differ in their affinity for nucleotide and protein ligands [57]. The R state is characterized by a high affinity for ATP and a low affinity for polypeptides, whereas the T state has a low affinity for ATP and a high affinity for polypeptides. Owing to the positive intraring cooperativity, each ring is either in the R form or in the T form. Thus, the GroEL tetradecamer can adopt the configurations TT, TR, and RR (fig. 3). In the absence of nucleotides, GroEL is preferentially in the TT state. Low concentrations of ATP shift the equilibrium to the RT state, in which the R ring is completely occupied with nucleotide, whereas the T ring is empty. Because of the negative interring cooperativity, the transition to the RR state only occurs at higher ATP concentrations (>100 μM). The cochaperone GroES, on the other hand, seems to reduce the negative interring cooperativity, since its binding to the RT state promotes the transition to the RR state [57]. This is consistent with the finding that in the 'ATP bullet' complex ($GroES_7 \cdot ATP_7 \cdot GroEL_7/GroEL_7$) the trans ring shows a decreased affinity for polypeptides [58].

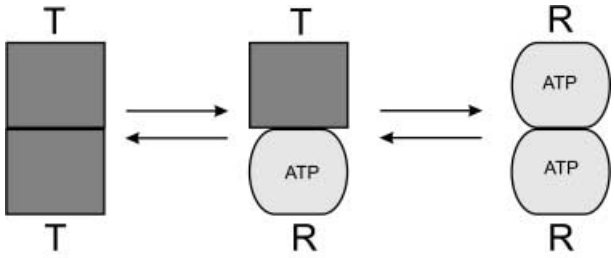


Figure 3. Model of nested cooperativity in GroEL [57]. Each rectangle (square or round) represents a single ring. Owing to the positive intraring cooperativity, all seven subunits within one ring adopt the same state. In the absence of ligands, GroEL is predominantly in the TT state (left). In the presence of low concentrations of ATP, the equilibrium is shifted towards the TR state (middle), because ATP preferentially binds to the R conformation. At higher concentrations of ATP, transition to the RR state occurs.

The functional cycle of the GroE chaperone

GroE-mediated folding requires the polypeptide substrate to participate in a cycle which can be dissected into three steps: capture, sequestration/folding and release [59]. Depending on the nature of the polypeptide, multiple rounds may be necessary for successful folding [60, 61]. The cycling ends when the polypeptide molecule has reached a conformation that is no longer recognized by GroEL. For some monomeric proteins like rhodanese, this exit point may be the native state [33, 62]. In general, however, it will be a committed state in which the protein has not yet reached its native conformation, but no longer requires the assistance of GroE [45, 63]. The reactions that lead from there to the native state may include further folding processes as well as oligomerization. Other molecules may adopt a conformation from which the native state is kinetically inaccessible. These dead-end products probably become degraded by cellular proteases [9].

The sequence of events during GroE cycling is best explained on the level of a single ring. The cycle starts when a polypeptide is captured by GroE (fig. 4, step 1). As mentioned above, a potential substrate is recognized by virtue of its exposed hydrophobic surfaces. The acceptor

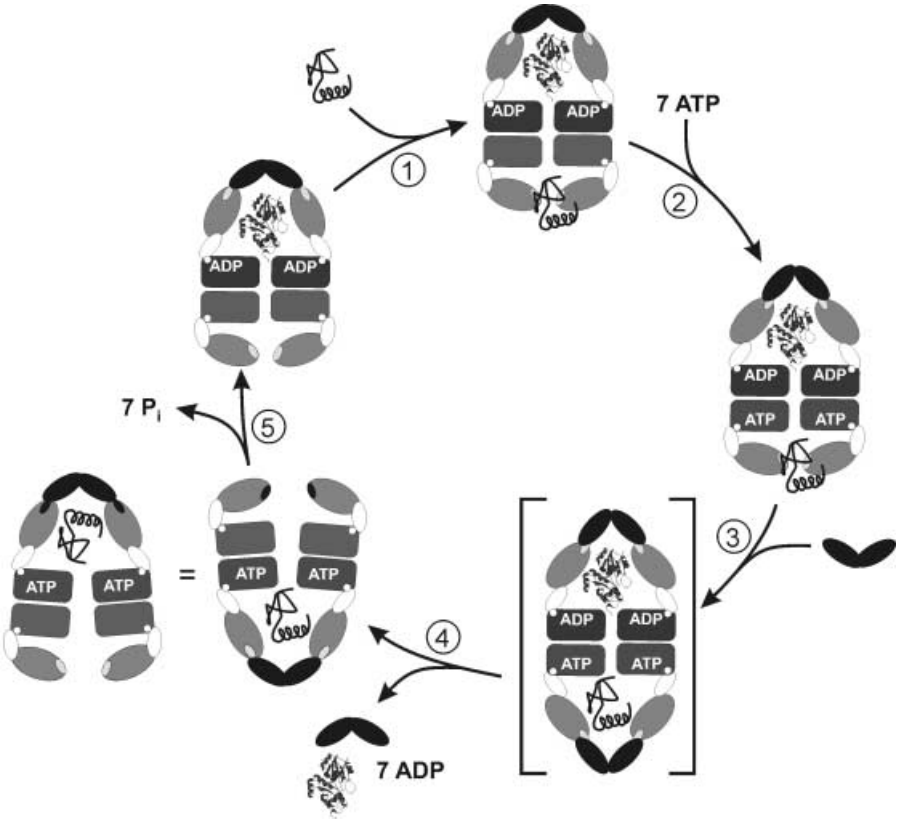


Figure 4. Model of the functional cycle of the GroE chaperone. Although GroEL is composed of two rings, the functional cycle is best described on the level of single rings, which represent the operational units of the chaperone. While both rings are active at the same time, they are in different phases of the cycle. Processing of an individual substrate polypeptide requires two revolutions of the GroE cycle during which the polypeptide remains associated with the same GroEL ring. For graphical reasons, the orientation of the GroE complex is reversed after step 4. The cycle of GroE-assisted folding can be dissected into three steps: capture, encapsulation/folding and release. During capture (1), a hydrophobic polypeptide is prevented from aggregation by binding to GroEL. The acceptor ring (bottom ring) is nucleotide free and therefore has a high affinity for the polypeptide. Binding of ATP (2) and GroES (3) to this ring induces a set of structural changes in GroEL. Most important, the affinity for the bound polypeptide is decreased, and it is released into the closed cavity where folding begins. Subsequent hydrolysis of ATP (5) induces a second conformational change in GroEL (top ring), which allows the bottom ring to bind polypeptide and initiate a new cycle. Upon binding of ATP and GroES in the next round, GroES is displaced from the top ring, and the substrate polypeptide is released (4). The formation of the symmetric complex shown in brackets is controversial.

state of GroE likely is the 'ADP bullet', in which the *trans* ring (i.e. the ring opposite of GroES) is in the high-affinity T form. Subsequent binding of ATP and GroES to this ring (fig. 4, steps 2 and 3) triggers a series of conformational changes in the GroEL molecule. As a result,

1) The affinity for both ADP and GroES in the *trans* ring decreases, causing the dissociation of both ligands (fig. 4, step 4) [64–66].
2) GroES covers the opening of the new *cis* cavity, thereby creating a closed compartment which sequesters the polypeptide substrate [33, 62].
3) The polypeptide binding site in the *cis* ring becomes buried within the cavity wall, causing the discharge of the polypeptide into the *cis* cavity.
4) The size of the cavity increases, giving the polypeptide sufficient room to undergo the structural rearrangements required for productive folding [23].

Once released from GroEL, the polypeptide will start to fold. Since it is still trapped in the central cavity, aggregation is no longer possible. In addition, the sterical restrictions resulting from confinement and the physical properties of the cavity wall may alter the energy landscape of the folding reaction [67, 68].

The formation of the symmetric intermediate shown in step 3 of the cycle is controversial. Though football-shaped GroE particles have been detected using electron microscopy (see section 2), it is not clear yet whether these molecules represent off-pathway products [69], or are indeed part of the chaperone cycle [70, 71]. The 'football' intermediate is an attractive concept since it would allow GroES to trap a 'new' polypeptide in the *cis* cavity, and at the same time trigger the release of the 'old' polypeptide from the *trans* ring.

Hydrolysis of the bound ATP (fig. 4, step 5), which takes ~10 s at 25 °C [55], represents the rate-limiting step in the

cycle and thus serves as a timer for encapsulation [58]. Once the ATP is hydrolyzed, the chaperone has completed its cycle and the next round starts, in which now the opposite ring will be charged with a polypeptide substrate. The release of the encapsulated protein occurs upon the subsequent binding of GroES/ATP (fig. 4, step 4).

At this stage, the ejected polypeptide is thought to undergo kinetic partitioning [64, 72]. Molecules that are not recognized by GroEL (native, committed or dead-end, see above) no longer participate in cycling. The remaining molecules may rebind and undergo another round of the GroE cycle, or bind to other molecular chaperones, or fold/assemble in bulk solution. The relative fractions of these species likely depend on the nature of the polypeptide as well as on the cellular context.

It is reasonable to assume that GroEL can process two substrates at a time, as shown in figure 4. According to this model, both rings (= operational units) are active, although they are in different phases of the chaperone cycle [58, 71]. As an example, the top ring after step 3 is loaded with a polypeptide that already has undergone folding in the cavity, and will become ejected in the next step. The polypeptide bound to the bottom ring, on the other hand, will be released into the cavity where folding is initiated. Thus, binding of GroES in step 3 has a dual function: it sequesters a 'new' polypeptide in the *cis* cavity, and it releases GroES and the processed polypeptide from the *trans* ring. In an alternative model, only one GroEL ring at a time is loaded with a polypeptide, whereas the second ring passes through the cycle in an empty state.

Owing to the limited volume of the cavity, GroE-mediated folding as shown in figure 4 is restricted to polypeptides smaller than ~ 60 kDa [73, 74]. Although larger proteins can bind to GroEL, they cannot become encapsulated underneath GroES. This is illustrated by the case of aconitase from yeast mitochondria, a monomeric enzyme of 82 kDa. Strains in which mt-hsp60, the yeast homologue of GroEL, was deleted were found to accumulate aggregated aconitase. Further analyses showed that both GroEL and GroES are required for the biogenesis of aconitase in vivo [75]. However, aconitase did not become encapsulated since GroES could only bind in *trans*, not in *cis* [76]. More important, GroES binding in *trans* was found to be essential for the release of GroEL-bound aconitase. It may thus reflect a general mechanism GroEL uses to dispose of substrates it cannot process [77].

GroE-mediated unfolding of proteins

A still open question is how GroE promotes the folding of proteins that are trapped in nonnative conformations. It is assumed that GroEL is capable of partially unfolding these proteins, thereby setting them back on the right track to the native state [78]. Several mechanisms have been suggested how GroEL accomplishes this task.

The most simple model, thermodynamic coupling [79], is based on the idea that for some folding intermediates structure formation and binding to GroEL may be antagonistic processes because they compete for the same (hydrophobic) residues. Since the amount of exposed hydrophobic surface generally decreases with the degree of folding, GroEL will preferentially bind to more unfolded conformations of a protein. Provided there is a rapid equilibrium between the various conformations of the polypeptide, GroEL will effectively unfold the protein by the law of mass action. This capability of GroEL has been demonstrated for a variety of relatively small proteins [79–82]. The coupling mechanism, however, has one important shortcoming. It would not allow a polypeptide to escape from a kinetic trap on its folding pathway, because, according to this model, all unfolding reactions occur in free solution at their intrinsic rates.

Based on experiments with Rubisco, it was also suggested that GroE can actively unfold a bound polypeptide [83]. According to this model, the movement of the apical domains, which occurs upon GroES binding (fig. 2C), may exert a mechanical stress on the bound polypeptide, thereby virtually tearing its structure apart. Importantly, this mechanism requires that the substrate be bound to multiple apical domains simultaneously [84]. Active unfolding, however, could yet not be observed with other stringent substrate proteins of GroE [45].

Perspectives

While our understanding of the basic mechanism(s) underlying GroE-assisted folding has increased considerably over the past years, there are still a number of issues that require further investigation. One of the most interesting problems that remains to be solved is the conformational changes a polypeptide undergoes when it is processed by the chaperone. This concerns unfolding reactions that occur upon binding of a polypeptide to GroEL as well as those unfolding events that may be associated with the subsequent binding of GroES (see above). Of equal importance is the question of what happens to the polypeptide while it folds in the cavity. Is the main purpose of sequestration to exclude other polypeptides and thus provide conditions of infinite dilution? Or does the cavity play a more active role, e.g. by lowering energy barriers on the folding pathway? Scientists have already begun to address these issues experimentally. The size of the GroE chaperone and the low conformational stability of the folding intermediates, however, make this a rather challenging undertaking.

Another issue concerns the general applicability of results obtained with a single 'model' substrate. Likely,

GroE provides a variety of 'tools' that promote the efficiency of protein folding, but not all of them may be observed when GroE processes a certain polypeptide substrate. This point is illustrated by the long-lasting debate about what is required to release a polypeptide from GroEL. In some cases, addition of ADP alone was found to be sufficient, in other cases ATP was required, while for a third set of protein substrates release was dependent on both ATP and GroES. These differences may simply reflect the heterogeneity of protein substrates GroE encounters in the cell. Thus, it will be crucial for our future understanding to identify the *E. coli* proteins that depend on GroE in their folding.

Acknowledgements. I thank Holger Grallert for help with preparation of figure 4. Financial support by the DFG is gratefully acknowledged.

1 Georgopoulos C., Hendrix R. W., Casjens S. R. and Kaiser A. D. (1973) Host participation in bacteriophage lambda head assembly. J. Mol. Biol. **76:** 45–60
2 Lorimer G. H. (2001) A personal account of chaperonin history. Plant Physiol. **125:** 38–41
3 Fayet O., Ziegelhoffer T. and Georgopoulos C. (1989) The groES and groEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. J. Bact. **171:** 1379–1385
4 Goloubinoff P., Gatenby A. A. and Lorimer G. H. (1989) GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphoshate carboxylase oligomers in *Escherichia coli*. Nature **337:** 44–47
5 Goloubinoff P., Christeller J. T., Gatenby A. A. and Lorimer G. H. (1989) Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. Nature **342:** 884–889
6 Buchner J., Schmidt M., Fuchs M. Jaenicke R. Rudolph R., Schmid F. X. et al. (1991) GroE facilitates refolding of citrate synthase by suppressing aggregation. Biochemistry **30:** 1586–1591
7 Höll-Neugebauer B., Rudolph R., Schmidt M. and Buchner J. (1991) Reconstitution of a heat shock effect in vitro: influence of GroE on the thermal aggregation of alpha-glucosidase from yeast. Biochemistry **30:** 11609–11614
8 Fenton W. A. and Horwich A. L. (1997) GroEL-mediated protein folding. Protein Sci. **6:** 743–760
9 Sigler P. B., Xu Z., Rye H. S., Burston S. G., Fenton W. A. and Horwich A. L. (1998) Structure and function in GroEL-mediated protein folding. Annu. Rev. Biochem. **67:** 581–608
10 Thirumalai D. and Lorimer G. H. (2001) Chaperonin-mediated protein folding. Ann. Rev. Biophys. Biomol. Struct. **30:** 245–269
11 Grallert H. and Buchner J. (2001) A structural view of the GroE chaperone cycle. J. Struct. Biol. **135:** 95–103
12 Walter S. and Buchner J. (2002) Molecular chaperones – cellular machines for protein folding. Angew. Chem. Int. Ed. Engl. **41:** 1098–1113
13 Weber F., Keppel F., Georgopoulos C., Hayer-Hartl M. K. and Hartl F. U. (1998) The oligomeric structure of GroEL/GroES is required for biologically significant chaperonin function in protein folding. Nat. Struct. Biol. **5:** 977–985
14 Hendrix R. W. (1979) Purification and properties of groE, a host protein involved in bacteriophage assembly. J. Mol. Biol. **129:** 375–392
15 Saibil H. R., Dong Z., Wood S. and auf der Mauer A. (1991) Binding of chaperonins. Nature **353:** 25–26
16 Langer T., Pfeifer G., Martin J., Baumeister W. and Hartl F. U. (1992) Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accomodates the protein substrate within its central cavity. EMBO J. **11:** 4757–4765
17 Braig K., Simon M., Furuya F., Hainfeld J. F. and Horwich A. L. (1993) A polypeptide bound by the chaperonin groEL is localized within a central cavity. Proc. Natl. Acad. Sci. USA **90:** 3978–3982
18 Chen S., Roseman A. M., Hunter A. S., Wood S., Burston S. G., Ranson N. A. et al. (1994) Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. Nature **371:** 261–264
19 Braig K., Otwinowski Z., Hedge R., Boisvert D. C., Joachimiak A., Horwich A. L. et al. (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. Nature **371:** 578–586
20 Buckle A. M., Zahn R. and Fersht A. R. (1997) A structural model for GroEL-polypeptide recognition. Proc. Natl. Acad. Sci. USA **94:** 3571–3575
21 Chen L. and Sigler P. B. (1999) The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. Cell **99:** 757–768
22 Fenton W. A., Kashi Y., Furtak K. and Horwich A. L. (1994) Residues in chaperonin GroEL required for polypeptide binding and release. Nature **371:** 614–619
23 Xu Z., Horwich A. L. and Sigler P. B. (1997) The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature **388:** 741–750
24 Hunt J. F., Weaver A. J., Landry S. J., Gierasch L. M. and Deisenhofer J. (1996) The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. Nature **379:** 37–45
25 Landry S. J., Zeilstra-Ryalls O., Fayet O., Georgopoulos C. and Gierasch L. M. (1993) Characterization of a functionally important mobile domain of GroES. Nature **364:** 255–258
26 Richardson A., Schwager F., Landry S. J. and Georgopoulos C. (2001) The importance of a mobile loop in regulating chaperonin/ co-chaperonin interaction: humans versus *Escherichia coli*. J. Biol. Chem. **276:** 4981–4987
27 Chandrasekhar G. N., Tilly K., Woolford C., Hendrix R. W. and Georgopoulos C. (1986) Purification and properties of the groES morphogenetic protein of *Escherichia coli*. J. Biol. Chem. **261:** 12414–12419
28 Jackson G. S., Staniforth R. A., Halsall D. J., Atkinson T., Holbrook J. J., Clarke A. R. et al. (1993) Binding and hydrolysis of nucleotides in the chaperonin catalytic cycle: Implications for the mechanism of assisted protein folding. Biochemistry **32:** 2554–2563
29 Roseman A. M., Chen S., White H., Braig K. and Saibil H. R. (1996) The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. Cell **87:** 241–251
30 Llorca O., Marco S., Carrascosa J. L. and Valpuesta J. M. (1994) The formation of symmetrical GroEL-GroES complexes in the presence of ATP. FEBS Lett. **345:** 181–186
31 Azem A., Kessel M. and Goloubinoff P. (1994) Characterization of a functional $GroEL_{14}(GroES_7)_2$ chaperonin hetero-oligomer. Science **265:** 653–656
32 Schmidt M., Rutkat K., Rachel R., Pfeifer G., Jaenicke R., Viitanen P. V. et al. (1994) Symmetric complexes of GroE chaperonins as part of the functional cycle. Science **265:** 656–659
33 Weissman J. S., Hohl C. M., Kovalenko O., Kashi Y., Chen S., Braig K. et al. (1995) Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. Cell **83:** 577–587
34 Viitanen P. V., Gatenby A. A. and Lorimer G. H. (1992) Purified chaperonin 60 (groEL) interacts with the nonnative states of a multitude of *Escherichia coli* proteins. Protein Sci. **1:** 363–369
35 Lorimer G. H. (1996) A quantitative assessment of the role of the chaperonin proteins in protein folding in vivo. FASEB J. **10:** 5–9

36 Houry W. A., Frishman D., Eckerskorn C., Lottspeich F. and Hartl F. U. (1999) Identification of in vivo substrates of the chaperonin GroEL. Nature **402:** 147–154
37 Lin Z., Schwarz F. P. and Eisenstein E. (1995) The hydrophobic nature of GroEL-substrate binding. J. Biol. Chem. **270:** 1011–1014
38 Katsumata K., Okazaki A. and Kuwajima K. (1996) Effect of GroEL on the re-folding kinetics of α-lactalbumin. J. Mol. Biol. **264:** 827–838
39 Aoki K., Taguchi H., Shindo Y., Yoshida M., Ogasahara K., Yutani K. et al. (1997) Calorimetric observation of a GroEL-protein binding reaction with little contribution of hydrophobic interaction. J. Biol. Chem. **272:** 32158–32162
40 Jaenicke R. and Seckler R. (1997) Protein misassembly in vitro. Adv. Protein Chem. **50:** 1–59
41 Fink A. L. (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid. Folding Des. **3:** R9–R23
42 Robinson C. V., Gross N., Eyles S. J., Ewbank J. J., Mayhew M., Hartl F. U. et al. (1994) Conformation of GroEL-bound α-lactalbumin probed by mass spectrometry. Nature **372:** 646–651
43 Lilie H. and Buchner J. (1995) Interaction of GroEL with a highly structured folding intermediate:Iterative binding cycles do not involve unfolding. Proc. Natl. Acad. Sci. USA **92:** 8100–8104
44 Goldberg M. S., Zhang J., Sondek S., Matthews C. R., Fox R. O. and Horwich A. L. (1997) Native-like structure of a protein-folding intermediate bound to the chaperonin GroEL. Proc. Natl. Acad. Sci. USA **94:** 1080–1085
45 Chen J., Walter S., Horwich A. L. and Smith D. L. (2001) Folding of malate dehydrogenase inside the GroEL-GroES cavity. Nat. Struct. Biol. **8:** 721–728
46 Landry S. J. and Gierasch L. (1991) The chaperonin GroEL binds a polypeptide in an a-helical conformation. Biochemistry **30:** 7359–7362
47 Schmidt M. and Buchner J. (1992) Interaction of GroE with an all-beta-protein. J. Biol. Chem. **267:** 16829–16833
48 Aoki K., Motojima F., Taguchi H., Yomo T. and Yoshida M. (2000) GroEL binds artificial proteins with random sequences. J. Biol. Chem. **275:** 13755–13758
49 Wang Z., Feng H., Landry S. J., Maxwell J. and Gierasch L. M. (1999) Basis of substrate binding by the chaperonin GroEL. Biochemistry **38:** 12537–12546
50 Falke S., Fisher M. T. and Gogol E. P. (2001) Structural changes in GroEL effected by binding a denatured protein substrate. J. Mol. Biol. **308:** 569–577
51 Weissman J. S., Rye H. S., Fenton W. A., Beechem J. M. and Horwich A. L. (1996) Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. Cell **84:** 481–490
52 Gray T. E. and Fersht A. R. (1991) Cooperativity in ATP hydrolysis by GroEL is increased by GroES. FEBS Lett. **292:** 254–258
53 Bochkareva E. S., Lissin N. M., Flynn G. C., Rothman J. E. and Girshovich A. S. (1992) Positive cooperativity in the functioning of molecular chaperone GroEL. J. Biol. Chem. **267:** 6796–6800
54 Yifrach O. and Horovitz A. (1994) Two lines of allosteric communication in the oligomeric chaperonin GroEL are revealed by the single mutation Arg196 $\rightarrow$ Ala. J. Mol. Biol. **243:** 397–401
55 Burston S. G., Ranson N. A. and Clarke A. R. (1995) The origins and consequences of asymmetry in the chaperonin reaction cycle. J. Mol. Biol. **249:** 138–152
56 Yifrach O. and Horovitz A. (1995) Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. Biochemistry **34:** 5303–5308
57 Horovitz A., Fridmann Y., Kafri G. and Yifrach O. (2001) Review: allostery in chaperonins. J. Struct. Biol. **135:** 104–114
58 Rye H. S., Roseman A. M., Chen S., Furtak K., Fenton W. A., Saibil H. R. et al. (1999) GroEL-GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. Cell **97:** 325–338
59 Ranson N. A., Burston S. G. and Clarke A. R. (1997) Binding, encapsulation and ejection: substrate dynamics during a chaperonin-assisted folding reaction. J. Mol. Biol. **266:** 656–664
60 Weissman J. S., Kashi Y., Fenton W. A. and Horwich A. L. (1994) GroEL-mediated protein folding proceeds by multiple rounds of binding and release of nonnative forms. Cell **78:** 693–702
61 Todd M. J., Lorimer G. H. and Thirumalai D. (1996) Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. Proc. Natl. Acad. Sci. USA **93:** 4030–4035
62 Mayhew M., da Silva A. C., Martin J., Erdjument-Bromage H., Tempst P. and Hartl F. U. (1996) Protein folding in the central cavity of the GroEL-GroES chaperonin complex. Nature **379:** 420–426
63 Grallert H., Rutkat K. and Buchner J. (2000) Limits of protein folding inside GroE complexes. J. Biol. Chem. **275:** 20424–20430
64 Todd M. J., Viitanen P. V. and Lorimer G. H. (1994) Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. Science **265:** 659–666
65 Hayer-Hartl M. K., Martin J. and Hartl F. U. (1995) Asymmetrical interaction of GroEL and GroES in the ATPase cycle of assisted protein folding. Science **269:** 836–841
66 Rye H. S., Burston S. G., Fenton W. A., Beechem J. M., Xu Z., Sigler P. B. et al. (1997) Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL. Nature **388:** 792–798
67 Betancourt M. R. and Thirumalai D. (1999) Exploring the kinetic requirements for enhancement of protein folding rates in the GroEL cavity. J. Mol. Biol. **287:** 627–644
68 Brinker A., Pfeifer G., Kerner M. J., Naylor D. J., Hartl F. U. and Hayer-Hartl M. K. (2001) Dual function of protein confinement in chaperonin-assisted protein folding. Cell **107:** 223–233
69 Hayer-Hartl M. K., Ewalt K. L. and Hartl F. U. (1999) On the role of symmetrical and asymmetrical chaperonin complexes in assisted protein folding. Biol. Chem. **380:** 531–540
70 Azem A., Diamant S., Kessel M., Weiss C. and Goloubinoff P. (1995) The protein-folding activity of chaperonins correlates with the symmetric $GroEL_{14}(GroES_7)_2$ heterooligomer. Proc. Natl. Acad. Sci. USA **92:** 12021X–12025
71 Sparrer H., Rutkat K. and Buchner J. (1997) Catalysis of protein folding by symmetric chaperone complexes. Proc. Natl. Acad. Sci. USA **94:** 1096–1100
72 Burston S. G., Weissman J. S., Farr G. W., Fenton W. A. and Horwich A. L. (1996) Release of both native and non-native proteins from a cis-only GroEL ternary complex. Nature **383:** 96–99
73 Ewalt K. L., Hendrick J. P., Houry W. A. and Hartl F. U. (1997) In vivo observation of polypeptide flux through the bacterial chaperonin system. Cell **90:** 491–500
74 Sakihawa C., Taguchi H., Makino Y. and Yoshida M. (1999) On the maximum size of proteins to stay and fold in the cavity of GroEL underneath GroES. J. Biol. Chem. **274:** 21251–21256
75 Dubaquie Y., Looser R., Fünfschilling U., Jeno P. and Rospert S. (1998) Identification of in vivo substrates of the yeast mitochondrial chaperonins reveals overlapping but non-identical requirement for hsp60 and hsp10. EMBO J. **15:** 5868–5876
76 Chaudhuri T. K., Farr G. W., Fenton W. A., Rospert S. and Horwich A. L. (2001) GroEL/GroES-mediated folding of a protein too large to be encapsulated. Cell **107:** 235–246
77 Sparrer H. and Buchner J. (1997) How GroES regulates binding of nonnative protein to GroEL. J. Biol. Chem. **272:** 14080–14086
78 Ranson N. A., Dunster N. J., Burston S. G. and Clarke A. R. (1995) Chaperonins can catalyse the reversal of early aggre-

gation steps when a protein misfolds. J. Mol. Biol. **250:** 581–586
79 Walter S., Lorimer G. H. and Schmid F. X. (1996) A thermodynamic coupling mechanism for GroEL-mediated unfolding. Proc. Natl. Acad. Sci. USA **93:** 9425–9430
80 Viitanen P. V., Donaldson G. K., Lorimer G. H., Lubben T. H. and Gatenby A. A. (1991) Complex interactions between the chaperonin 60 molecular chaperone and dihydrofolate reductase. Biochemistry **30:** 9716–9723
81 Zahn R., Spitzfaden C., Ottiger M., Wüthrich K. and Plückthun A. (1994) Destabilization of the complete protein secondary structure on binding to the chaperone GroEL. Nature **368:** 261–265
82 Zahn R., Perrett S., Stenberg G. and Fersht A. R. (1996) Catalysis of amide proton exchange by the molecular chaperones GroEL and SecB. Science **271:** 642–645
83 Shtilerman M., Lorimer G. H. and Englander S. W. (1999) Chaperonin function: folding by forced unfolding. Science **284:** 822–825
84 Farr G. W., Furtak K., Rowland M. B., Ranson N. A., Saibil H. R. and Horwich A. L. (2000) Multivalent binding of nonnative substrate proteins by the chaperonin GroEL. Cell **100:** 561–573
85 Boisvert D. C., Wang J., Otwinowski Z., Howich A. L. and Sigler P. B. (1996) The 2.4 Å crystal structure of the bacterial chaperonin GroEL complexed with ATPγS. Nat. Struct. Biol. **3:** 170–177

REVIEW

doi:10.1038/nature10317

Molecular chaperones in protein folding and proteostasis

F. Ulrich Hartl[1], Andreas Bracher[1] & Manajit Hayer-Hartl[1]

Most proteins must fold into defined three-dimensional structures to gain functional activity. But in the cellular environment, newly synthesized proteins are at great risk of aberrant folding and aggregation, potentially forming toxic species. To avoid these dangers, cells invest in a complex network of molecular chaperones, which use ingenious mechanisms to prevent aggregation and promote efficient folding. Because protein molecules are highly dynamic, constant chaperone surveillance is required to ensure protein homeostasis (proteostasis). Recent advances suggest that an age-related decline in proteostasis capacity allows the manifestation of various protein-aggregation diseases, including Alzheimer's disease and Parkinson's disease. Interventions in these and numerous other pathological states may spring from a detailed understanding of the pathways underlying proteome maintenance.

Proteins are the most versatile and structurally complex biological macromolecules. They are involved in almost every biological process. Mammalian cells typically express in excess of 10,000 different protein species, which are synthesized on ribosomes as linear chains of up to several thousand amino acids. To function, these chains must generally fold into their 'native state', an ensemble of a few closely related three-dimensional structures[1,2]. How this is accomplished and how cells ensure the conformational integrity of their proteome in the face of acute and chronic challenges constitute one of the most fundamental and medically relevant problems in biology.

Central to this problem is that proteins must retain conformational flexibility to function, and thus are only marginally thermodynamically stable in their physiological environment. A substantial fraction of all proteins in eukaryotic cells (20–30% of the total in mammalian cells) even seem to be inherently devoid of any ordered three-dimensional structure and adopt folded conformations only after interaction with binding partners[3]. Aberrant behaviour of some of these metastable proteins, such as tau and α-synuclein, can give rise to the formation of fibrillar aggregates that are associated with dementia and Parkinson's disease. Thus, protein quality control and the maintenance of proteome homeostasis (known as proteostasis) are crucial for cellular and organismal health. Proteostasis is achieved by an integrated network of several hundred proteins[4], including, most prominently, molecular chaperones and their regulators, which assist in *de novo* folding or refolding, and the ubiquitin–proteasome system (UPS) and autophagy system, which mediate the timely removal of irreversibly misfolded and aggregated proteins. Deficiencies in proteostasis have been shown to facilitate the manifestation or progression of numerous diseases, such as neurodegeneration and dementia, type 2 diabetes, peripheral amyloidosis, lysosomal storage disease, cystic fibrosis, cancer and cardiovascular disease. A major risk factor for many of these ailments is advanced age. Indeed, studies in model organisms indicate that ageing is linked to a gradual decline in cellular proteostasis capacity[5,6].

Here we discuss recent insights into the mechanisms of chaperone-assisted protein folding and proteome maintenance. We focus on how proteins use the chaperone machinery to navigate successfully the complex folding-energy landscape in the crowded cellular environment. Understanding these reactions will guide future efforts to define the proteostasis network as a target for pharmacological intervention in diseases of aberrant protein folding.

Fundamental role of molecular chaperones

Many small proteins refold after their removal from denaturant *in vitro*, in the absence of other components or an energy source. This signifies that the amino-acid sequence, encoded in the DNA, contains all of the necessary information to specify the three-dimensional structure of a protein[1]. However, research over the past couple of decades has firmly established that in the cellular environment, many proteins require molecular chaperones to fold efficiently and on a biologically relevant timescale[7]. Why is this extra layer of complexity necessary?

Although small proteins may fold at very fast speeds[8] (within microseconds), in dilute buffer solutions, larger, multidomain proteins may take minutes to hours to fold[9], and often even fail to reach their native states *in vitro*. The folding of such proteins becomes considerably more challenging *in vivo*, because the cellular environment is highly crowded, with total cytosolic protein reaching concentrations of 300–400 g l^{-1}. The resultant excluded volume effects, although enhancing the functional interactions between macromolecules, also strongly increase the tendency of non-native and structurally flexible proteins to aggregate[10]. It seems likely, therefore, that the fundamental requirement for molecular chaperones arose very early during the evolution of densely crowded cells, owing to the need to minimize protein aggregation during folding and maintain proteins in soluble, yet conformationally dynamic states. Moreover, as mutations often disrupt the ability of a protein to adopt a stable fold[11], it follows that the chaperone system provides a crucial buffer, allowing the evolution of new protein functions and phenotypic traits[11,12].

Some basics on protein folding and how it can go awry

Because the number of possible conformations a protein chain can adopt is very large, folding reactions are highly complex and heterogeneous, relying on the cooperation of many weak, non-covalent interactions. In the case of soluble proteins, hydrophobic forces are particularly important in driving chain collapse and the burial of non-polar amino-acid residues within the interior of the protein (see ref. 13 for a discussion of membrane protein folding). Considerable progress has been made in recent years in understanding these reactions through biophysical experiments and theoretical analyses[1,2]. In the current model, polypeptide chains are thought to explore funnel-shaped potential energy surfaces as they progress, along several

[1]Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany.

downhill routes, towards the native structure (Fig. 1). Chain collapse and the progressive increase in the number of native interactions rapidly restrict the conformational space that needs to be searched en route to the native state. However, the free-energy surface that must be navigated is often rugged, which means that the molecules must cross substantial kinetic barriers during folding. As a consequence, partially folded states may become transiently populated as kinetically trapped species. Such folding intermediates are the rule for proteins larger than 100 amino acids (~90% of all proteins in a cell), which have a strong tendency to undergo rapid hydrophobic collapse into compact globular conformations[2]. The collapse may lead either to disorganized globules lacking specific contacts and retaining large configurational entropy or to intermediates that may be stabilized by non-native interactions (misfolded states). In the former case, the search for crucial native contacts within the globule will limit folding speed, whereas in the latter, the breakage of non-native contacts may be rate-limiting[1] (Fig. 1). The propensity of proteins to populate globular intermediates with a high degree of flexibility may increase with larger, topologically more complex domain folds that are stabilized by many long-range interactions (such as α/β domain architectures). Such proteins are often highly chaperone dependent[14].

Partially folded or misfolded states are problematic because they tend to aggregate in a concentration-dependent manner (Fig. 1). This is due to the fact that these forms typically expose hydrophobic amino-acid residues and regions of unstructured polypeptide backbone to the solvent — features that become buried in the native state[15]. Like intramolecular folding, aggregation is largely driven by hydrophobic forces and primarily results in amorphous structures (Fig. 1). Alternatively, fibrillar aggregates called amyloid may form, defined by β-strands that run perpendicular to the long fibril axis (cross-β structure). Although many proteins can adopt these highly ordered, thermodynamically stable structures under conditions *in vitro*[16], the formation of these aggregates *in vivo* is strongly restricted by the chaperone machinery, suggesting that they may become more widespread under stress or when protein quality control fails. Importantly, the formation of fibrillar aggregates is often accompanied by the formation of soluble oligomeric states, which are thought to have key roles in diseases of aberrant folding[16] (Fig. 1). The toxicity of these less ordered and rather heterogeneous forms has been suggested to correlate with the exposure of sticky, hydrophobic surfaces and accessible peptide-backbone structure that is not yet integrated into a stable cross-β core[17]. The soluble oligomers must undergo considerable rearrangement to form fibrils, the thermodynamic end state of the aggregation process, and may thus be comparable to the kinetically trapped intermediates in folding (Fig. 1). Notably, some common structural epitopes have been detected on the prefibrillar oligomers of different polypeptides[18], but how these features are linked with toxicity is not yet understood. Such information is urgently needed to develop treatments for the numerous pathological states associated with protein aggregation.

Major chaperone classes

We define a molecular chaperone as any protein that interacts with, stabilizes or helps another protein to acquire its functionally active conformation, without being present in its final structure[7,19]. Several different classes of structurally unrelated chaperones exist in cells, forming cooperative pathways and networks. Members of these protein families are often known as stress proteins or heat-shock proteins (HSPs), as they are upregulated under conditions of stress in which the concentrations of aggregation-prone folding intermediates increase. Chaperones are usually classified according to their molecular weight (HSP40, HSP60, HSP70, HSP90, HSP100 and the small HSPs). They are involved in a multitude of proteome-maintenance functions, including *de novo* folding, refolding of stress-denatured proteins, oligomeric assembly, protein trafficking and assistance in proteolytic degradation. The chaperones that participate broadly in *de novo* protein folding and refolding, such as the HSP70s, HSP90s and the chaperonins (HSP60s), are multicomponent molecular machines that promote folding through

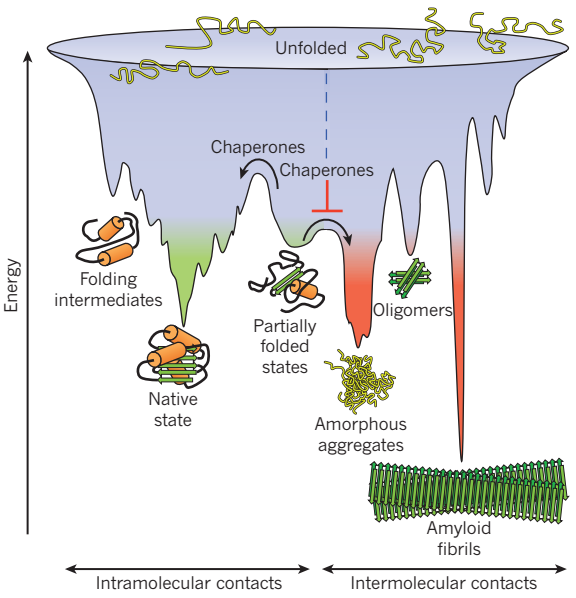


Figure 1 | Competing reactions of protein folding and aggregation. Scheme of the funnel-shaped free-energy surface that proteins explore as they move towards the native state (green) by forming intramolecular contacts (modified from refs 19 and 95). The ruggedness of the free-energy landscape results in the accumulation of kinetically trapped conformations that need to traverse free-energy barriers to reach a favourable downhill path. *In vivo*, these steps may be accelerated by chaperones[39,41,42]. When several molecules fold simultaneously in the same compartment, the free-energy surface of folding may overlap with that of intermolecular aggregation, resulting in the formation of amorphous aggregates, toxic oligomers or ordered amyloid fibrils (red). Fibrillar aggregation typically occurs by nucleation-dependent polymerization. It may initiate from intermediates populated during *de novo* folding or after destabilization of the native state (partially folded states) and is normally prevented by molecular chaperones.

ATP- and cofactor-regulated binding and release cycles. They typically recognize hydrophobic amino-acid side chains exposed by non-native proteins and may functionally cooperate with ATP-independent chaperones, such as the small HSPs, which function as 'holdases', buffering aggregation.

In the ATP-dependent mechanism of chaperone action, *de novo* folding and protein refolding is promoted through kinetic partitioning (Fig. 2). Chaperone binding (or rebinding) to hydrophobic regions of a non-native protein transiently blocks aggregation; ATP-triggered release allows folding to proceed. Importantly, although the HSP70s and the chaperonins both operate by this basic mechanism, they differ fundamentally in that the former (like all other ATP-dependent chaperones) release the substrate protein for folding into bulk solution, whereas the cylindrical chaperonins allow the folding of single protein molecules enclosed in a cage. The two systems act sequentially, whereby HSP70 interacts upstream with nascent and newly synthesized polypeptides and the chaperonins function downstream in the final folding of those proteins that fail to reach native state by cycling on HSP70 alone[20,21] (Figs 2 and 3). In the following sections, we will use the HSP70, chaperonin and HSP90 models to illustrate the basic mechanisms of the major cytosolic protein-folding machines. Client-specific chaperones that function downstream of folding in mediating the assembly of oligomeric complexes are not discussed (see, for example, refs 22 and 23).

The HSP70 system

The constitutively expressed (HSC70, also known as HSPA8) and stress-inducible forms of HSP70 are central players in protein folding and proteostasis control. Increasing HSP70 levels has also proven effective in preventing toxic protein aggregation in disease models[24].

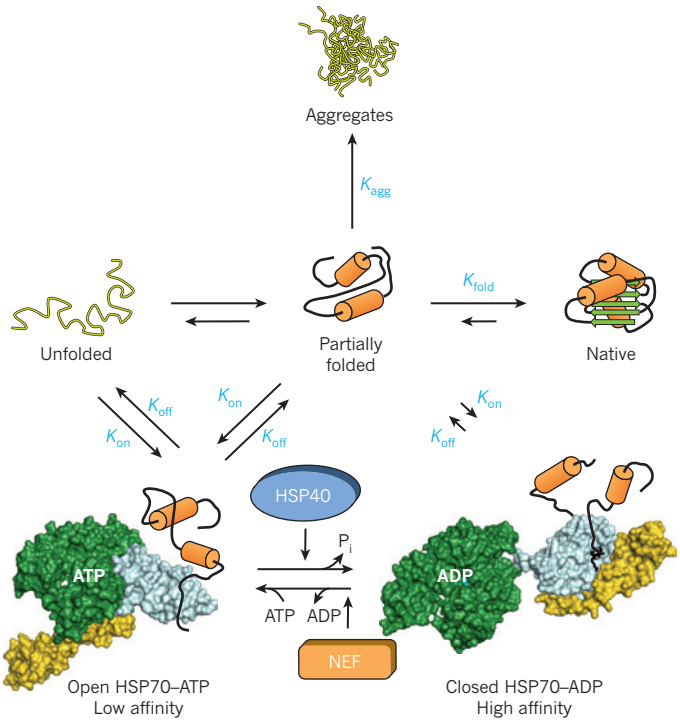


Figure 2 | The HSP70 chaperone cycle. HSP70 is switched between high- and low-affinity states for unfolded and partially folded protein by ATP binding and hydrolysis. Unfolded and partially folded substrate (nascent chain or stress-denatured protein), exposing hydrophobic peptide segments, is delivered to ATP-bound HSP70 (open; low substrate affinity with high on-rates and off-rates) by one of several HSP40 cofactors. The hydrolysis of ATP, which is accelerated by HSP40, results in closing of the α-helical lid of the peptide-binding domain (yellow) and tight binding of substrate by HSP70 (closed; high affinity with low on-rates and off-rates). Dissociation of ADP catalysed by one of several nucleotide-exchange factors (NEFs) is required for recycling. Opening of the α-helical lid, induced by ATP binding, results in substrate release. Folding is promoted and aggregation is prevented when both the folding rate constant (K_{fold}) is greater than the association constant (K_{on}) for chaperone binding (or rebinding) of partially folded states, and K_{on} is greater than intermolecular association by the higher-order aggregation rate constant K_{agg} ($K_{fold} > K_{on} > K_{agg}$) (kinetic partitioning). For proteins that populate misfolded states, K_{on} may be greater than K_{fold} ($K_{fold} \leq K_{on} > K_{agg}$). These proteins are stabilized by HSP70 in a non-aggregated state, but require transfer into the chaperonin cage for folding[14,20]. After conformational stress, K_{agg} may become faster than K_{on}, and aggregation occurs ($K_{agg} > K_{on} \geq K_{fold}$), unless chaperone expression is induced via the stress-response pathway. Structures in this figure relate to Protein Data Bank (PDB) accession codes 1DKG, 1DKZ, 2KHO and 2QXL. P_i, inorganic phosphate.

The ATP-dependent reaction cycle of HSP70 is regulated by chaperones of the HSP40 (also known as DnaJ) family and nucleotide-exchange factors[25,26]. Some of these factors are also involved in linking chaperone functions with the UPS and autophagy for the removal of misfolded proteins[27]. Binding and release by HSP70 is achieved through the allosteric coupling of a conserved amino-terminal ATPase domain with a carboxy-terminal peptide-binding domain, the latter consisting of a β-sandwich subdomain and an α-helical lid segment[25] (Fig. 2). The β-sandwich recognizes extended, ~seven-residue segments enriched in hydrophobic amino acids, preferentially when they are framed by positively charged residues[28]. Such segments occur on average every 50–100 amino acids in proteins, and the exposure of these fragments correlates with the aggregation propensity of the protein[29]. The α-helical lid and a conformational change in the β-sandwich domain regulate the affinity state for the peptide in an ATP-dependent manner[25]. In the ATP-bound state, the lid adopts an open conformation, resulting in high on rates and off rates for the peptide. Hydrolysis of ATP to ADP is strongly accelerated by HSP40, leading to lid closure and stable peptide binding (low on rates and off rates for the peptide substrate) (Fig. 2). HSP40 also interacts directly with unfolded polypeptides and can recruit HSP70 to protein substrates[20,26]. After ATP hydrolysis, a nucleotide-exchange factor binds to the HSP70 ATPase domain and catalyses ADP–ATP exchange, resulting in lid opening and substrate release. Release allows fast-folding molecules to bury hydrophobic residues, whereas molecules that need longer than a few seconds for folding will rebind to HSP70, thereby avoiding aggregation. HSP70 (re)binding may also result in conformational remodelling, perhaps removing kinetic barriers to the folding process[30].

Proteins that are unable to partition to fast-folding trajectories after HSP70 cycling may be transferred into the specialized environment of the chaperonin cage for folding. Among these are several essential proteins, such as actins and tubulins[31], which encounter high energetic barriers in folding and are completely unable to reach their native states spontaneously, even in dilute solution *in vitro*.

The chaperonins

Chaperonins are large double-ring complexes of ~800–900 kDa that function by globally enclosing substrate proteins up to ~60 kDa for folding. Group I chaperonins (also known as HSP60s in eukaryotes and GroEL in bacteria) have seven-membered rings in bacteria, mitochondria and chloroplasts, and functionally cooperate with HSP10 proteins (GroES in bacteria), which form the lid of the folding cage. The group II chaperonins in archaea (thermosome) and the eukaryotic cytosol (TRiC, also known as CCT) usually have eight-membered rings. They are independent of HSP10 factors.

The GroEL–GroES chaperonin system of *Escherichia coli* has been studied most extensively[19,32] (Fig. 3). GroEL interacts with at least 250 different cytosolic proteins. Most of these are between 20 and 50 kDa in size and have complex α/β or α+β domain topologies, such as the TIM barrel fold[14,33]. These proteins are stabilized by many long-range interactions and are thought to populate flexible, kinetically trapped folding intermediates exposing hydrophobic surfaces[34,35]. The apical domains of GroEL present hydrophobic amino-acid residues for substrate binding in the ring centre. Subsequent folding depends on global substrate encapsulation by GroES (Fig. 3). GroES binding is ATP regulated and is associated with a marked conformational change of GroEL that leads to the formation of a cage with a highly hydrophilic, net-negatively-charged inner wall[19,32,36]. Encapsulated protein is free to fold in this environment for ~10 seconds — the time needed for ATP hydrolysis in the GroES-bound ring (*cis* ring). Protein substrate leaves the cage after GroES dissociation, which is allosterically triggered by ATP binding in the opposite ring (*trans* ring). Not-yet folded substrate rapidly rebinds to GroEL for further folding attempts.

Enclosing unfolded protein, one molecule at a time, avoids disruption of folding by aggregation or (re)binding to upstream chaperones. In addition, an effect of steric confinement probably modulates the folding-energy landscape. Although the chaperonin functions as a passive-aggregation prevention device for some proteins[32,37], encapsulation can also accelerate folding substantially[37–39]. This rate acceleration may be due to steric confinement, entropically destabilizing collapsed yet flexible folding intermediates, and promoting their conversion to more compact, native-like conformations. As shown recently, the effect of the folding cage may be comparable to the role of disulphide bonds in restricting conformational space in the folding of secretory proteins[39]. Furthermore, repeated unfolding events in successive binding and release cycles have been suggested to reverse misfolded, kinetically trapped states that are stabilized by non-native interactions[40–42]. Thus, the chaperonins may be able to remove both entropic and enthalpic barriers in rugged free-energy landscapes of folding (Fig. 1).

TRiC, the group II chaperonin in the eukaryotic cytosol, consists of eight paralogous subunits per ring[31,43,44]. All group II chaperonins deviate from GroEL in that their apical domains contain finger-like

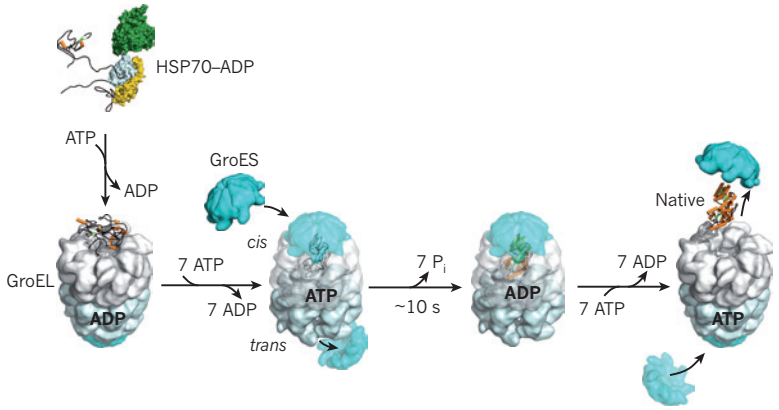


Figure 3 | Folding in the GroEL–GroES chaperonin cage. Substrate binding to GroEL (after transfer from HSP70) may result in local unfolding[42]. ATP binding then triggers a conformational rearrangement of the GroEL apical domains. This is followed by the binding of GroES (forming the *cis* complex) and substrate encapsulation for folding. At the same time, ADP and GroES dissociate from the opposite (*trans*) GroEL ring, allowing the release of substrate that had been enclosed in the former *cis* complex[32] (omitted from the figure for simplicity). The new substrate remains encapsulated, free to fold, for the time needed to hydrolyse the seven ATP molecules in the newly formed *cis* complex (~10 s). Binding of ATP and GroES to the *trans* ring causes the opening of the *cis* complex. A symmetrical GroEL–$(GroES)_2$ complex may form transiently. Structural model is based on PDB accession 1AON.

protrusions, which act as an iris-like, built-in lid and replace the function of GroES. These segments open and close in an ATP-dependent protein-encapsulation cycle, similar in principle to that of GroEL–GroES[44]. However, the TRiC reaction cycle is much slower than that of GroEL, probably providing a substantially longer period of protein encapsulation and folding in the cage[45]. TRiC interacts with approximately 10% of newly synthesized cytosolic proteins, including actin and tubulins[31,43]. Interestingly, TRiC also functions in preventing the accumulation of toxic aggregates by the Huntington's disease protein[46–48].

The HSP90 system

HSP90 forms a proteostasis hub that controls numerous important signalling pathways in eukaryotic cells[49]. These pleiotropic functions include, among others, cell-cycle progression, telomere maintenance, apoptosis, mitotic signal transduction, vesicle-mediated transport, innate immunity and targeted protein degradation. Indeed, the evolution and maintenance of these functional networks is thought to depend on the ability of HSP90 to buffer the effects of structurally destabilizing mutations in the underlying protein complexes, thereby allowing the acquisition of new traits[12].

HSP90 functions downstream of HSP70 in the structural maturation and conformational regulation of numerous signal-transduction molecules, such as kinases and steroid receptors[49,50]. It cooperates in this process with several regulators and co-chaperones, many of which use tetratricopeptide repeat (TPR) domains to dock onto HSP90. For example, the TPR protein HOP provides a direct link between HSP70 and HSP90, allowing substrate transfer[51]. Although the mechanism by which HSP90 and its cofactors mediate conformational changes in substrate proteins is not yet understood[52], recent crystal structures of full-length HSP90s provided long-awaited information[53,54]. HSP90 functions as a dimer of subunits that are assembled by their C-terminal domains. An N-terminal domain binds and hydrolyses ATP and is joined to the C-terminal domain by a middle domain (Fig. 4). The middle domain participates in substrate binding and interacts with the co-chaperone AHA1. Similar to other chaperones, the HSP90 dimer undergoes an ATP-driven reaction cycle that is accompanied by considerable structural rearrangement[25] (Fig. 4). ATP binding leads to the dimerization of the N-terminal domains, forming the HSP90 'molecular clamp'. This results in a compaction of the HSP90 dimer, in which the individual monomers twist around each other. After hydrolysis, the ATPase domains dissociate, and the HSP90 monomers separate N-terminally. Various cofactors regulate this cycle: CDC37, which delivers certain kinase substrates to HSP90, inhibits the ATPase activity, and HOP inhibits N-terminal dimerization. AHA1 stimulates ATP hydrolysis, whereas p23 stabilizes the dimerized form of HSP90 before ATP hydrolysis. These factors are thought to adjust the kinetic properties of the cycle to achieve certain conformational transitions in HSP90-bound substrates, as well as their release from HSP90.

How HSP90 recruits different types of substrate protein with the help of various co-chaperones remains enigmatic. HSP90 appears to have several substrate-interaction regions, and the binding strength seems to be strongly influenced by the structural flexibility of the substrate[52], in line with the proposed role of HSP90 as an evolutionary capacitor in protecting mutated protein variants from degradation[12]. Because several HSP90 substrates are kinases with well-documented roles in tumour development, the inhibition of HSP90 with drugs

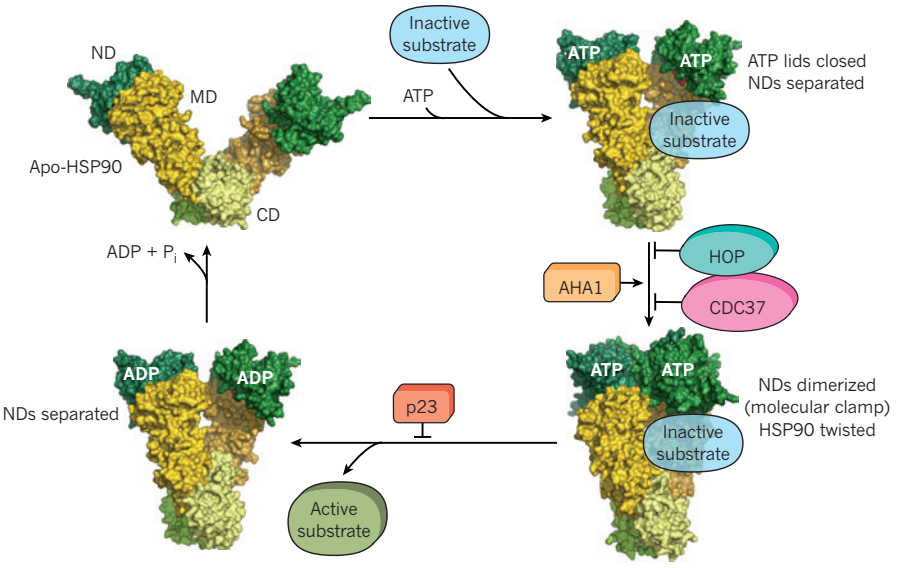


Figure 4 | ATPase cycle of the HSP90 chaperone system. Clockwise from top left, ATP binding to the N-terminal ATPase domain (ND) of apo-HSP90 induces a conformational change and the closure of the ATP lid in the ND. After lid closure, the NDs dimerize, forming the closed HSP90 dimer (molecular clamp) with twisted subunits. This metastable conformation is committed for ATP hydrolysis. After hydrolysis, the NDs dissociate. The inactive substrate molecule interacts mostly with the middle domain (MD) and is conformationally activated as HSP90 proceeds through the ATPase cycle. The cofactors CDC37, HOP, AHA1 and p23 accelerate or slow certain steps of the cycle. Structures relate to PDB accessions 2IOQ, 2O1U, 2CG9 and 2O1V. CD, C-terminal ATPase domain.

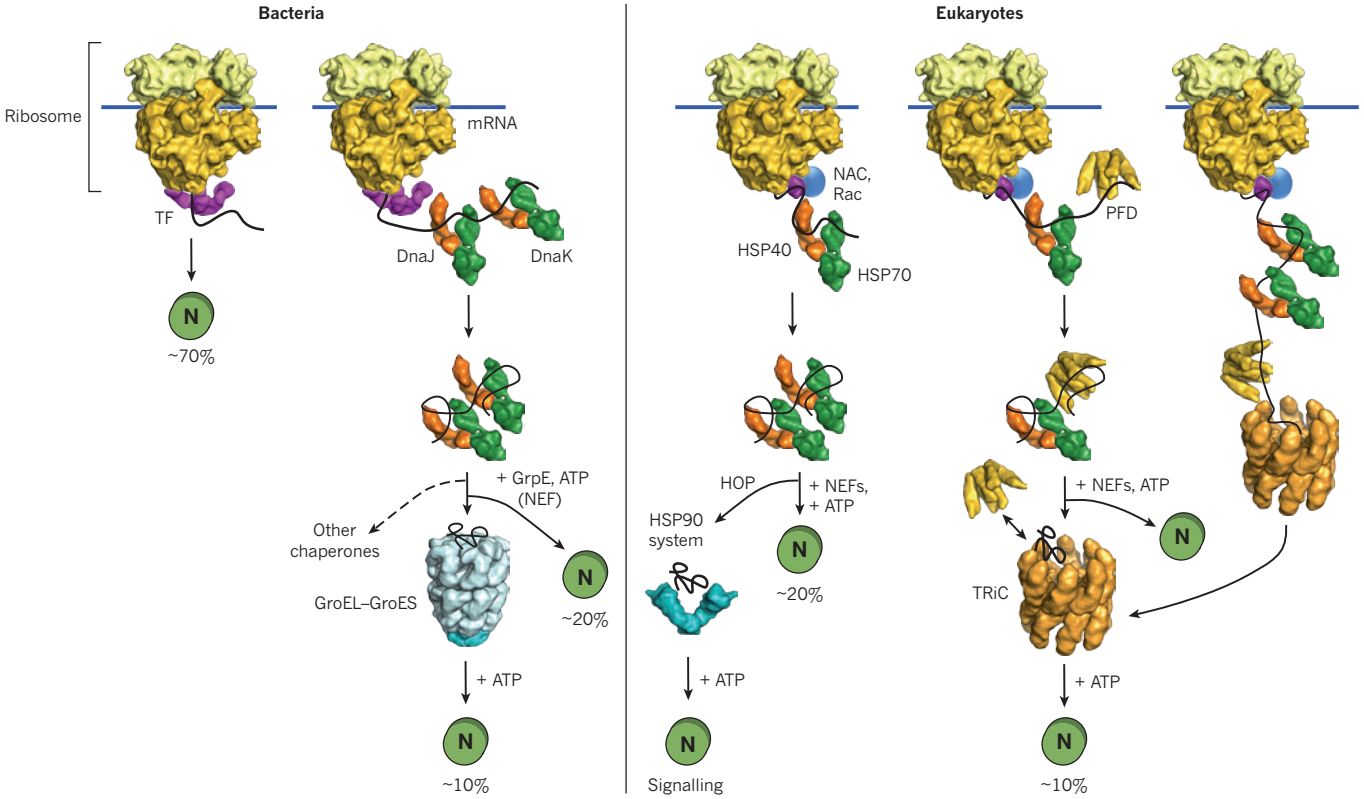


Figure 5 | Organization of chaperone pathways in the cytosol. In bacteria (left) and eukaryotes (right), chaperones that function in stabilizing nascent polypeptides on ribosomes and in initiating folding cooperate with machinery that acts downstream in completing folding[7,19–21,31]. The number of interacting substrates is indicated as a percentage of the total proteome. The first category of factors includes chaperones that bind in close proximity to the ribosomal polypeptide exit site, such as trigger factor (TF) in bacteria and specialized HSP70 complexes (ribosome-associated complex (Rac) in *Saccharomyces cerevisiae*, MPP11 and HSP70L1 in mammalian cells) and nascent-chain-associated complex (NAC) in eukaryotes). These chaperones bind hydrophobic chain segments. Non-ribosome-bound members of the HSP70 family (DnaK in bacteria and HSC70 in eukaryotes) function as second-tier chaperones for longer nascent chains, mediating co- or post-translational folding. They also distribute subsets of proteins to downstream chaperones, such as the chaperonins (GroEL in bacteria and TRiC in eukaryotes)[14,20,21] and HSP90[52]. Substrate transfer from HSC70 to HSP90 is promoted by the coupling protein HOP. Dashed arrow indicates that pathway is not well established. Structures relate to PDB accessions 1DKG, 1DKZ, 2KHO, 1W26, 3IYF, 2AVY, 2AW4, 1FXK, 3LKX, 2IOQ, 2QWR and 1NLT. N, native protein; GrpE, protein GrpE; mRNA, messenger RNA; PFD, prefoldin.

such as geldanamycin has emerged as a promising strategy for the treatment of certain cancers[55]. These drugs specifically inhibit the ATPase function of HSP90. They will probably prove useful not only in cancer therapy but also in the treatment of viral diseases, owing to the fact that various pathogenic viruses hijack the HSP90 system and use it for capsid assembly[56]. However, the global inhibition of HSP90 is likely to result in a marked derangement of cellular circuitry, and it would be desirable to find ways to inhibit only specific aspects of HSP90 function.

From ribosome to folded protein

The vectorial synthesis of polypeptides on the ribosome has important implications in the folding process that are only partly understood. Key questions concern the stage at which the nascent chain begins to fold and the extent to which the translation process modulates the free-energy landscape of folding. In addressing these issues, it is useful to first consider small, single-domain proteins, which tend to fold spontaneously *in vitro*. The translation process for such proteins seems to increase the risk of misfolding and aggregation considerably, because an incomplete nascent polypeptide is unable to fold into a stable native conformation[57,58] and the local concentration of nascent chains in the context of polyribosomes is very high. Furthermore, the exit channel of the large ribosomal subunit, which is ~100 Å long but at most 20 Å wide, is unfavourable to folding beyond α-helices and small tertiary elements that may begin to form near the tunnel exit[59–61]; it thus prevents the C-terminal 30–40 amino-acid residues of the chain from participating in long-range interactions that are necessary for cooperative domain folding. As a consequence, productive folding may occur only after the complete protein has emerged from the ribosome[57,62]. Because translation is relatively slow (~4–20 amino acids s^{-1}), nascent chains are exposed in partially folded, aggregation-sensitive states for considerable periods of time. Moreover, non-native intrachain contacts formed during translation or interactions with the highly charged ribosomal surface could delay folding after completion of synthesis. For these reasons, nascent chains are thought to interact co-translationally with ribosome-bound chaperones, which inhibit their premature (mis)folding and maintain the nascent chain in a non-aggregated, folding-competent state (Fig. 5). For example, the bacterial trigger factor[63] binds to the small titin I27 chain (~120 amino acids) throughout translation[64], presumably delaying chain collapse until the complete β-sandwich domain has emerged from the ribosome and is available for folding. Moreover, the aggregation of nascent chains is disfavoured by the densely packed, pseudohelical arrangement of ribosomes in polyribosome complexes — an organization that maximizes the distance between nascent-chain exit sites on adjacent ribosomes[65].

Although single-domain proteins will reach their native state post-translationally, multidomain proteins may undergo domain-wise co-translational folding, as independently folding structural units (~50–300 amino acids in length) emerge sequentially from the

ribosome[66,67]. This process avoids non-native interdomain contacts, thus smoothing the folding-energy landscape for large proteins[66,68]. Sequential domain folding during translation, which is highly efficient on eukaryotic ribosomes, probably promoted the explosive evolution of complex multidomain proteins in eukaryotes[66,68]. Co-translational folding is thought to be aided by the slower elongation speed of eukaryotic ribosomes (~4 amino acids s^{-1} in eukaryotes versus ~20 amino acids s^{-1} in bacteria) and as a result of various adaptations of the folding machinery. For example, eukaryotic ribosomes bind specialized HSP70 chaperone complexes (Fig. 5) and the binding and release of the canonical HSC70 from nascent chains may be coordinated with translation speed so as to support domain-wise folding. The eukaryotic chaperonin TRiC is recruited to nascent chains by HSC70 (ref. 69) and other upstream factors, such as prefoldin[31], allowing co-translational folding. Moreover, fine-tuning of co-translational folding may be achieved by translational pausing at rare codons[70]. Overall, the eukaryotic translation and chaperone machinery has been highly optimized through evolution, ensuring efficient folding for the bulk of newly synthesized proteins[71].

The chaperone pathways operating in the endoplasmic reticulum (ER) follow analogous organizational principles, but specialized machinery is used in disulphide-bond formation and the glycosylation of many secretory proteins[72].

Proteome maintenance and the proteostasis network

Although it is generally accepted that the chaperone machinery is required for initial protein folding, we are only beginning to appreciate the extent to which many proteins depend on macromolecular assistance throughout their cellular lifetime to maintain or regain their functionally active conformations. Compared with prokaryotes, the proteomes of eukaryotic cells are highly complex, comprising a much greater number and diversity of multidomain proteins. In the dynamic cellular environment, these proteins constantly face numerous challenges to their folded states; these result from post-translational modifications (phosphorylation and acetylation), changes in cell physiology and alterations in the composition and concentration of small-molecule ligands that may influence protein stability[4]. Moreover, 20–30% of all proteins in mammalian cells are intrinsically unstructured[3]; that is, they may adopt defined three-dimensional conformations only after binding to other macromolecules or membrane surfaces. Such proteins probably require assistance to avoid aberrant interactions and aggregation, particularly when their concentration is increased and they are not in complexes with partner molecules[73].

These considerations help to explain why cells must invest in an extensive network of factors, comprising ~800 proteins in human cells (~200 chaperones and co-chaperones and ~600 UPS and autophagy components), which cooperate to maintain the conformational integrity of the proteome and provide adaptation to changes in the environment. This proteostasis network integrates general and specialized chaperone components for proper protein folding and trafficking with the machinery for disaggregation and proteolytic degradation of irreversibly misfolded proteins (the UPS and the autophagy system) (Fig. 6). The remarkable complexity of the system arises from the expansion, in multicellular organisms, of the diversity of regulatory components for the major chaperone systems (HSP70 and HSP90)[26] and of factors functionally coupling these chaperones with the UPS and the autophagy system[27,74,75]. For example, various HSP70 cofactors, such as the BCL2-associated athanogene (BAG) family proteins and certain HSP40s, contain ubiquitin-like or ubiquitin-interacting domains[74]. The HSP70 and HSP90 cofactor known as carboxyl terminus of Hsp70-interacting protein (CHIP) has E3 ubiquitin ligase activity and channels certain mutant or damaged proteins towards proteasomal degradation[74]. Notably, CHIP is only one of several hundred different E3 ligases, which reflects the enormous importance of proteolytic pathways for proteostasis and cell regulation. Interestingly, whereas the clearance of misfolded protein species by the UPS requires that these molecules are maintained in a non-aggregated state by chaperones, disposal by autophagy is thought to involve active mechanisms to force such molecules into larger, presumably less toxic, aggregates[76,77]. These inclusions are often deposited at specific subcellular sites close to the microtubule-organizing centre, referred to as the aggresome[78].

The proteostasis network is regulated by several interconnected signalling pathways, some of which are stress responsive and ensure that cellular protein folding and/or degradation is adapted to avoid the accumulation of misfolded and aggregation-prone species (Fig. 6). These pathways include the cytosolic stress response and the unfolded protein response of the ER and mitochondria, as well as signalling pathways that control ribosome biogenesis and translational capacity (Box 1). How the inputs from these different branches are coordinated and fine-tuned is only partly understood, but proteostasis capacity and responsiveness to stress may vary considerably in different cell types[79].

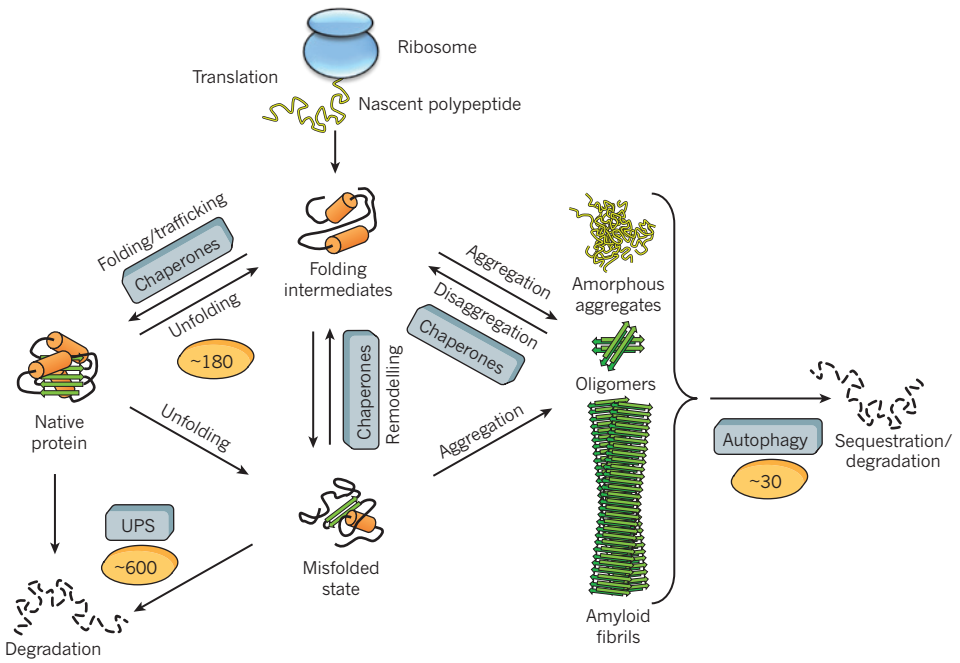


Figure 6 | Protein fates in the proteostasis network. The proteostasis network integrates chaperone pathways for the folding of newly synthesized proteins, for the remodelling of misfolded states and for disaggregation with the protein degradation mediated by the UPS and the autophagy system. Approximately 180 different chaperone components and their regulators orchestrate these processes in mammalian cells, whereas the UPS comprises ~600 and the autophagy system ~30 different components. The primary effort of the chaperone system is in preventing aggregation, but machinery for the disaggregation of aggregated proteins has been described in bacteria and fungi, involving oligomeric AAA^{+}-proteins such as HSP104 and the *E. coli* molecular chaperone protein ClpB, which cooperate with HSP70 chaperones[25]. A similar activity has been detected in metazoans, but the components involved have not yet been defined[83].

BOX 1

Signalling pathways in proteostasis and ageing

The expression of stress-inducible chaperone proteins (such as HSP70, HSP40, HSP90 and small HSPs) in the cytosol is governed by the heat-shock response[96]. The genes encoding these proteins are transcriptionally regulated by the HSF-1 and FOXO (DAF-16 in *C. elegans*) transcription factors.

The unfolded protein response (UPR)[72] of the ER adjusts the folding capacity of the secretory pathway by upregulating ER chaperones and/or attenuating protein synthesis by means of the transcription factors IRE1, PERK and ATF6.

The mitochondrial UPR[97,98] is activated by conformational stress in mitochondria and increases resistance to oxidative damage.

Ageing and longevity pathways are coupled to the regulation of stress-protective pathways[5,99]. Specifically, the upregulation of stress-protection factors such as chaperones by HSF-1 and FOXO is required for the lifespan-extending effect of mutations in the insulin and insulin-like growth factor I (IGF-I) receptor pathway. Autophagy, a process required for the recycling of organelles and the removal of large protein aggregates, is also necessary for lifespan extension and youthfulness in *C. elegans*. Autophagy is downregulated by the mammalian target of rapamycin (TOR) kinase when nutrients are plentiful[99] and is upregulated by FOXO[81]. Dietary restriction, which extends lifespan in model organisms, is also coupled with HSF-1 and FOXO activation[81,100].

Proteostasis collapse in ageing and disease

The accumulation of misfolded and/or oxidized proteins in cells during ageing is a challenge to the proteostasis system and eventually results in the deposition of aggregates, as shown in model organisms such as *Caenorhabditis elegans* and *Drosophila*[80,81]. The inability of cells to restore normal proteostasis may result in disease, and even in cell death. Indeed, numerous diseases are now recognized to be associated with aberrant protein folding and are usually categorized as loss-of-function or toxic gain-of-function diseases, although specific pathological states often show elements of both groups. The former are generally caused by inherited mutations and include numerous disorders such as cystic fibrosis, lysosomal storage diseases and α1-antitrypsin deficiency. The latter, gain-of-function disorders, include type 2 diabetes and the major neurodegenerative conditions (Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis and Alzheimer's disease) and are either sporadic or caused by mutations that render specific proteins more aggregation prone. These gain-of-function diseases are typically age related and are caused by the accumulation of amyloid or amyloid-like aggregates of the disease protein. A plausible explanation for the late onset of these diseases is provided by recent evidence from model organisms that the signalling pathways that regulate proteostasis are integrated with the genetic and epigenetic pathways that control longevity[82,83] (Box 1). Thus, the age-related decline in proteostasis and specifically in the inability to upregulate chaperones in response to conformational stresses would trigger disease manifestation and, in turn, accelerate proteostasis collapse[81,84,85].

Although the toxic principle operating in these disorders is far from understood, a consensus is emerging that soluble oligomeric aggregates, which may be 'on-pathway' or 'off-pathway' towards fibril formation, are the primary cytotoxic species[16,18] (Fig. 1). One prominent hypothesis suggests that these oligomers expose promiscuous hydrophobic surfaces that can mediate aberrant interactions with several other proteins or with cellular membranes[16,17]. In support of this proposal, a recent proteomics study in human cells showed that certain metastable proteins are targeted preferentially by such interactions, resulting in their co-aggregation with the amyloidogenic disease protein[86]. The co-aggregating proteins are generally large in size and are enriched in intrinsically unstructured regions, properties that are coupled with a high degree of functionality. Accordingly, they tend to occupy essential hub positions in cellular protein networks, including transcriptional regulation, translation and maintenance of cell architecture, suggesting that their sequestration by the amyloid aggregates results in multifactorial toxicity. An interesting manifestation of this toxicity mechanism is the recent demonstration that aggregating mutant p53 may exert dominant oncogenic potential by sequestering wild-type p53 into co-aggregates, resulting in a complete loss of p53 function[87].

Aggregate toxicity may be exacerbated by the inability of affected cells to adequately respond to stress stimuli[86]. This is consistent with recent evidence that aberrantly folded protein species may interfere with central proteostasis functions, including protein folding and clearance mechanisms[88,89]. Notably, the overexpression of members of the HSP70 system has been shown to inhibit the formation of toxic oligomers and to prevent the formation of amyloid aggregates for different disease proteins[24,90,91]. In the case of polyglutamine-repeat proteins, which cause Huntington's disease and several related neurodegenerative disorders, HSP70 cooperates with the chaperonin TRiC to prevent the accumulation of potentially toxic oligomers[47], which is reminiscent of the functional cooperation between these chaperone systems in *de novo* protein folding.

On the basis of these findings, the pharmacological upregulation of chaperone function promises to open up new strategies for the treatment of numerous pathological states associated with aberrant folding and aggregation. Proof-of-principle experiments using small-molecule compounds to increase chaperone synthesis and rebalance proteostasis (for example, by activating heat-shock transcription factor-1 (HSF-1)-regulated pathways) have already demonstrated efficacy in loss-of-function and toxic gain-of-function disease models[5,6,92,93]. Likewise, recently identified proteasome activators[94] have the potential to accelerate the clearance of toxic protein species, particularly when applied in combination with chaperone upregulation. Unlike conventional drugs, such 'proteostasis regulators' would not be disease-specific or protein-specific, and thus may be applicable to a whole group of related diseases — a new concept in medical practice.

Outlook

Studies over the past two decades have provided fascinating insight into the mechanics of chaperone-assisted protein folding, but there are still major gaps in our understanding of how the pathways of folding in the cell differ from those studied in the test tube. Progress is being held back by the problem that the sophisticated biophysical methods used to characterize folding intermediates *in vitro* are not easily transferable to the *in vivo* situation. Major innovation potential can thus be expected from the development of advanced imaging techniques, eventually allowing us to monitor conformational changes in a single polypeptide chain as it emerges from the ribosome, performs its biological function and is finally degraded in the living cell. Much research will also be stimulated by the emerging concept that molecular chaperones function as the central element of a much larger cellular network of proteostasis control, comprising, in addition, the protein biogenesis machinery as well as the UPS and the autophagy system. Unravelling the complex regulatory circuitry of this network and understanding why it loses its grip during ageing will pose a major challenge for years to come. Solving this problem will require a broad systems-biology approach relying on a combination of ribosome profiling, quantitative proteomics and computational modelling. How cells react to conformational stress or proteostasis deficiency at the proteome level is unclear. Key questions include determining how certain aberrantly folding proteins aggregate

into toxic species whereas others are degraded, how the composition of the proteosome changes during ageing, what the signature of a youthful proteome is, and how we can find ways to maintain it for longer as we age. Addressing these and related issues not only offers great opportunities for intervention with numerous, currently incurable diseases but will also eventually reveal the fundamentally important relationship between proteostasis and longevity. ■

1. Dobson, C. M., Sali, A. & Karplus, M. Protein folding — a perspective from theory and experiment. *Angew. Chem. Int. Edn Engl.* **37,** 868–893 (1998).
2. Bartlett, A. I. & Radford, S. E. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nature Struct. Mol. Biol.* **16,** 582–588 (2009).
3. Dunker, A. K., Silman, I., Uversky, V. N. & Sussman, J. L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **18,** 756–764 (2008).
4. Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W. & Balch, W. E. Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem.* **78,** 959–991 (2009).
5. Morimoto, R. I. Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. *Genes Dev.* **22,** 1427–1438 (2008).
6. Balch, W. E., Morimoto, R. I., Dillin, A. & Kelly, J. W. Adapting proteostasis for disease intervention. *Science* **319,** 916–919 (2008).
7. Hartl, F. U. Molecular chaperones in cellular protein folding. *Nature* **381,** 571–580 (1996).
8. Kubelka, J., Hofrichter, J. & Eaton, W. A. The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* **14,** 76–88 (2004).
9. Herbst, R., Schafer, U. & Seckler, R. Equilibrium intermediates in the reversible unfolding of firefly (*Photinus pyralis*) luciferase. *J. Biol. Chem.* **272,** 7099–7105 (1997).
10. Ellis, R. J. & Minton, A. P. Protein aggregation in crowded environments. *Biol. Chem.* **387,** 485–497 (2006).
11. Tokuriki, N. & Tawfik, D. S. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* **459,** 668–671 (2009).
12. Rutherford, S. L. & Lindquist, S. Hsp90 as a capacitor for morphological evolution. *Nature* **396,** 336–342 (1998).
This seminal study puts forward the idea that chaperones function in buffering the otherwise deleterious consequences of mutations.
13. Skach, W. R. Cellular mechanisms of membrane protein folding. *Nature Struct. Mol. Biol.* **16,** 606–612 (2009).
14. Kerner, M. J. *et al.* Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* **122,** 209–220 (2005).
15. Eichner, T., Kalverda, A. P., Thompson, G. S., Homans, S. W. & Radford, S. E. Conformational conversion during amyloid formation at atomic resolution. *Mol. Cell* **41,** 161–172 (2011).
This exciting paper describes, at atomic resolution, the structural features of a non-native folding intermediate that are critical for amyloidogenic aggregation.
16. Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75,** 333–366 (2006).
17. Bolognesi, B. *et al.* ANS binding reveals common features of cytotoxic amyloid species. *ACS Chem. Biol.* **5,** 735–740 (2010).
This paper provides evidence that the exposure of hydrophobic surfaces by oligomeric aggregation intermediates correlates with their toxicity.
18. Kayed, R. *et al.* Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300,** 486–489 (2003).
19. Hartl, F. U. & Hayer-Hartl, M. Converging concepts of protein folding *in vitro* and *in vivo*. *Nature Struct. Mol. Biol.* **16,** 574–581 (2009).
20. Langer, T. *et al.* Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature* **356,** 683–689 (1992).
21. Frydman, J., Nimmesgern, E., Ohtsuka, K. & Hartl, F. U. Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* **370,** 111–117 (1994).
22. Ellis, R. J. Molecular chaperones: assisting assembly in addition to folding. *Trends Biochem. Sci.* **31,** 395–401 (2006).
23. Liu, C. *et al.* Coupled chaperone action in folding and assembly of hexadecameric Rubisco. *Nature* **463,** 197–202 (2010).
24. Auluck, P. K., Chan, H. Y. E., Trojanowski, J. Q., Lee, V. M. Y. & Bonini, N. M. Chaperone suppression of α-synuclein toxicity in a *Drosophila* model for Parkinson's disease. *Science* **295,** 865–868 (2002).
25. Mayer, M. P. Gymnastics of molecular chaperones. *Mol. Cell* **39,** 321–331 (2010).
26. Kampinga, H. H. & Craig, E. A. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Rev. Mol. Cell Biol.* **11,** 579–592 (2010).
27. Arndt, V. *et al.* Chaperone-assisted selective autophagy is essential for muscle maintenance. *Curr. Biol.* **20,** 143–148 (2010).
28. Rüdiger, S., Buchberger, A. & Bukau, B. Interaction of Hsp70 chaperones with substrates. *Nature Struct. Biol.* **4,** 342–349 (1997).
A key paper in understanding how chaperones recognize their substrates.
29. Rousseau, F., Serrano, L. & Schymkowitz, J. W. H. How evolutionary pressure against protein aggregation shaped chaperone specificity. *J. Mol. Biol.* **355,** 1037–1047 (2006).
30. Sharma, S. K., De los Rios, P., Christen, P., Lustig, A. & Goloubinoff, P. The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. *Nature Chem. Biol.* **6,** 914–920 (2010).
31. Frydman, J. Folding of newly translated proteins *in vivo*: the role of molecular chaperones. *Annu. Rev. Biochem.* **70,** 603–647 (2001).
32. Horwich, A. L. & Fenton, W. A. Chaperonin-mediated protein folding: using a central cavity to kinetically assist polypeptide chain folding. *Q. Rev. Biophys.* **42,** 83–116 (2009).
33. Fujiwara, K., Ishihama, Y., Nakahigashi, K., Soga, T. & Taguchi, H. A systematic survey of *in vivo* obligate chaperonin-dependent substrates. *EMBO J.* **29,** 1552–1564 (2010).
34. Raineri, E., Ribeca, P., Serrano, L. & Maier, T. A more precise characterization of chaperonin substrates. *Bioinformatics* **26,** 1685–1689 (2010).
35. Tartaglia, G. G., Dobson, C. M., Hartl, F. U. & Vendruscolo, M. Physicochemical determinants of chaperone requirements. *J. Mol. Biol.* **400,** 579–588 (2010).
36. Xu, Z. H., Horwich, A. L. & Sigler, P. B. The crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ chaperonin complex. *Nature* **388,** 741–749 (1997).
37. Brinker, A. *et al.* Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* **107,** 223–233 (2001).
38. Tang, Y. C. *et al.* Structural features of the GroEL–GroES nano-cage required for rapid folding of encapsulated protein. *Cell* **125,** 903–914 (2006).
39. Chakraborty, K. *et al.* Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. *Cell* **142,** 112–122 (2010).
40. Thirumalai, D. & Lorimer, G. H. Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **30,** 245–269 (2001).
41. Lin, Z., Madan, D. & Rye, H. S. GroEL stimulates protein folding through forced unfolding. *Nature Struct. Mol. Biol.* **15,** 303–311 (2008).
42. Sharma, S. *et al.* Monitoring protein conformation along the pathway of chaperonin-assisted folding. *Cell* **133,** 142–153 (2008).
43. Munoz, I. G. *et al.* Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. *Nature Struct. Mol. Biol.* **18,** 14–19 (2011).
44. Douglas, N. R. *et al.* Dual action of ATP hydrolysis couples lid closure to substrate release into the Group II chaperonin chamber. *Cell* **144,** 240–252 (2011).
45. Reissmann, S., Parnot, C., Booth, C. R., Chiu, W. & Frydman, J. Essential function of the built-in lid in the allosteric regulation of eukaryotic and archaeal chaperonins. *Nature Struct. Mol. Biol.* **14,** 432–440 (2007).
46. Kitamura, A. *et al.* Cytosolic chaperonin prevents polyglutamine toxicity with altering the aggregation state. *Nature Cell Biol.* **8,** 1163–1170 (2006).
47. Behrends, C. *et al.* Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. *Mol. Cell* **23,** 887–897 (2006).
48. Tam, S., Geller, R., Spiess, C. & Frydman, J. The chaperonin TRiC controls polyglutamine aggregation and toxicity through subunit-specific interactions. *Nature Cell Biol.* **8,** 1155–1162 (2006).
49. Taipale, M., Jarosz, D. F. & Lindquist, S. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nature Rev. Mol. Cell Biol.* **11,** 515–528 (2010).
50. McClellan, A. J. *et al.* Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches. *Cell* **131,** 121–135 (2007).
51. Scheufler, C. *et al.* Structure of TPR domain–peptide complexes: critical elements in the assembly of the Hsp70–Hsp90 multichaperone machine. *Cell* **101,** 199–210 (2000).
52. Wandinger, S. K., Richter, K. & Buchner, J. The Hsp90 chaperone machinery. *J. Biol. Chem.* **283,** 18473–18477 (2008).
53. Ali, M. M. U. *et al.* Crystal structure of an Hsp90–nucleotide–p23/Sba1 closed chaperone complex. *Nature* **440,** 1013–1017 (2006).
54. Shiau, A. K., Harris, S. F., Southworth, D. R. & Agard, D. A. Structural analysis of *E. coli* hsp90 reveals dramatic nucleotide-dependent conformational rearrangements. *Cell* **127,** 329–340 (2006).
55. Neckers, L. Heat shock protein 90: the cancer chaperone. *J. Biosci.* **32,** 517–530 (2007).
56. Geller, R., Vignuzzi, M., Andino, R. & Frydman, J. Evolutionary constraints on chaperone-mediated folding provide an antiviral approach refractory to development of drug resistance. *Genes Dev.* **21,** 195–205 (2007).
This seminal study describes the requirement of HSP90 in viral assembly, outlining a strategy for antiviral treatment based on HSP90 inhibition.
57. Eichmann, C., Preissler, S., Riek, R. & Deuerling, E. Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc. Natl Acad. Sci. USA* **107,** 9111–9116 (2010).
58. Cabrita, L. D., Hsu, S. T., Launay, H., Dobson, C. M. & Christodoulou, J. Probing ribosome-nascent chain complexes produced *in vivo* by NMR spectroscopy. *Proc. Natl Acad. Sci. USA* **106,** 22239–22244 (2009).
59. Lu, J. L. & Deutsch, C. Folding zones inside the ribosomal exit tunnel. *Nature Struct. Mol. Biol.* **12,** 1123–1129 (2005).
60. Woolhead, C. A., McCormick, P. J. & Johnson, A. E. Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* **116,** 725–736 (2004).
61. O'Brien, E. P., Hsu, S.-T. D., Christodoulou, J., Vendruscolo, M. & Dobson, C. M. Transient tertiary structure formation within the ribosome exit port. *J. Am. Chem. Soc.* **132,** 16928–16937 (2010).
62. Elcock, A. H. Molecular simulations of cotranslational protein folding: fragment stabilities, folding cooperativity, and trapping in the ribosome. *PLoS Comput Biol.* **2,** e98 (2006).
63. Ferbitz, L. *et al.* Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* **431,** 590–596 (2004).
64. Kaiser, C. M. *et al.* Real-time observation of trigger factor function on translating ribosomes. *Nature* **444,** 455–460 (2006).
65. Brandt, F. *et al.* The native 3D organization of bacterial polysomes. *Cell* **136,** 261–271 (2009).

66. Netzer, W. J. & Hartl, F. U. Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **388,** 343–349 (1997).
67. Frydman, J., Erdjument-Bromage, H., Tempst, P. & Hartl, F. U. Co-translational domain folding as the structural basis for the rapid *de novo* folding of firefly luciferase. *Nature Struct. Biol.* **6,** 697–705 (1999).
68. Agashe, V. R. *et al.* Function of trigger factor and DnaK in multidomain protein folding: increase in yield at the expense of folding speed. *Cell* **117,** 199–209 (2004).
69. Cuellar, J. *et al.* The structure of CCT–Hsc70$_{NBD}$ suggests a mechanism for Hsp70 delivery of substrates to the chaperonin. *Nature Struct. Mol. Biol.* **15,** 858–864 (2008).
70. Zhang, G. & Ignatova, Z. Generic algorithm to predict the speed of translational elongation: implications for protein biogenesis. *PLoS ONE* **4,** e5036 (2009).
71. Vabulas, R. M. & Hartl, F. U. Protein synthesis upon acute nutrient restriction relies on proteasome function. *Science* **310,** 1960–1963 (2005).
72. Buchberger, A., Bukau, B. & Sommer, T. Protein quality control in the cytosol and the endoplasmic reticulum: brothers in arms. *Mol. Cell* **40,** 238–252 (2010).
73. Vavouri, T., Semple, J. I., Garcia-Verdugo, R. & Lehner, B. Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. *Cell* **138,** 198–208 (2009).
74. Arndt, V., Rogon, C. & Hohfeld, J. To be, or not to be — molecular chaperones in protein degradation. *Cell. Mol. Life Sci.* **64,** 2525–2541 (2007).
75. Gamerdinger, M. *et al.* Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *EMBO J.* **28,** 889–901 (2009).
76. Kaganovich, D., Kopito, R. & Frydman, J. Misfolded proteins partition between two distinct quality control compartments. *Nature* **454,** 1088–1095 (2008).
77. Iwata, A., Riley, B. E., Johnston, J. A. & Kopito, R. R. HDAC6 and microtubules are required for autophagic degradation of aggregated huntingtin. *J. Biol. Chem.* **280,** 40282–40292 (2005).
78. Kopito, R. R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* **10,** 524–530 (2000).
79. Kern, A., Ackermann, B., Clement, A. M., Duerk, H. & Behl, C. HSF1-controlled and age-associated chaperone capacity in neurons and muscle cells of *C. elegans*. *PLoS ONE* **5,** e8568 (2010).
80. David, D. C. *et al.* Widespread protein aggregation as an inherent part of aging in *C. elegans*. *PLoS Biol.* **8,** e1000450 (2010).
81. Demontis, F. & Perrimon, N. FOXO/4E-BP signaling in *Drosophila* muscles regulates organism-wide proteostasis during aging. *Cell* **143,** 813–825 (2010).
82. Morley, J. F. & Morimoto, R. I. Regulation of longevity in *Caenorhabditis elegans* by heat shock factor and molecular chaperones. *Mol. Biol. Cell* **15,** 657–664 (2004).
This pioneering study provides important insight into the relationship between molecular chaperone functions and longevity.
83. Cohen, E., Bieschke, J., Perciavalle, R. M., Kelly, J. W. & Dillin, A. Opposing activities protect against age-onset proteotoxicity. *Science* **313,** 1604–1610 (2006).
An exciting study demonstrating that active disaggregation and the forced formation of large inclusions prevent the accumulation of toxic aggregate species in *C. elegans*.
84. Ben-Zvi, A., Miller, E. A. & Morimoto, R. I. Collapse of proteostasis represents an early molecular event in *Caenorhabditis elegans* aging. *Proc. Natl Acad. Sci. USA* **106,** 14914–14919 (2009).
85. Cohen, E. *et al.* Reduced IGF-1 signaling delays age-associated proteotoxicity in mice. *Cell* **139,** 1157–1169 (2009).
86. Olzscha, H. *et al.* Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* **144,** 67–78 (2011).
This paper demonstrates the existence of a metastable sub-proteome that is at risk of co-aggregating with amyloid-forming disease proteins.
87. Xu, J. *et al.* Gain of function of mutant p53 by coaggregation with multiple tumor suppressors. *Nature Chem. Biol.* **7,** 285–295 (2011).
This interesting study expands the range of diseases promoted by proteostasis deficiency to cancer.
88. Bence, N. F., Sampat, R. M. & Kopito, R. R. Impairment of the ubiquitin–proteasome system by protein aggregation. *Science* **292,** 1552–1555 (2001).
This key paper demonstrates that protein aggregation can interfere with protein degradation.
89. Gidalevitz, T., Ben-Zvi, A., Ho, K. H., Brignull, H. R. & Morimoto, R. I. Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* **311,** 1471–1474 (2006).
90. Schaffar, G. *et al.* Cellular toxicity of polyglutamine expansion proteins: mechanism of transcription factor deactivation. *Mol. Cell* **15,** 95–105 (2004).
91. Lotz, G. P. *et al.* Hsp70 and Hsp40 functionally interact with soluble mutant huntingtin oligomers in a classic ATP-dependent reaction cycle. *J. Biol. Chem.* **285,** 38183–38193 (2010).
92. Sittler, A. *et al.* Geldanamycin activates a heat shock response and inhibits huntingtin aggregation in a cell culture model of Huntington's disease. *Hum. Mol. Genet.* **10,** 1307–1315 (2001).
93. Mu, T. W. *et al.* Chemical and biological approaches synergize to ameliorate protein-folding diseases. *Cell* **134,** 769–781 (2008).
94. Lee, B.-H. *et al.* Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* **467,** 179–184 (2010).
This study describes the first drug-like molecule that can activate proteasome function, thus providing a means to enhance the clearance of aberrantly folded proteins.
95. Jahn, T. R. & Radford, S. E. The Yin and Yang of protein folding. *FEBS J.* **272,** 5962–5970 (2005).
96. Vabulas, R. M., Raychaudhuri, S., Hayer-Hartl, M. & Hartl, F. U. Protein folding in the cytoplasm and the heat shock response. *Cold Spring Harb. Perspect. Biol.* **2,** a004390 (2010).
97. Ryan, M. T. & Hoogenraad, N. J. Mitochondrial-nuclear communications. *Annu. Rev. Biochem.* **76,** 701–722 (2007).
98. Haynes, C. M. & Ron, D. The mitochondrial UPR — protecting organelle protein homeostasis. *J. Cell Sci.* **123,** 3849–3855 (2010).
99. Kenyon, C. J. The genetics of ageing. *Nature* **464,** 504–512 (2010).
100. Westerheide, S. D., Anckar, J., Stevens, S. M., Jr, Sistonen, L. & Morimoto, R. I. Stress-inducible regulation of heat shock factor 1 by the deacetylase SIRT1. *Science* **323,** 1063–1066 (2009).

Acknowledgements We thank W. Balch, A. Dillin, J. Kelly, R. Morimoto and P. Reinhart for discussions about proteostasis, and thank the members of our laboratory for comments on the manuscript. We apologize to all those whose important work could not be cited owing to space limitations.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to F.U.H. (uhartl@biochem.mpg.de).

Molecular Chaperone Functions in Protein Folding and Proteostasis

Yujin E. Kim,* Mark S. Hipp,* Andreas Bracher, Manajit Hayer-Hartl, and F. Ulrich Hartl

Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany; email: uhartl@biochem.mpg.de

Annu. Rev. Biochem. 2013. 82:323–55

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

This article's doi: 10.1146/annurev-biochem-060208-092442

*These authors contributed equally.

Keywords

chaperonin, degradation, Hsc70/Hsp70, Hsp90, protein folding, trigger factor

Abstract

The biological functions of proteins are governed by their three-dimensional fold. Protein folding, maintenance of proteome integrity, and protein homeostasis (proteostasis) critically depend on a complex network of molecular chaperones. Disruption of proteostasis is implicated in aging and the pathogenesis of numerous degenerative diseases. In the cytosol, different classes of molecular chaperones cooperate in evolutionarily conserved folding pathways. Nascent polypeptides interact cotranslationally with a first set of chaperones, including trigger factor and the Hsp70 system, which prevent premature (mis)folding. Folding occurs upon controlled release of newly synthesized proteins from these factors or after transfer to downstream chaperones such as the chaperonins. Chaperonins are large, cylindrical complexes that provide a central compartment for a single protein chain to fold unimpaired by aggregation. This review focuses on recent advances in understanding the mechanisms of chaperone action in promoting and regulating protein folding and on the pathological consequences of protein misfolding and aggregation.

Contents

INTRODUCTION 324
Protein Folding and Aggregation ... 324
Protein Folding In Vivo 326
THE MOLECULAR CHAPERONE CONCEPT 327
Chaperone-Mediated Folding by Kinetic Partitioning 327
Role in Protein Evolution 328
CHAPERONE PATHWAYS IN THE CYTOSOL 328
Chaperone Action on the Ribosome 328
Chaperones Acting Downstream of the Ribosome 330
CHAPERONE PARADIGMS 331
Trigger Factor 331
The Hsp70 System 332
The Chaperonin System 334
The Hsp90 System 338
PROTEOSTASIS NETWORK 340
Significance of Conformational Maintenance 341
Degradation 341
Proteostasis Collapse in Aging and Disease 343
The Proteostasis Network as a Drug Target 344
OUTLOOK 344

Molecular chaperone: any protein that interacts with and aids in the folding or assembly of another protein without being part of its final structure

Proteostasis: the state of balanced proteome found in healthy cells

Amyloid: disease-associated, fibrillar aggregates with cross-β structure

INTRODUCTION

The successful execution of cellular processes depends on the coordinated interactions of proteins. In humans, an estimated 20,000 to 25,000 different proteins are responsible for most aspects of biological function. Following synthesis on ribosomes as linear sequences of amino acids, the vast majority of proteins must fold into well-defined three-dimensional structures (their native states) to attain functionality. Although some newly translated proteins are able to fold spontaneously, a substantial fraction of proteins are less efficient at folding and vulnerable to misfolding, a problem exacerbated by the highly crowded cellular environment (1). In particular, large proteins with complex structures may expose hydrophobic amino acid residues to the solvent during folding, rendering them susceptible to nonnative (off-pathway) interactions that lead to aggregation (2, 3). To counteract these nonnative interactions, cells have a network of molecular chaperones that assist in de novo folding and maintain preexisting proteins in their native states (4–7). A key role of molecular chaperones is preventing protein aggregation, especially under conditions of cellular stress. Moreover, the molecular chaperone network functions in diverse aspects of protein quality control, including protein unfolding and disaggregation and targeting terminally misfolded proteins for proteolytic degradation.

Imbalances in protein homeostasis (proteostasis) are observed in an increasing number of disease states, emphasizing the importance of cellular protein quality control (8). The predominant feature of these disorders is protein misfolding as manifested by the formation of intracellular and/or extracellular deposits of aggregated proteins. Examples include the formation of intracellular inclusions containing aggregated α-synuclein in Parkinson's disease or huntingtin in Huntington's disease, as well as the extracellular β-amyloid plaques in Alzheimer's disease (9). Deficiencies in proteostasis are also observed in many other age-related diseases, such as type II diabetes, peripheral amyloidosis, cancer, and cardiovascular diseases. Indeed, studies using model organisms demonstrate that a gradual decline in cellular proteostasis capacity occurs with aging (10).

Here, we review recent advances in understanding the role of molecular chaperones in protein folding and proteostasis maintenance. We focus our discussion on the cytosolic chaperone networks and the pathological consequences of their disruption. For a discussion of protein folding in the secretory pathway and the folding of membrane proteins, we refer the reader to recent reviews (11, 12, 13).

Protein Folding and Aggregation

The folded three-dimensional structures of most proteins represent a compromise between

thermodynamic stability and the conformational flexibility required for function. Consequently, proteins are often marginally stable in their physiological environment and thus susceptible to misfolding and aggregation (2, 3). In addition, a substantial fraction of proteins in eukaryotic cells (~30%) are classified as intrinsically unstructured and contain regions thought to adopt ordered structure only upon interaction with binding partners (14). Such proteins may be metastable and prone to aggregation.

The pioneering studies of Anfinsen revealed that small denatured proteins refold spontaneously in vitro, demonstrating that the three-dimensional structure of a protein is encoded in its amino acid sequence (2, 3). Much progress in recent years has helped us understand how exactly the linear sequence of amino acids encodes the native state of a protein and directs its folding process. Because the number of possible conformations a protein chain can adopt is very large, folding reactions are highly complex and heterogeneous, relying on the cooperation of multiple weak, noncovalent interactions. Among these, hydrophobic forces are critical in driving chain collapse and the burial of nonpolar amino acid residues within the interior of the folding protein. Polypeptide chains are thought to explore funnel-shaped potential energy surfaces as they progress toward the native structure along several downhill paths rather than a single defined pathway (**Figure 1**) (2). Rapid chain collapse and the incremental formation of native contacts limit the conformational space that must be searched en route to the native state. However, the rugged free-energy surface navigated during folding often requires molecules to cross substantial kinetic barriers. As a result, kinetically trapped folding intermediates and misfolded states may be transiently populated. Misfolded states are characterized by the presence of nonnative interactions that must be resolved prior to correct folding. Productive folding intermediates may display a high degree of configurational flexibility, increasing the search time required for the formation of native intrachain contacts. The propensity of proteins to

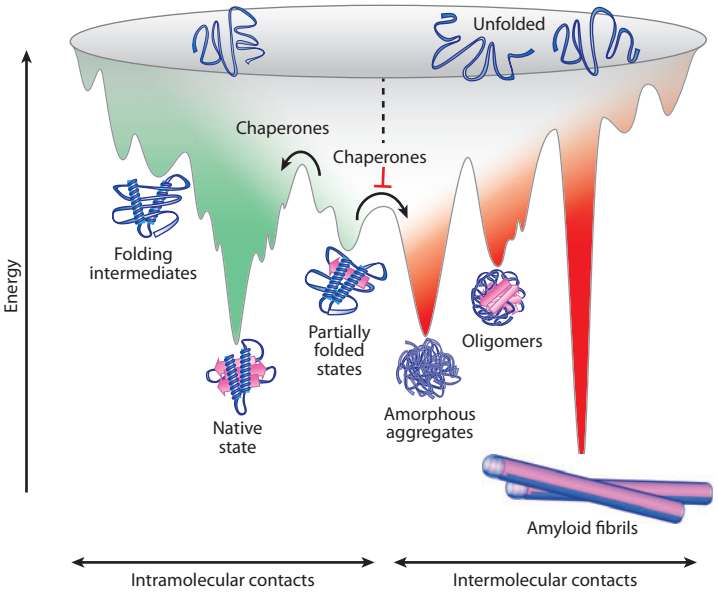


Figure 1

Competing reactions of protein folding and aggregation. Proteins fold by sampling various conformations in a folding energy landscape. Energetically favorable intramolecular interactions (*green*) are stabilized as the protein progresses on a downhill path through the landscape toward the native state. Energetically favorable but nonnative conformations result in populations of kinetically trapped states that occupy low-energy wells (partially folded states or misfolded states). Chaperones assist these states in overcoming free energy barriers and prevent intermolecular interactions (*red*) leading to aggregation (amorphous aggregates, β-sheet-rich oligomers, and amyloid fibrils), thus promoting folding to the native state. Figure adapted and modified from References 17 and 237.

populate such entropically stabilized intermediates increases with larger, more topologically complex domain folds that are stabilized by many long-range interactions (such as α/β domain architectures). Such proteins are often highly chaperone dependent (15).

Partially folded or misfolded states typically expose hydrophobic amino acid residues and regions of unstructured polypeptide backbone to the solvent, the features that mediate aggregation in a concentration-dependent manner (**Figure 1**) (9, 16). Although aggregation primarily leads to amorphous structures largely driven by hydrophobic forces, it may also lead to the formation of amyloid-like fibrils (**Figure 1**). These fibrillar aggregates are characterized by β-strands that run perpendicular to the long

Protein folding: the process by which the extended, newly synthesized polypeptide chain collapses into its functional three-dimensional conformation

fibril axis (cross-β structure) and are associated with diseases of aberrant protein folding. Many proteins can adopt these highly ordered, thermodynamically stable structures in vitro, but molecular chaperones restrict their formation in vivo (9, 16, 17). The formation of soluble oligomeric intermediates often accompanies the formation of fibrillar aggregates (**Figure 1**). The pronounced toxicity of these less-ordered and rather heterogeneous forms likely correlates with the exposure of interaction-prone hydrophobic surfaces, which in the fibril are integrated into a compact cross-β core (18). We can view certain oligomers as kinetically trapped intermediates that must undergo considerable structural rearrangement to form fibrils, the thermodynamic end state of aggregation (**Figure 1**). The relative impact of fibrillar and oligomeric aggregates on proteostasis and cell health is currently under intense investigation.

Protein Folding In Vivo

Attempts to understand protein folding in vivo must take into consideration the dramatic differences in physical properties that exist between the cellular environment and the conditions of test-tube refolding (19, 20). Compared with the dilute conditions in vitro, the cellular environment is highly crowded, containing concentrations of 300–400 mg ml^{-1} of protein and other macromolecules (1, 21). Macromolecular crowding results in excluded volume effects, limiting the entropic freedom of the folding polypeptide chains and favoring compact nonnative states (1, 22). In addition, macromolecular crowding enhances protein aggregation (amorphous and fibrillar) by increasing the affinities between interacting macromolecules including folding intermediates (1).

Protein folding in vivo is further complicated by its coupling with translation and by the fact that many newly synthesized polypeptides must be transported into subcellular compartments, such as the endoplasmic reticulum (ER) or the mitochondria (23), prior to folding. The vectorial translation process from the N terminus to C terminus places considerable restrictions on the energy landscape of in vivo folding (24, 25). The exit tunnel of the large ribosomal subunit, ~100 Å long and ~20 Å wide, precludes folding beyond the formation of α-helices or small tertiary structural elements that may begin to form near the tunnel exit (26–29). Thus, the C-terminal 30–40 amino acid residues of the nascent chain cannot participate in the long-range interactions necessary for cooperative domain folding. Consequently, productive folding is delayed until a complete protein domain (~50–300 amino acid residues), or substantial segments thereof, has emerged from the ribosome (30–35). Whereas single-domain proteins complete folding posttranslationally (after chain termination and release from the ribosome), proteins consisting of multiple domains may fold cotranslationally as the domains emerge sequentially from the ribosome. Sequential folding of the domains prevents the formation of unproductive intermediates resulting from nonnative interactions between concomitantly folding domains (36, 37). For the multidomain protein firefly luciferase, sequential domain folding results in a dramatic acceleration of folding speed (6, 7, 36). The slower translation speed in eukaryotes (~4 amino acids s^{-1}) compared with bacteria (~20 amino acids s^{-1}) (38), together with evolutionary adaptations of the chaperone machinery, may facilitate cotranslational folding for domains with slower folding kinetics and thus may have contributed to the explosive evolution of multidomain proteins in eukaryotes. Although domain size was conserved during evolution, the average size of proteins increased from ~35 kDa in bacteria such as *Escherichia coli* to ~52 kDa in humans (36). Translational pausing may also enhance the efficiency of cotranslation folding, but the significance of this phenomenon in vivo is still under investigation (34, 39–42).

The fastest translation speeds are slow compared with the rapid kinetics of folding observed for small protein domains in vitro, some of which fold on the microsecond to millisecond timescale (2, 3). A nascent chain

of average length (~300 amino acid residues in *E. coli*) will be exposed on the ribosome in an unfolded state for ~15 s, during which it may engage in nonnative intra- and interchain interactions. Contrary to the previous belief that polysomes enhance aggregation by increasing the local concentration of nascent chains, recent studies suggest that the three-dimensional organization of individual ribosomes in polysomes maximizes the distance between nascent chains, thus reducing the probability of unproductive interactions (43).

THE MOLECULAR CHAPERONE CONCEPT

We define a molecular chaperone as any protein that interacts with and aids in the folding or assembly of another protein without being part of its final structure (4). Chaperones are classified into different groups on the basis of sequence homology. Many are stress proteins or heat shock proteins (Hsps), as their synthesis is induced under conditions of stress (e.g., heat shock or oxidative stress), which structurally destabilize a subset of cellular proteins. Members of the various groups of chaperones were initially named according to their molecular weight: Hsp40s, Hsp60s, Hsp70s, Hsp90s, Hsp100s, and the small Hsps. Besides their fundamental role in de novo protein folding, chaperones are involved in various aspects of proteome maintenance, including assistance in macromolecular complex assembly, protein transport and degradation, and aggregate dissociation and refolding of stress-denatured proteins.

Chaperone-Mediated Folding by Kinetic Partitioning

Chaperones that function broadly in de novo folding and refolding (i.e., the chaperonins, Hsp70s, and Hsp90s) are ATP regulated and recognize segments of exposed hydrophobic amino acid residues, which are later buried in the interior of the natively folded protein. Binding to hydrophobic segments enables these chaperones to recognize the nonnative states of many different proteins. Folding is then promoted during ATP- and cochaperone-regulated cycles of binding and release of nonnative protein (**Figure 2**). In this mechanism of kinetic partitioning, (re)binding to chaperones blocks aggregation and reduces

Assembly: the association of two or more protein molecules in a functional complex

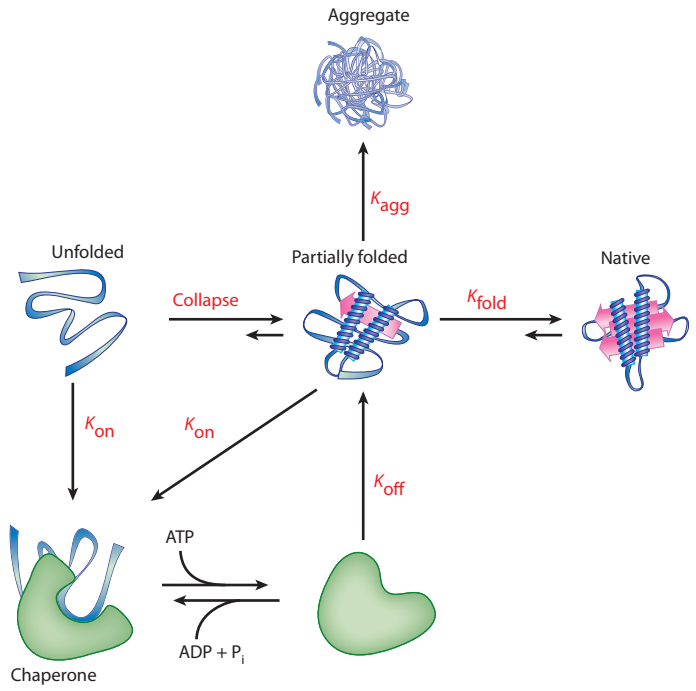


Figure 2

Molecular chaperones promote protein folding through kinetic partitioning. Under physiological conditions, the unfolded state of a protein undergoes rapid collapse to a partially folded, compact intermediate. Although chain collapse restricts the conformational space that must be searched en route to the native state, the collapsed folding intermediates are often aggregation-prone, kinetically trapped states. Many molecular chaperones switch between high- and low-affinity states for unfolded and partially folded proteins in a manner regulated by ATP binding and hydrolysis. Folding proceeds efficiently when the folding rate constant (K_{fold}) is greater than the rate constants for chaperone binding (K_{on}) and aggregation (K_{agg}). Binding or rebinding of nonnative protein to chaperone allows kinetic partitioning, preventing aggregation and favoring folding. When K_{on} is greater than K_{fold}, the chaperone system functions as a holdase, stabilizing the protein in a nonaggregated state for transfer to other chaperone systems or for degradation. During exposure to conformational stress, K_{agg} becomes greater than K_{on}, resulting in aggregation unless the cell upregulates its chaperone capacity through stress response pathways. K_{off} is the rate constant for protein release from chaperone and P_i is inorganic phosphate. Figure adapted and modified from Reference 38.

Stress proteins: cellular proteins whose expression is induced or increased under conditions of conformational stress; not all molecular chaperones are stress proteins

Aggregate: the association of two or more protein molecules in a nonfunctional state

Chaperonin: a class of structurally related molecular chaperones forming large, double-ring complexes that transiently enclose substrate protein for folding

TF: trigger factor

NAC: nascent-chain-associated complex

Nascent polypeptide: the polypeptide chain emerging from the ribosome during translation

RAC: ribosome-associated complex

the concentration of free folding intermediates, whereas transient release of bound hydrophobic regions is necessary for folding to proceed (**Figure 2**). ATP-independent chaperones, such as the small Hsps, may function as additional holdases that buffer aggregation. Efficient folding is achieved when the rate of folding is faster than the rates of aggregation or chaperone rebinding. If folding is slower, then the protein may be transferred to a chaperone system with different mechanistic properties, as exemplified by the sequential cooperation between Hsp70 chaperones and the cylindrical chaperonins (discussed below). Alternatively, transfer to the degradation machinery may occur. Aggregation occurs if the concentration of folding intermediates exceeds the available chaperone capacity. Such a situation generally results in the induction of the cellular stress response, which increases the abundance of stress-regulated chaperones.

Role in Protein Evolution

The general function of chaperones in assisting protein folding is significant in facilitating the structural evolution of proteins. By maintaining nonnative proteins in a soluble, folding-competent state, chaperones are thought to buffer mutations in proteins that would otherwise preclude their folding, thus broadening the range of mutant proteins subject to Darwinian selection (44, 45). After selection of a mutant protein with favorable functional properties, secondary mutations may improve its folding efficiency and solubility, allowing the protein to become less chaperone dependent and increase in abundance. Interestingly, proteins that depend highly on a specific chaperone system, such as the *E. coli* chaperonin GroEL, are of less than average abundance and often have nonessential functions (15). Conversely, highly abundant proteins with essential functions tend to be less dependent on a specific chaperone and may use multiple chaperone systems to optimize folding yield (46). Notable exceptions include the photosynthetic enzyme ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCO) as well as the cytoskeletal proteins actin and tubulin. Although highly abundant, these proteins are obligate substrates of the chaperonin system for folding (4, 6). Presumably, RuBisCO, actin, and tubulin reside in an evolutionarily trapped state in which further mutations that reduce chaperonin dependence are incompatible with function.

CHAPERONE PATHWAYS IN THE CYTOSOL

The general organization of cytosolic chaperone pathways is highly conserved throughout evolution (**Figure 3**). In all domains of life (bacteria, archaea, and eukarya), ribosome-binding chaperones [e.g., trigger factor (TF), nascent-chain-associated complex (NAC), and specialized Hsp70s] interact first with the nascent polypeptide, followed by a second tier of chaperones without direct affinity for the ribosome (the classical Hsp70 system). Folding may begin cotranslationally and finish posttranslationally upon chain release from the ribosome or after transfer to downstream chaperones (e.g., the chaperonins and Hsp90 system) (**Figure 3**). Recent system-wide and bioinformatic approaches identified the substrate interactome of several major chaperones, revealing the cooperative organization of the chaperone network (15, 47–57).

Chaperone Action on the Ribosome

As discussed above, the nascent polypeptide chain is topologically restricted until a complete protein domain is synthesized and emerges from the ribosomal tunnel. Ribosome-binding chaperones prevent emerging nascent chains from engaging in unfavorable intra- and intermolecular interactions during translation, typically by shielding exposed hydrophobic segments. The ribosome-associated molecular chaperones include TF (in prokaryotes) and specialized Hsp70 complexes such as ribosome-associated complex (RAC; in *Saccharomyces cerevisiae*), MPP11 and Hsp70L1

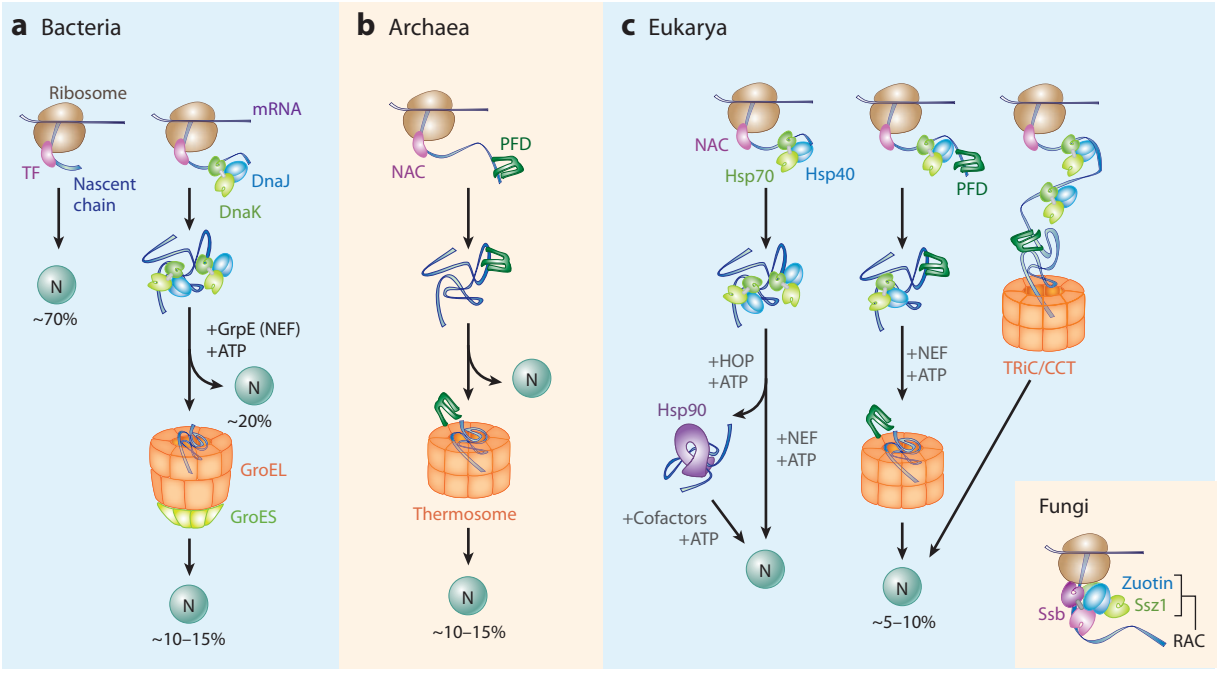


Figure 3

Organization of chaperone pathways in the cytosol. In bacteria (*a*), archaea (*b*), and eukarya (*c*), ribosome-bound chaperones [trigger factor (TF) in bacteria, nascent-chain-associated complex (NAC) in archaea and eukarya] aid folding cotranslationally by binding to hydrophobic segments on the emerging nascent chains. For longer nascent chains, members of the heat shock protein (Hsp)70 family (DnaK in bacteria and Hsp70 in eukarya), together with Hsp40s and nucleotide exchange factors (NEFs), mediate co- and posttranslational folding. In archaea lacking the Hsp70 system, prefoldin (PFD) assists in folding downstream of NAC. Partially folded substrates may be transferred to the chaperonins [GroEL-GroES in bacteria, thermosome in archaea, and tailless complex polypeptide-1 (TCP-1) ring complex (TRiC)/chaperonin-containing TCP-1 (CCT) in eukarya]. The Hsp90 system also receives its substrates from heat shock cognate 70 (Hsc70) and mediates their folding with additional cofactors. The insert in panel *c* shows the ribosome-binding chaperone system, the ribosome-associated complex (RAC), in fungi. RAC consists of Ssz1 (a specialized Hsp70) and zuotin (Hsp40) and assists nascent chain folding together with another Hsp70 isoform, Ssb. Percentages indicate the approximate protein flux through the various chaperones. Figure adapted and modified from Reference 7.

(in mammals), and NAC (in archaea and eukaryotes) (**Figure 3**) (17, 58, 59).

TF binds to the large ribosomal subunit at the opening of the ribosomal exit tunnel (60–62) and interacts with most newly synthesized cytosolic proteins as well as a subset of secretory proteins (55, 56, 58, 59). In vitro, TF binds to nascent chains as short as ~60 amino acid residues, presumably when the first hydrophobic segment of the chain has emerged (58, 63). In vivo, TF preferentially binds ribosomes when nascent chains have reached ~100 amino acids in length (55), thereby permitting prior interactions of ribosome-binding targeting factors (e.g., signal recognition particle) (64, 65) and modifying enzymes (e.g., peptide deformylase) (66) with the nascent chains. Release of TF from the nascent chain is ATP independent and permits folding or polypeptide transfer to downstream chaperones such as DnaK, the major Hsp70 in bacteria. Although the combined deletion of the genes encoding TF and DnaK is lethal at temperatures above 30°C, *E. coli* cells tolerate individual deletions of TF or DnaK, indicating that these proteins are functionally redundant (56, 58, 67).

In eukaryotes, RAC and NAC may fulfill a role similar to TF, although they are not structurally related to TF. RAC comprises the Hsp70-like protein Ssz1 (also known as Pdr13)

NEF: nucleotide exchange factor

TRiC: tailless complex polypeptide-1 ring complex

and the Hsp70 cochaperone zuotin (Hsp40) (**Figure 3*c***) (58, 59, 68–71). In fungi, RAC cooperates with ribosome-binding isoforms of Hsp70, Ssb1, or Ssb2. NAC, a heterodimeric complex of α- (33 kDa) and β- (22 kDa) subunits, also associates with ribosomes (via the β-subunit) and short nascent chains (54, 59, 72, 73). The exact role of NAC in folding or protein quality control is not established. In yeast, NAC function appears to be partially redundant with that of Ssb (71), reminiscent of the interplay between TF and DnaK in bacteria. Both Ssb-RAC and NAC may have a role in ribosome biogenesis (71). Archaea have a homolog of α-NAC.

Chaperones Acting Downstream of the Ribosome

In bacteria and eukaryotic cells, the classical Hsp70s have a central role in the cytosolic chaperone network (6, 17). They interact with a multitude of nascent and newly synthesized polypeptides but have no direct affinity for the ribosome (56, 74). The nascent-chain-interacting Hsp70 chaperones include DnaK in bacteria and some archaea, Ssa1–4 in yeast, and the constitutively expressed heat shock cognate 70 (Hsc70) in metazoan and mammalian cells (4, 58). Hsp70 chaperones function with cochaperones of the Hsp40 family (also known as DnaJ proteins or J proteins) and nucleotide exchange factors (NEFs) (**Figure 3**) (75). In addition to protecting nascent chains against aberrant interactions, the Hsp70-Hsp40 chaperone systems assist folding co- or post-translationally through ATP-regulated cycles of substrate binding and release (by kinetic partitioning; **Figure 2**) and mediate polypeptide chain transfer to downstream chaperones. Remarkably, most species of archaea lack the Hsp70 chaperone system. The chaperone prefoldin (also known as the Gim complex, GimC) may substitute for Hsp70 in these cases (7). Prefoldin, a hexamer of ~14–23 kDa α- and β-subunits with long α-helical coiled coils, binds in an ATP-independent manner to certain nascent chains and mediates their transfer to the cylindrical chaperonin complex for the final stages of folding. In eukaryotes, prefoldin participates in the chaperonin-assisted folding of actin and tubulin (6, 7).

Proteins that are unable to utilize Hsp70 for folding are transferred to the chaperonin or the Hsp90 system. Chaperonins (also referred to as Hsp60s) are large double-ring complexes of 800–1,000 kDa with a central cavity, which permits protein molecules to fold in an isolated compartment protected from the aggregation-promoting cytosol (**Figure 3**) (4, 5). The chaperonins are structurally classified into group I and group II (76, 77). Group I chaperonins include GroEL in bacteria, Hsp60 in mitochondria, and Cpn60 in chloroplasts. They cooperate with lid-shaped cochaperones (GroES, Hsp10, Cpn10/20) to encapsulate substrates. The group II chaperonins include the archaeal thermosome and its eukaryotic homolog tailless complex polypeptide-1 (TCP-1) ring complex (TRiC), also known as chaperonin-containing TCP-1 (CCT), which have a built-in lid. GroEL-GroES assist folding posttranslationally, whereas TRiC may assist folding co- and posttranslationally (6). TRiC binds to nascent chains and cooperates with Hsp70 in the cotranslational folding of multidomain proteins (78). Investigators have demonstrated a direct interaction between Hsp70 and TRiC (79). The chaperonins interact with 10–15% of newly synthesized polypeptides in bacteria and archaea (15, 52, 53) and 5–10% in eukarya (50). The obligate substrates of GroEL typically include proteins with complex domain folds, which tend to populate kinetically trapped folding intermediates (15, 53). Established TRiC substrates include the cytoskeletal proteins actin and tubulin as well as several proteins with β-propellers/WD40 repeats (49, 50).

In the eukaryotes, many signaling proteins are transferred from Hsp70 to the ATP-dependent Hsp90 chaperone system for completion of folding and conformational regulation (**Figure 3*c***) (45, 48, 80). Examples of Hsp90-mediated conformational regulation include nuclear hormone receptors, which Hsp90 stabilizes in a conformation poised for hormone binding and activation. Substrate

transfer to Hsp90 is mediated by the Hsp90 organizing protein (HOP; also known as Sti1 and p60), which uses multiple tetratricopeptide repeat (TPR) domains to bridge Hsp70 and Hsp90 (81). Various cochaperones facilitate the interaction of Hsp90 with steroid hormone receptors and multiple protein kinases (45, 48). Accordingly, Hsp90 affects many key cellular processes, including cell cycle progression, steroid signaling, calcium signaling, protein trafficking, protein secretion, the immune response, and the heat shock response (HSR) (45, 48, 82). Pharmacologic inhibition of Hsp90 with geldanamycin and derivatives results in the downregulation of many kinases (83) and is in clinical development for cancer therapy (84).

The molecular chaperone network is central to cellular protein quality control through its involvement in the conformational maintenance of proteins, the dissociation of aggregates, and the degradation of misfolded proteins. In yeast and other fungi, the cytosolic Hsp70 system cooperates with the AAA+ (ATPases associated with various cellular activities) chaperone Hsp104 in dissociating and refolding aggregated proteins. Hsp104 is homologous to bacterial ClpB, which functions with DnaK in protein disaggregation (85). Various cochaperones of Hsp70 and Hsp90 escort terminally misfolded proteins to the protein degradation machinery (ubiquitin-proteasome system or autophagy) (86).

CHAPERONE PARADIGMS

Research has defined several mechanistic paradigms of chaperone function in protein folding. In the following sections we discuss TF as an ATP-independent chaperone as well as the Hsp70 system, the chaperonins, and Hsp90 as ATP-dependent paradigms. Structural and functional data strongly support the mechanistic models of each of these systems.

Trigger Factor

Bacterial TF is an abundant ~50 kDa protein that binds to ribosomes and interacts with most nascent polypeptides (7, 55, 56, 59, 63, 87, 88). The crystal structure of *E. coli* TF reveals an elongated structure consisting of three domains, an N-terminal ribosome-binding domain, a peptidylprolyl *cis/trans* isomerase (PPIase) domain, and a C-terminal domain (positioned between the N-terminal and PPIase domains) (**Figure 4*a***) (61). The N-terminal

PPIase: peptidylprolyl *cis/trans* isomerase

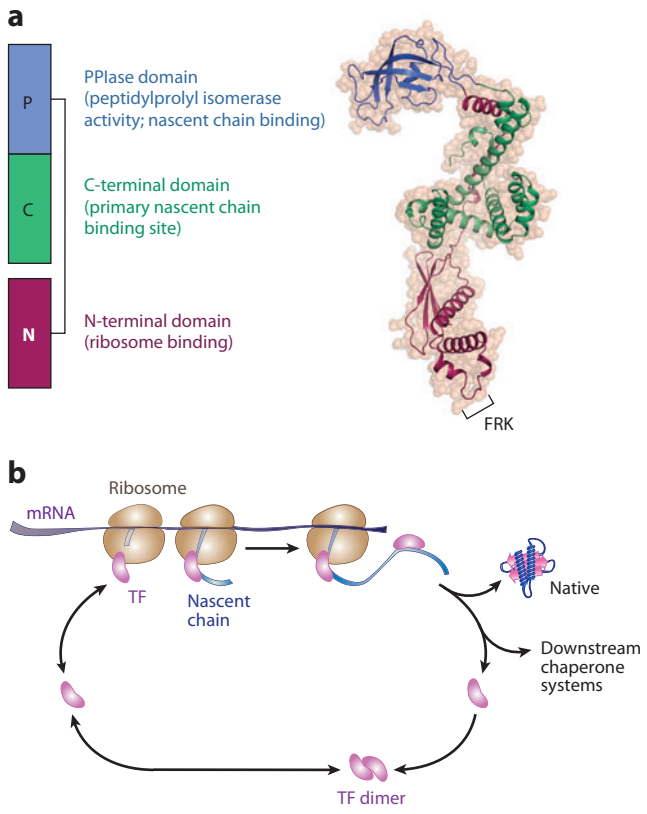


Figure 4

Structure and functional cycle of trigger factor (TF). (*a*) Structure of TF. The N-terminal domain of TF (*magenta*) contains the Phe-Arg-Lys (FRK) ribosome-binding loop and connects by a long linker to the peptidylprolyl *cis/trans* isomerase (PPIase) domain (*blue*). The C-terminal domain (*green*) is in the center of the protein and provides the main binding site for the nascent chain substrate [Protein Data Bank (PDB) 1W26]. (*b*) TF reaction cycle. Free TF exists in the cytosol as a dimer in rapid equilibrium with monomers. Monomeric TF binds to ribosomes with translating nascent chains. Hydrophobic sequence motifs in the nascent chain regulate TF-nascent chain interaction. Binding to the elongating nascent chain may persist as TF dissociates from the ribosome (with $t_{1/2} \sim 10$ s) (63), allowing a second TF molecule to bind to the ribosome at the polypeptide exit site. Concurrent with release from TF, the newly synthesized polypeptide folds into its native state or is transferred to downstream chaperones.

NBD: nucleotide-binding domain

SBD: substrate-binding domain

domain mediates binding of monomeric TF to the large ribosomal subunit at proteins L23 and L29 in close proximity to the polypeptide exit site (60, 61, 89, 90). L23 is essential for TF binding and is thought to signal the progression of the nascent chain through the ribosomal exit tunnel (89), whereas L29 has an auxiliary function (60). The PPIase domain connects to the N domain via a long linker and is most distal to the ribosome docking site. It belongs to the family of FK506 binding proteins and recognizes stretches of eight amino acids that are enriched in basic and aromatic residues (7). Studies have measured PPIase activity in vitro, but its contribution to folding in vivo has remained unclear, as the domain is dispensable for TF function (59, 67, 91, 92). Cross-linking experiments showed that the PPIase domain interacts with longer nascent chains, presumably representing an auxiliary substrate recognition site (62, 87). The centrally located C-terminal domain is structurally similar to the N-terminal domain of the periplasmic chaperone SurA and provides the primary binding site for the nascent chain (87, 93).

Cells contain an excess of TF compared with ribosomes; the non-ribosome-bound fraction is in rapid monomer-dimer equilibrium (**Figure 4*b***) (63, 94). TF monomer aids de novo folding through ATP-independent cycles of binding and release from both the ribosome and the nascent chain. Binding to the ribosome (mean residence time of ~10–15 s) is a prerequisite for interaction with the nascent chain. The disposition of the bound peptide to bury hydrophobic regions during folding likely drives the eventual release of TF from the nascent chain. Accordingly, TF slows hydrophobic chain collapse and delays cotranslational folding (37, 95, 96). Furthermore TF may remain bound to certain polypeptides after their release from the ribosome, which is consistent with a role for TF as a holdase in ribosome assembly (97, 98). Several studies indicate that TF retains a high degree of conformational flexibility on the ribosome during interactions with the emerging nascent chain (62, 63, 89, 90). This flexibility likely enables TF to accommodate a wide range of polypeptides (63, 88). Although it primarily binds hydrophobic chain segments (63, 88), TF also interacts with small basic proteins, including many ribosomal proteins (56, 88, 98).

The Hsp70 System

Hsp70 chaperones are a ubiquitous class of proteins. They are involved in a wide range of protein quality control functions, including de novo protein folding, refolding of stress-denatured proteins, protein transport, membrane translocation, and protein degradation.

Structure and reaction cycle. Hsp70 consists of an N-terminal nucleotide-binding domain (NBD) and a C-terminal substrate-binding domain (SBD) connected by a highly conserved hydrophobic linker region (**Figure 5*a***). The N-terminal domain has an actin-like fold; it consists of two lobes, each containing two subdomains, with the nucleotide-binding cleft situated in between (5, 99). The SBD consists of a β-sandwich subdomain and an α-helical lid with the substrate binding site located in the β-sandwich subdomain (5). The SBD binds to 5–7-residue peptide segments enriched in hydrophobic amino acids and typically flanked by positively charged residues. The peptide binds in an extended conformation mediated by hydrogen bonds between the SBD and the peptide backbone and by van der Waals contacts with the hydrophobic side chains (5, 100).

Conformational changes in the NBD upon ATP binding and hydrolysis are allosterically coupled to the SBD, regulating peptide binding and release. Binding of ATP to the NBD triggers the attachment of the hydrophobic interdomain linker and the α-helical lid of the SBD to the NBD, which opens the peptide binding pocket (101, 111), as initially revealed in structures of the Hsp70 homolog Hsp110 (**Figure 5*a***) (103–105). Hydrolysis of ATP to ADP triggers the detachment of the lid from the NBD and the closing of the SBD; NBD and SBD are loosely held together by the linker in a dynamic random coil conformation (106, 107).

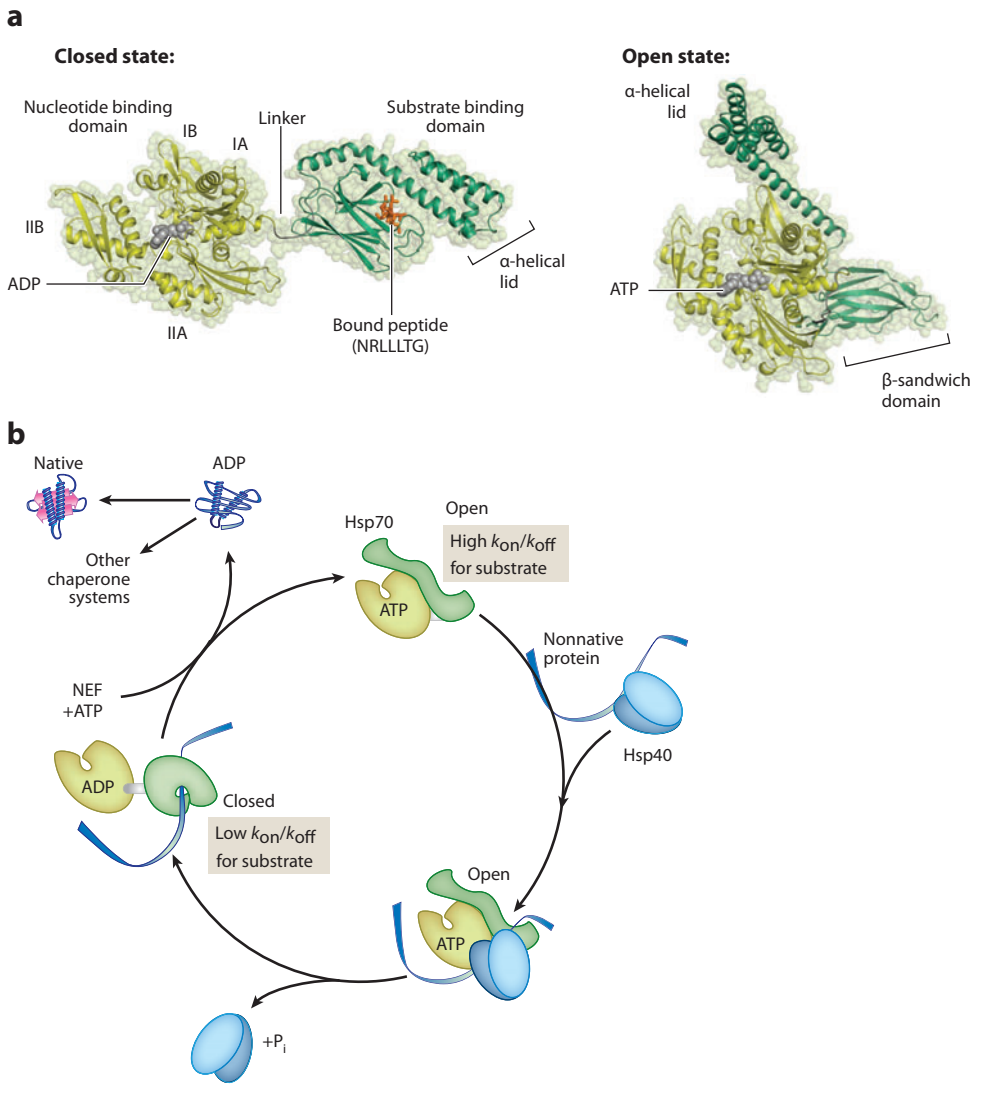


Figure 5

Structure and functional cycle of Hsp70. (*a*) Structure of Hsp70. Hsp70 consists of two domains, the nucleotide-binding domain (NBD) and the substrate-binding domain (SBD), connected by a conserved linker. The closed state of *Escherichia coli* DnaK (*left*; PDB 2KHO) was solved using a combination of solution nuclear magnetic resonance spectroscopy and crystal structures of the individual domains (106). The structure illustrates the ADP-bound NBD separated by a linker from the SBD. The α-helical lid of the SBD is closed over the substrate peptide (NRLLLTG) bound in the pocket of the β-sandwich domain. The open state is illustrated by the crystal structure of ATP-bound Sse1 (yeast Hsp110; *right*; PDB 2QXL). The β-sandwich domain contacts subdomain IA of the NBD, whereas the α-helical lid contacts subdomains IA and IB. (*b*) Reaction cycle. ATP binding to the NBD stabilizes the open state of Hsp70, facilitating the binding of substrate protein recruited to Hsp70 by Hsp40 cochaperone. The open state has fast on and off rates for substrate peptide. Hsp40 stimulates ATP hydrolysis on Hsp70, resulting in the closing of the SBD α-helical lid over the bound substrate peptide. The closed state has slow on and off rates for substrate peptide. NEFs stimulate the release of ADP from the NBD, and ATP binding causes substrate release.

Details of this allosteric mechanism are under intense investigation (108). The ATP-bound open state has high on and off rates for peptide substrate, whereas the ADP-bound closed state has low on and off rates (**Figure 5*b***) (5, 100). In turn, substrate binding increases the rate of ATP hydrolysis (109–111). Substrate release allows folding (i.e., the burial of hydrophobic residues) to proceed. Proteins unable to fold rapidly upon dissociation from Hsp70 may rebind, transfer to downstream chaperones, or be transferred to the degradation machinery.

Cochaperones. Hsp40 (J protein) and NEF cochaperones regulate the Hsp70 reaction cycle (38, 100). The Hsp40 proteins constitute a large family with more than 40 members in humans (75). All of them contain a J domain, which binds to the N-terminal ATPase domain of Hsp70 and the adjacent linker region (112, 113). Canonical Hsp40s (members of classes I and II) function as chaperones independently and recruit Hsp70 to nonnative substrate proteins. Other Hsp40s (class III) are more diverse and combine the J domain with a variety of functional modules (75, 114, 115). The interaction with Hsp70 strongly stimulates (>1,000-fold) the hydrolysis of Hsp70-bound ATP to ADP, resulting in stable substrate binding by Hsp70 in the closed conformation (38, 100). Subsequent binding of an NEF to the NBD of Hsp70 catalyzes the exchange of ADP to ATP, opening the SBD and thereby triggering substrate release (**Figure 5*b***).

Bacteria have one Hsp70 NEF, GrpE, whereas eukaryotic cells contain several structurally unrelated families of NEFs, including the Bcl-2-associated athanogene (BAG) domain proteins, HspBP1 and Hsp110 proteins (116–119). Crystal structures of Hsp70-NEF complexes suggest that all NEFs, except HspBP1, stabilize the Hsp70 NBD in a conformation with an open nucleotide-binding cleft (104, 105, 120–123). The most abundant and ubiquitous eukaryotic NEFs are the Hsp110 proteins, which are structurally related to Hsp70 (103–105). Hsp110 may function as holdases for nonnative proteins and cooperate with Hsp70 and Hsp40 in protein disaggregation (104, 124, 125).

Proper regulation of the ATPase cycle is crucial for efficient Hsp70 function. Hsp40 and NEF proteins are present at lower levels relative to their partner Hsp70. They provide a means of diversifying Hsp70 function and establish substrate specificities for the Hsp70 machinery. For example, the human genome encodes 41 different J-domain proteins compared with 11 Hsp70s and 13 NEFs (75). Interestingly, inactivating mutations in the NEF Sil1 cause the neurodegenerative disease Marinesco-Sjögren syndrome, which highlights the critical importance of the nucleotide exchange function (126, 127).

The Chaperonin System

Chaperonins are large double-ring complexes of 7–9 ~60 kDa subunits per ring. They are unique among molecular chaperones in that they encapsulate their substrate proteins, one molecule at a time, in a central cavity for folding unimpaired by aggregation (4, 5, 17). Two groups of distantly related chaperonins can be distinguished (6, 76, 77).

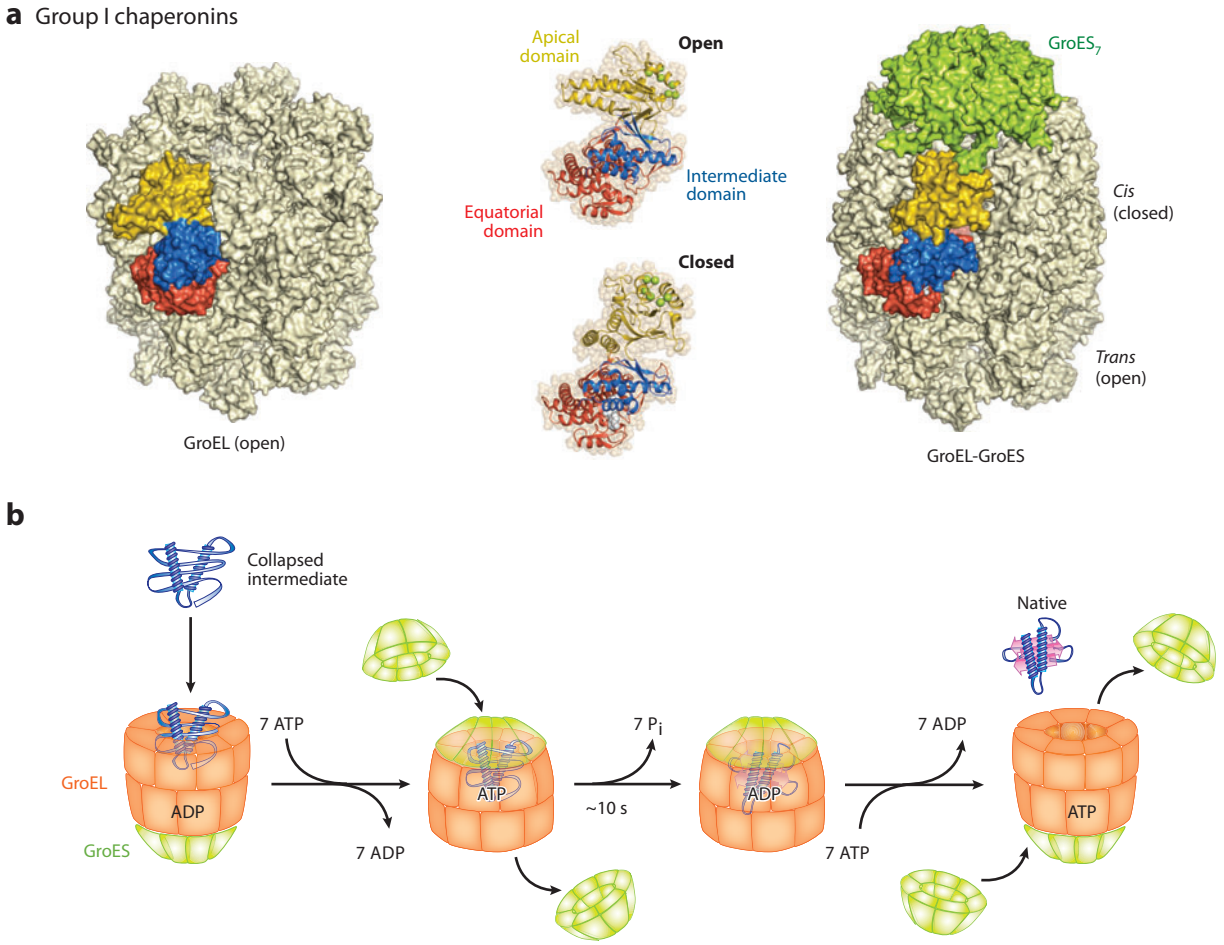
Group I chaperonin GroEL. Group I chaperonins are present in the bacterial cytosol (GroEL), the mitochondrial matrix (Hsp60), and the stroma of chloroplasts (Cpn60). They have seven subunits per ring and are defined by their functional requirement for lid-shaped cochaperones (GroES in bacteria, Hsp10 in mitochondria, and Cpn10/Cpn20 in chloroplasts). The GroEL-GroES system of *E. coli* has been studied most extensively. The subunits of GroEL contain an equatorial ATPase domain, an intermediate hinge domain, and an apical domain (**Figure 6*a***) (77, 100). The apical domains form the entrance to the GroEL cavity and expose hydrophobic amino acid residues, which mediate substrate binding. GroEL-bound substrates typically interact with multiple apical domains (77, 128) and populate an ensemble of compact and locally expanded states that lack stable tertiary interactions, similar to a molten globule (4, 129–131). Binding to GroEL prevents aggregation of these flexible

folding intermediates, and subsequent folding depends critically on the global encapsulation of the substrate in the chaperonin cavity by the cochaperone GroES (7, 132–135). GroES is a heptameric ring of ~10 kDa subunits that binds to the apical GroEL domains, capping the GroEL cylinder (**Figure 6*a*,*b***) (4, 77, 100).

GroEL and GroES undergo a complex binding-and-release cycle that is allosterically regulated by ATP binding and hydrolysis in the GroEL subunits (**Figure 6*b***) (4, 77, 100, 136). Cooperative binding of ATP to GroEL initiates a series of conformational changes that trigger the association of GroES, followed by substrate release from hydrophobic binding sites into a GroES-capped, hydrophilic folding chamber (137). Proteins up to ~60 kDa can be encapsulated and are free to fold in the cage for ~10 s (at 25°C) (132), the time needed for ATP hydrolysis in the GroES-capped ring (*cis* ring). The protein substrate leaves GroEL upon GroES dissociation, which is induced by ATP binding in the opposite ring (*trans* ring) (**Figure 6*b***) (77). Folding intermediates still exposing hydrophobic regions rapidly rebind to GroEL for repeated folding cycles. The exact coordination of the two GroEL rings in the folding cycle is still under investigation (138). Proteins that exceed the size limit of the GroEL-GroES cage may utilize the Hsp70 system for folding (15, 37, 56) or undergo cycles of binding and release from GroEL without GroES encapsulation (77).

In addition to providing an isolated folding environment, other mechanistic elements of the chaperonin cycle contribute to optimizing the rate and yield of the folding process. Repeated events of substrate unfolding in successive binding and release cycles may reverse kinetically trapped states (iterative annealing) (139, 140). Studies have documented unfolding upon binding with Förster resonance energy transfer (FRET) using fluorescence-labeled substrate proteins (130, 140). Additionally, some active unfolding may occur as a result of ATP-dependent movement of the GroEL apical domains (130, 140). However, the significance of iterative annealing is unclear considering that a single round of substrate binding and encapsulation (using a single-ring mutant of GroEL that binds GroES stably) results in substrate protein folding with equal efficiency and kinetics as achieved through multiple cycles of binding and release (7, 132, 141). Whereas partial unfolding upon initial binding may dissociate nonnative interactions in kinetically trapped folding intermediates, the release of protein from the GroEL apical domains may follow a stepwise mechanism in which less tightly bound hydrophobic regions are released first (130). This stepwise release may contribute to avoiding nonnative interactions during protein collapse.

In addition (or as an alternative) to the mechanisms discussed above, growing experimental and theoretical evidence suggests that the GroEL-GroES cage promotes folding by sterically confining folding intermediates (7, 132, 141–145). This model assigns an active role to the chaperonin cage, as opposed to the view that it functions solely as a passive aggregation-prevention device (146). In support of the confinement model, electron microscopy demonstrated that substrates fully occupy the limited volume of the GroEL-GroES cage (134). The resulting constraints on the substrate protein will inevitably affect its folding energy landscape. Indeed, evidence shows that encapsulation in the GroEL-GroES cage accelerates folding up to tenfold compared with the rate of spontaneous folding (measured without interference of aggregation) (7, 132, 141, 147). In addition to steric confinement, mutational analysis demonstrated that the charged residues of the GroEL cavity wall are critical for the observed acceleration (132, 133, 148). According to molecular dynamics simulations, these polar residues accumulate ordered water molecules in their vicinity, thereby generating a local environment in which a substrate protein may bury exposed hydrophobic residues more effectively (144). This would result in the entropic destabilization of flexible folding intermediates in a manner similar to the role of intramolecular disulfide bonds in promoting the folding of secretory proteins. In this context, it is interesting



to note that oxidizing folding compartments supporting disulfide bond formation (e.g., the bacterial periplasm or ER) do not contain chaperonins (141).

Obligate GroEL substrates typically have complex α/β or α + β domain topologies, such as the $(\beta/\alpha)_8$ TIM barrel fold (15, 46, 53), which are stabilized by many long-range interactions. Such proteins must navigate complex folding energy landscapes and overcome entropic barriers to reach their native states (141); hence, they may benefit from steric confinement during folding. Moreover, obligate GroEL substrates often have relatively low sequence hydrophobicity, consistent with inefficient hydrophobic collapse and population of aggregation-prone intermediates that are recognized by GroEL (46, 141, 149). Recent evidence indicates that GroEL-GroES can also accelerate the folding of proteins containing trefoil-knotted structures, a complex topological arrangement that would plausibly form more efficiently in a confining environment (150). Future studies are needed to define precisely what structural properties distinguish GroEL-dependent from GroEL-independent proteins.

Group II chaperonin TRiC/CCT. Although all chaperonins share a common cylindrical architecture, substantial structural differences exist between group I and group II chaperonins (**Figure 6*a*,*c***). For example, group II chaperonins have apical protrusions that function as a built-in lid in place of a separate GroES-like cochaperone. Also, group II chaperonins contain eight or nine subunits per ring, which are stacked directly opposite one another in the two rings (77, 151–158). In contrast, each of the seven subunits of group I chaperonins interdigitates between two subunits of the opposite ring. In many cases, group II chaperonins are hetero-oligomeric, containing up to eight paralogous subunits per ring in a defined order in the case of the eukaryotic chaperonin TRiC (159, 160). Similar to GroEL, group II chaperonins also cycle between open and closed states, and substrate encapsulation is essential for folding (161). But in contrast to GroEL, ATP hydrolysis, not ATP binding, triggers cavity closure, and ATP hydrolysis transition state analogs (e.g., ADP-aluminum fluoride) stabilize the closed state (136, 161, 162). During the transition to the closed state, the apical domains of adjacent subunits undergo pairwise

Figure 6

Structure and function of the chaperonin systems. (*a*) Structure of group I chaperonin. The crystal structures of GroEL (PDB 1SS8) (*left*) and the asymmetrical GroEL (7 ADP)-GroES complex (PDB 1PF9) (*right*) appear with GroES shown in green and one subunit of GroEL colored to indicate its domain structure (equatorial nucleotide-binding domain in *red*; intermediate hinge domain in *blue*; and the apical substrate and GroES-binding domain in *yellow*). The conformational differences between the GroEL subunits in the open state (GroEL and *trans* ring of GroEL-GroES complex) and in the closed state (*cis* ring of GroEL-GroES complex) appear in ribbon representations of single subunits (*middle*). The green spheres represent hydrophobic residues on helices 8 and 9 of the apical domain that are involved in substrate binding in the open conformation and in GroES binding in the closed state. (*b*) GroEL-GroES reaction cycle. Substrate protein as a collapsed folding intermediate is bound by the open GroEL ring of the asymmetrical GroEL-GroES complex, shown in panel *a*. Binding of ATP to each of the seven GroEL subunits causes a conformational change in the apical domains, which results in the exposure of the GroES binding residues, allowing substrate encapsulation in the *cis* complex. ADP and GroES dissociate from the opposite ring (*trans* ring) together with the previously bound substrate. The newly encapsulated substrate is free to fold in the GroEL cavity during the time needed to hydrolyze the seven ATP molecules bound to the *cis* ring (~10 s). ATP binding followed by GroES binding to the *trans* ring triggers GroES dissociation from the *cis* ring, releasing the substrate protein. (*c*) Structure of group II chaperonin. The crystal structures of the open form of the homo-oligomeric *Methanococcus maripaludis* thermosome (PBD 3KFK) (*left*) and the closed form of TRiC/CCT (PDB 4D8Q) (*right*) are shown. The eight paralogous TRiC/CCT subunits appear in different colors, showing both homotypic (*blue* subunits) and heterotypic contacts (*green-beige*, *blue-beige*) between the top and bottom rings. One subunit (CCT3 in the case of TRiC) is colored to indicate its domain structure (equatorial nucleotide-binding domain in *red*; intermediate hinge domain in *green*; and the apical substrate-binding domain in *yellow*). The conformational differences between the open and the closed state, taken from the thermosome crystal structures (PBD 3KFK for the open and PDB 3KFB for the closed state) (*middle*), appear in ribbon representation. In place of a GroES-like cochaperone, the group II chaperonins have an extended apical domain that functions as a built-in lid.

association, resulting in intermediate structures with pseudo-fourfold symmetry (158). Closure is completed when the tips of the helical protrusions form a mixed eight-stranded β-barrel structure around the apical pore.

TRiC interacts with a wide range of cytosolic proteins (49, 50). Prominent substrates include actin and tubulin, which are strictly dependent on the chaperonin for folding (4, 77). A recent crystal structure of bovine TRiC in the open conformation bound to a folding intermediate of tubulin suggests that the substrate initially interacts with loops in the apical and equatorial domains of the TRiC subunits exposed toward the central cavity (157). Helical motifs at the interface between adjacent apical domains have also been implicated in substrate interactions (163). Differences in binding specificities among the different subunits may be important in binding and folding a range of structurally diverse proteins. Accordingly, all eight TRiC subunits are essential in *S. cerevisiae*. Interestingly, the cavity wall exhibits a pronounced segregation of positive and negative surface charges into opposing halves, a feature that might be functionally important (160). The reaction cycle of TRiC is slower than that of GroEL, providing a substantially longer period of protein encapsulation and folding in the cage (164). Also, the iris-like closing mechanism enables the encapsulation of TRiC-dependent domains in the context of large multidomain proteins that cannot be encapsulated in their entirety (164a). Such a mechanism would circumvent the size constraints of chaperonin-assisted folding and may have facilitated the evolution of eukaryotic multidomain proteins with complex architectures. TRiC also interacts with N-terminal fragments of mutant huntingtin that contain an expanded polyglutamine repeat sequence (165–168). Binding to TRiC modulates the aggregation properties of this protein and reduces its cytotoxicity.

The Hsp90 System

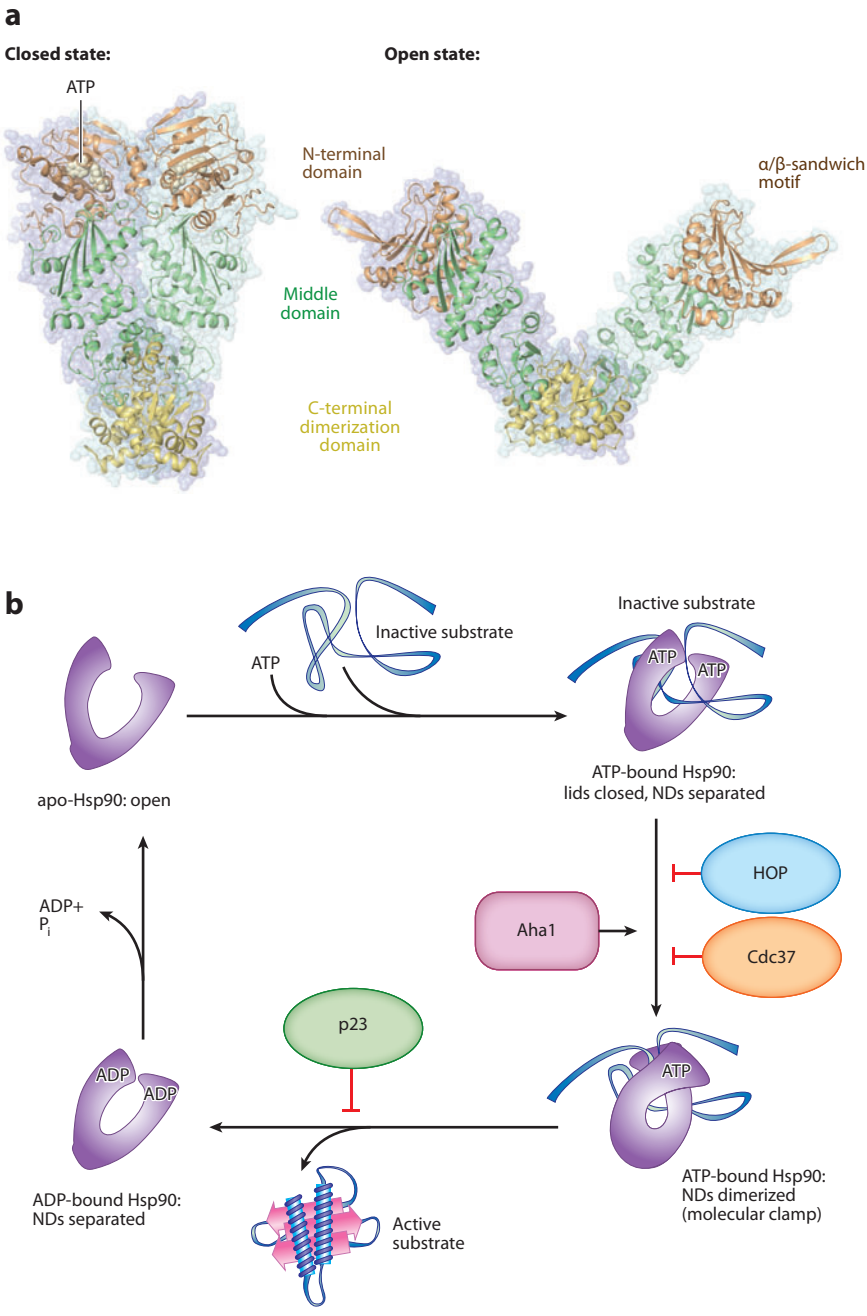
The Hsp90 chaperone system has a central role in cell regulation. Among its substrates are multiple signaling molecules, which are delivered to Hsp90 by Hsp70 chaperones and other cofactors.

Structure and reaction cycle. The Hsp90 chaperones structurally belong to the gyrase, histidine kinase, and MutL superfamily of ATPases (169). Crystal structures of homodimeric Hsp90 molecules from bacteria (170), yeast (171), and mammals (172) illustrate both open and closed functional states of the chaperone (**Figure 7*a***). In addition, complete or partial structures of Hsp90 in complex with different cochaperones or small molecule inhibitors were obtained by X-ray crystallography and cyroelectron microscopy (81, 169, 171, 173–176). Collectively, these structures reveal a high degree of flexibility in Hsp90 conformations (177), consistent with the diversity of Hsp90 client proteins (45).

Figure 7

Structure and functional cycle of the Hsp90 system. (*a*) Structure of Hsp90. Crystal structures of the Hsp90 dimer in the ATP-bound closed state (*Saccharomyces cerevisiae*; PDB 2CG9) (*left*) and the nucleotide-free open state (*E. coli*; PDB 2IOQ) (*right*) are shown with the nucleotide-binding N-terminal domain in orange, middle domain in green, and C-terminal domain in yellow. (*b*) Hsp90 reaction cycle. Inactive substrate protein binds to ATP-bound Hsp90. In this state the ATP lids are closed and the N-terminal domains are separated. In the next step, the N-terminal domains dimerize, forming the closed Hsp90 dimer (referred to as a molecular clamp) with twisted subunits. This metastable state is committed to ATP hydrolysis, upon which the N-terminal domains dissociate. The bound substrate protein is conformationally activated as Hsp90 proceeds through the cycle. Cofactors such as Cdc37 and the Hsp90 organizing protein (HOP) slow the ATP hydrolysis step of the cycle, whereas the activator of Hsp90 ATPase (Aha1) enhances ATP hydrolysis. The cofactor p23 stabilizes the closed dimer to slow the release of substrate protein from Hsp90.

Hsp90 consists of three domains: the highly conserved N-terminal domain, the middle domain, and the C-terminal domain (**Figure 7*a***) (169). The N-terminal domain contains the ATP binding site, which comprises a two-layer α/β-sandwich structure. Inhibition of ATP binding and hydrolysis either by mutagenesis of the N domain or by specific inhibitors, such as geldanamycin and radicicol, demonstrated the functional requirement of the Hsp90



Proteostasis network (PN): the collection of cellular components involved in proteostasis maintenance

ATPase (81, 169). The N-terminal domain contains a long, highly charged loop segment with regulatory function located in proximity to the middle domain (178, 179). The middle domain, composed of α-β-α motifs, is essential for interactions with substrate proteins and regulation of ATP hydrolysis (179a, 180, 181). The C-terminal domain of Hsp90 contains the interaction site for the assembly of the functional Hsp90 dimer, a mixed four-helix bundle (173). The extreme C terminus of Hsp90 contains the MEEVD sequence motif, which mediates interactions with numerous cochaperones containing TPR domains (81).

Despite an abundance of structural information, the detailed mechanism by which Hsp90 recognizes kinases and other substrate proteins and facilitates their folding or conformational regulation is not well understood. Similar to other chaperones, the Hsp90 dimer undergoes an ATP-regulated reaction cycle accompanied by extensive structural rearrangements (**Figure 7*b***) (100, 182). In this process, the open, V-shaped state of Hsp90 receives inactive client protein and then converts into the closed form, often described as a molecular clamp formed by the N domains (176, 182). This reaction is driven by the combined effects of ATP binding, ATP hydrolysis, posttranslational modifications (183, 184), and interactions with multiple cochaperones (**Figure 7*b***). Formation of the closed conformation results in a compaction of the Hsp90 dimer in which the individual subunits twist around each other (171). After hydrolysis the Hsp90 N termini separate, releasing the client protein in an active state (**Figure 7*b***). Recent electron microscopy and FRET studies using ATP and nonhydrolyzable ATP analogs suggest that the N-terminally clamped, closed state and the open conformation are in a more dynamic equilibrium than previously believed (185–187).

Cochaperones. Hsp90 intimately cooperates with multiple cochaperones that regulate different steps of the reaction cycle (**Figure 7*b***). Cochaperones HOP (Sti1) and Cdc37 (p50) stabilize the open conformation of the Hsp90 dimer (81, 174, 175, 188), inhibit ATP hydrolysis, and facilitate substrate protein binding. HOP mediates client transfer from the Hsp70-Hsp40 system, whereas Cdc37 functions as an adapter for kinases. Conversely, p23 (Sba1) couples the Hsp90 ATPase activity to efficient polypeptide dissociation. Two molecules of p23 bind to the N domains of closed Hsp90 and presumably stabilize the ATP-bound state (81, 171, 189). The activator of Hsp90 ATPase (Aha1) binds asymmetrically to the Hsp90 middle domain, stimulating ATP hydrolysis and inducing transition to the closed state (180, 181, 185, 190). In addition, Hsp90 interacts via its C terminus with a range of TPR domain–containing cochaperones. These factors often contain PPIase domains (Cyp40, FKBP51, and FKBP52) and participate in Hsp90-mediated client protein folding (45).

PROTEOSTASIS NETWORK

The successful folding of newly synthesized proteins and their conformational maintenance are essential in sustaining a functional proteome. In addition, the cellular concentration, localization, and activity of each protein must be carefully controlled in response to intrinsic and environmental stimuli. Although research has made major advances in elucidating the mechanics of individual chaperones, we are far from understanding how the various chaperone systems cooperate as a network and function in conjunction with the protein transport and degradation machineries to ensure proteome integrity. The term proteostasis describes this state of healthy proteome balance, whereas proteostasis network (PN) refers to the collection of cellular components involved in proteostasis maintenance (8). Failure of proteostasis is implicated in disease and the deleterious effects of aging (10). Molecular chaperones, through their ability to recognize incorrectly folded proteins, have multiple key roles in the PN (**Figure 8*a***).

The PN is regulated by interconnected pathways that respond to specific forms of

cellular stress, including the cytosolic heat shock response (HSR) (191), the unfolded protein response (UPR) in the endoplasmic reticulum (192), and the mitochondrial UPR (193). Additionally, PN regulation is integrated with pathways involved in inflammation, response to oxidative stress, caloric restriction/starvation, and longevity. The PN of mammalian cells consists of ~1,300 different proteins involved in protein biogenesis (~400), conformational maintenance (~300), and degradation (~700), with many proteins being part of more than one pathway (**Figure 8*b***). Different cell types vary in their proteostasis capacity and thus in their stress sensitivity and vulnerability to protein aggregation (194, 195).

Significance of Conformational Maintenance

After their initial folding and assembly, many proteins remain reliant on molecular chaperones throughout their cellular lifetime to maintain their functionally active conformations. This is consistent with the notion that proteins with key cellular functions are often structurally dynamic and may be expressed at levels at which they are poorly soluble (14, 46). Many of the chaperone systems discussed in the previous sections function not only in de novo folding but also in conformational maintenance, i.e., they prevent aggregation of misfolded proteins and mediate their refolding. Specific proteins may interact with as many as 25 different types of chaperones throughout their lifetime, as shown in yeast (51). Pulse-chase labeling and quantitative proteomics have described the contributions of the bacterial Hsp70 and chaperonin systems to conformational maintenance (15, 56). Upregulation of chaperones under conditions of conformational stress, such as heat shock or oxidative stress, expands the cellular capacity for the prevention of aggregation. Failure of conformational maintenance is particularly relevant to the onset of age-related degenerative disorders, which typically involve protein aggregation (9).

Degradation

A central feature in the organization of the PN is the tight interconnection of molecular chaperone functions with the pathways of protein degradation, which serve to remove nonfunctional, misfolded, or aggregated proteins that may otherwise disrupt proteostasis. The PN branch of degradation includes the ubiquitin-proteasome system (UPS) and machinery of autophagy (23, 196–200). Approximately 700 proteins are implicated in protein degradation, reflecting the fundamental importance of these pathways in cell regulation and protein homeostasis.

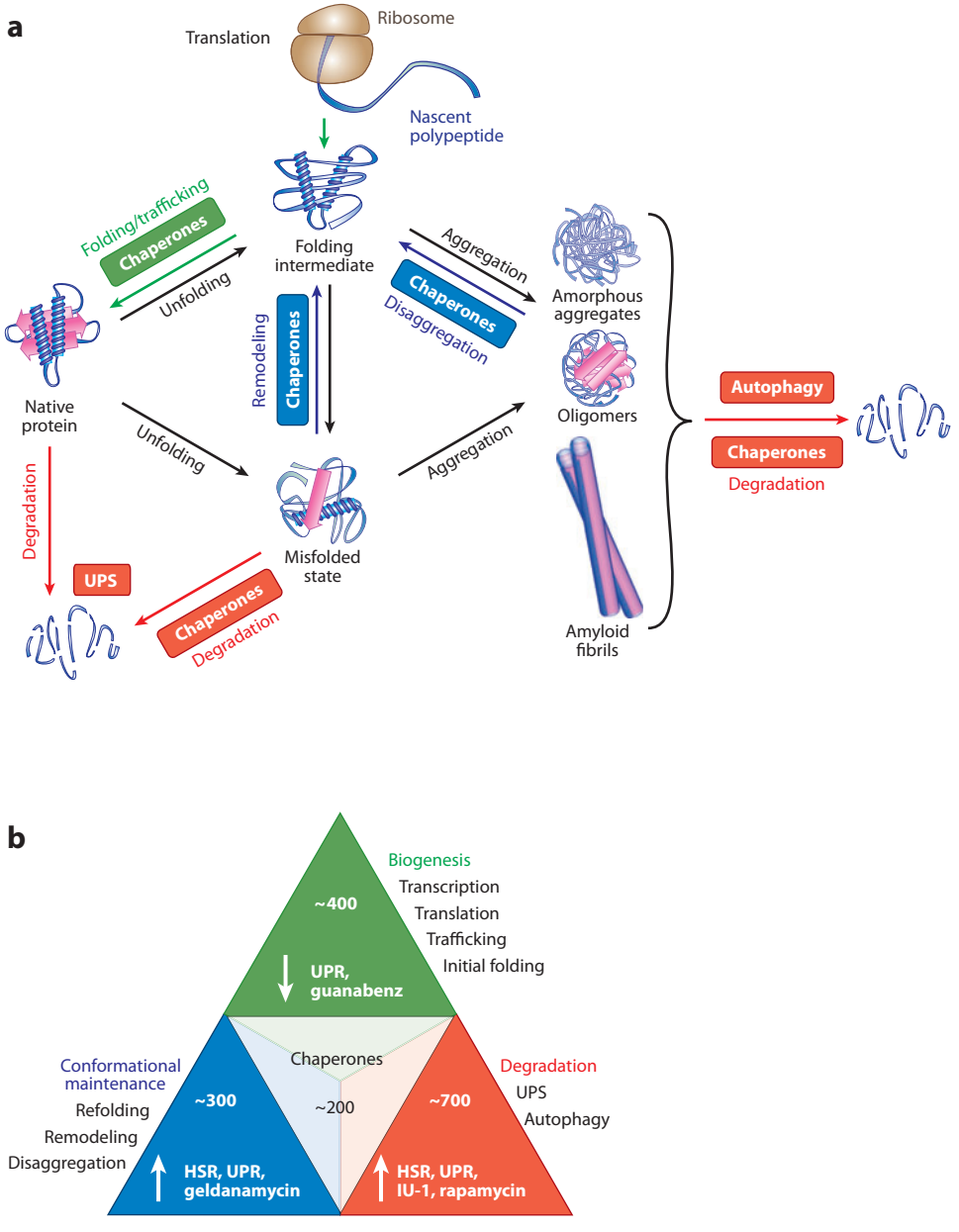
Degradation via the UPS depends on protein unfolding by the 26S proteasome (201) and generally requires that chaperones maintain target proteins in a nonaggregated state. Chaperones cooperate with various E3 ubiquitin ligases in recognizing and targeting misfolded proteins for proteasomal degradation. For example, Hsp70 and Hsp90 cooperate with the U-box-dependent ubiquitin ligase CHIP (C terminus of Hsc70-interacting protein) and a variety of other cofactors (such as BAG1 and BAG3) to ubiquitinate client proteins (23, 86, 198). CHIP interacts with either Hsc70 or Hsp90 via its TPR domain (86, 198, 202). CHIP also cooperates with the E2 enzyme Ubc13-Uev1a to form noncanonical Lys63-linked polyubiquitin chains, which suggests an additional role in targeting proteins for destruction via autophagy (202, 203). BAG1 and BAG3 associate (via their BAG domain) with the NBD of Hsp70 and also interact with CHIP to promote the ubiquitination of Hsp70-bound client proteins (86, 198, 204). BAG1 targets proteins for degradation by the UPS, whereas BAG3 mediates degradation by macroautophagy.

Aggregated proteins that cannot be unfolded for proteasomal degradation may be removed by autophagy and lysosomal/vacuolar degradation. Loss of autophagy causes inclusion body formation and neurodegeneration, even in the absence of additional stress, demonstrating the importance of this pathway for proteostasis (205). Aggregate removal by

UPR: unfolded protein response

UPS: ubiquitin-proteasome system

autophagy entails the sorting and concentration of small protein aggregates to specific sites in the cytosol (206–209) with the participation of chaperones such as Hsp42 (small Hsp) (209). These deposition sites include the aggresome, to which components mediating autophagic vacuole formation are recruited (206, 210, 211). Additional pathways of autophagy include chaperone-mediated autophagy (CMA) and chaperone-assisted selective autophagy (CASA) (198, 200, 212). In CMA, Hsc70 and certain cochaperones bind to a KFERQ sequence motif present in approximately 30% of all cytoplasmic proteins. This complex binds the lysosomal membrane protein LAMP2, followed by translocation of the substrate



protein across the lysosomal membrane for degradation (200). Whereas CMA is ubiquitin-independent, CASA uses ubiquitination as a signal for degradation in a pathway that involves the ubiquitin ligase CHIP, Hsc70, and BAG3, as well as the autophagic ubiquitin adaptor p62 (213). CASA is reportedly required for the removal of damaged skeletal muscle proteins such as filamin (213). Notably, chaperones such as Hsp70/Hsc70 act in parallel in the different branches of the PN, and our understanding of how they switch their function between initial folding and conformational maintenance to degradation is still rudimentary.

Proteostasis Collapse in Aging and Disease

As shown in *Caenorhabditis elegans*, *Drosophila*, and the mouse, the ability of cells and tissues to maintain proteostasis declines during aging, concurrent with the capacity to respond to conformational stresses (214–220). Why this is the case is still unclear, but one proposed explanation is that multicellular organisms place less value on protecting the somatic proteome against internal and external stress once propagation of the germ line is certain. The gradual decline in proteostasis capacity would then result in the accumulation of misfolded (or oxidized) proteins, leading to the deposition of aggregates, cellular toxicity, and cell death (214, 219, 221). Accordingly, age is a universal risk factor for a range of degenerative diseases associated with protein misfolding and aggregate deposition.

The diseases of aberrant protein folding associated with aging are usually categorized as toxic gain-of-function disorders and include type 2 diabetes and the major neurodegenerative diseases (Parkinson's, Huntington's, and Alzheimer's disease, as well as amyotrophic lateral sclerosis) (17, 222). They either occur sporadically or are dominantly inherited. Classical examples of the latter type are Huntington's disease and related disorders, in which the age of onset is inversely correlated with the length of an expanded polyglutamine tract in the disease protein (214). As aggregation propensity increases with polyglutamine length, manifestation of neuronal degeneration may occur when available proteostasis capacity is no longer sufficient to prevent the formation of toxic aggregates. This is supported by experiments demonstrating that the onset of polyglutamine toxicity in *C. elegans* correlates with age-dependent proteostasis decline (214, 215, 217, 220). The accumulation of protein aggregates in turn exerts pressure on the PN, further accelerating its decline. This is exemplified by the observation that toxic polyglutamine repeat proteins interfere with normal protein clearance by the UPS and the conformational maintenance of metastable proteins by the chaperone network (17, 195, 223, 224). Furthermore, toxic protein aggregation also

Figure 8

The proteostasis network. (*a*) Protein fates in the proteostasis network (PN). The PN integrates chaperone pathways for the folding of newly synthesized proteins, the remodeling of misfolded states, and disaggregation with protein degradation. Panel *a* adapted and modified from Reference 16. (*b*) Central role of molecular chaperones. The three branches of the proteostasis network are interconnected by the functions of molecular chaperones. The approximate number of proteins in each branch as well as the number of chaperone components (including cofactors) is indicated. Activation of the cytosolic heat shock response (HSR) and the unfolded protein response (UPR) of the endoplasmic reticulum generally increases proteostasis capacity in all three branches. Various heat shock factor 1 activators can pharmacologically induce the HSR (38, 229, 230). Treatment with guanabenz results in attenuation of translation and increases proteostasis capacity by reducing the load of potentially misfolding proteins (226). The small molecule compound IU-1 inhibits protein deubiquitination and increases degradation by the ubiquitin-proteasome system (UPS) (236). The drug rapamycin activates autophagy by inhibiting the kinase mammalian target of rapamycin (mTOR) (199).

compromises the cellular response to stress stimuli (221).

The Proteostasis Network as a Drug Target

Considering the impact of proteostasis imbalance in age-related degenerative diseases, returning the cell to a more youthful state by pharmacologically boosting proteostasis capacity is a promising therapeutic strategy. Whereas ligand compounds can stabilize specific disease proteins against aggregation (225), activating proteostasis could benefit a wide variety of diseases and might also delay the deleterious effects of aging (8). In principle, we might achieve this by manipulating the three branches of the PN: biogenesis, conformational maintenance, and degradation—either individually or in combination (**Figure 8*b***). In the biogenesis branch, attenuation of translation may be beneficial by reducing the load of misfolding proteins. The antihypertensive drug guanabenz demonstrates this; besides being an α_2 receptor antagonist, it also stabilizes the translation initiation factor 2 (eIF2) in its inactive, phosphorylated state (226). Furthermore, boosting cellular chaperone capacity can increase the efficiency of folding or degradation of proteins carrying destabilizing mutations and inhibit their aggregation (227). For example, small molecules (e.g., geldanamycin) that activate heat shock factor 1, the main transcriptional regulator of the cytosolic stress response, increase the effective concentration of cytosolic chaperones and suppress the aggregation of various disease proteins (8, 38, 228–230). This approach is based on multiple lines of evidence demonstrating that overexpression of chaperones such as Hsp70 and Hsp40 prevents the aggregation and toxicity of huntingtin and α-synuclein (38, 231–234). The Hsp70 system acts synergistically with the cytosolic chaperonin TRiC to prevent aggregation of proteins with expanded polyglutamine tracts (165–168). Finally, activating the UPS or inducing autophagy can increase the clearance of potentially toxic proteins (**Figure 8*b***) (212, 235, 236).

OUTLOOK

Studies over the past decade have revealed fascinating insights into the structures of a variety of chaperone systems and the mechanisms by which they assist in protein folding. However, most of these advances are derived from analyses in vitro, and consequently, our understanding of how the pathways of folding in the cell differ from those studied in the test tube is still incomplete. Moreover, for most newly synthesized proteins, the relevant quantitative parameters of folding (rate, yield, and overall efficiency) are unknown. Likewise, we are just beginning to understand how the cellular environment influences protein folding and stability and how translation affects the folding process. For example, what is the role of translational pausing in protein folding and trafficking? Much future work will also be directed toward developing an integrated view of the different aspects of the PN, with particular regard to the cooperation between folding and degradation machineries. Solving this problem will require a broad systems biology approach relying on a combination of ribosome profiling, quantitative proteomics, and computational modeling. How do cells react to conformational stresses or proteostasis deficiencies at the proteome level? Which proteins are prone to misfolding, and why do certain proteins aggregate into toxic species whereas others get degraded? How does the composition of the proteome change during aging, what are the signatures of a youthful proteome, and can we find ways to preserve it longer as we age? Addressing these and related questions offers not only a deeper understanding of cell biology but also the prospect of great medical benefits should we be able to intervene in the numerous, presently incurable diseases of protein aggregation and proteostasis deficiency.

SUMMARY POINTS

1. The efficient biogenesis of proteins in the densely crowded cellular environment depends on molecular chaperones to avoid protein misfolding and aggregation.
2. Chaperones promote folding and inhibit aggregation through cycles of binding and release of nonnative proteins (often ATP regulated) that allow kinetic partitioning.
3. Different classes of molecular chaperones may cooperate in sequential pathways.
4. Nascent-chain-binding chaperones prevent misfolding during translation; folding occurs either immediately upon completion of synthesis or after transfer to downstream chaperones, such as the chaperonins, which complete the folding process.
5. Major ATP-dependent chaperone paradigms in the cytosol include the Hsp70 and Hsp90 systems as well as the chaperonins.
6. The chaperonins are cylindrical, ATP-dependent folding machines that encapsulate a single protein chain, allowing it to overcome kinetic folding barriers while being protected against aggregation.
7. Molecular chaperones function as central elements of the large cellular network of proteostasis control, which comprises the protein biogenesis machinery as well as the ubiquitin proteasome and autophagy systems for protein degradation.
8. Understanding the organization of this network and its regulation during stress and aging will help in developing new strategies for the treatment of a range of age-related degenerative diseases associated with protein aggregation.

FUTURE ISSUES

1. How do the pathways of protein folding in the cell differ from those studied in vitro and how does translation affect the folding process?
2. Can we determine the rates, yields, and overall efficiencies of protein folding at the proteome level?
3. How do molecular chaperones of the various branches of the PN cooperate in maintaining proteome integrity?
4. How does the proteome composition change during proteostasis, stress, and aging?
5. What are the mechanisms underlying the cellular toxicity of protein misfolding and aggregation?
6. Can pharmacological chaperone activation serve as a strategy to combat diseases associated with protein misfolding and aggregation?

DISCLOSURE STATEMENT

F.U.H. and M.S.H. are paid consultants of Proteostasis Therapeutics, Inc.

ACKNOWLEDGMENTS

We thank S. Pinkert for providing the bioinformatic analysis of the PN. We apologize that references to original work older than ten years (prior to 2002) could not be cited owing to space limitations. For a description of these contributions, the reader is referred to relevant review articles.

LITERATURE CITED

1. Ellis RJ, Minton AP. 2006. Protein aggregation in crowded environments. *Biol. Chem.* 387:485–97
2. Dobson CM, Sali A, Karplus M. 1998. Protein folding: a perspective from theory and experiment. *Angew. Chem. Int. Ed.*37:868–93
3. Bartlett AI, Radford SE. 2009. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* 16:582–88
4. Hartl FU. 1996. Molecular chaperones in cellular protein folding. *Nature* 381:571–79
5. Bukau B, Horwich AL. 1998. The Hsp70 and Hsp60 chaperone machines. *Cell* 92:351–66
6. Frydman J. 2001. Folding of newly translated proteins in vivo: the role of molecular chaperones. *Annu. Rev. Biochem.* 70:603–47
7. Hartl FU, Hayer-Hartl M. 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295:1852–58
8. Balch WE, Morimoto RI, Dillin A, Kelly JW. 2008. Adapting proteostasis for disease intervention. *Science* 319:916–19
9. Chiti F, Dobson CM. 2006. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75:333–66
10. Morimoto RI. 2008. Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. *Genes Dev.* 22:1427–38
11. Skach WR. 2009. Cellular mechanisms of membrane protein folding. *Nat. Struct. Mol. Biol.* 16:606–12
12. Braakman I, Bulleid NJ. 2011. Protein folding and modification in the mammalian endoplasmic reticulum. *Annu. Rev. Biochem.* 80:71–99
13. Araki K, Nagata K. 2011. Protein folding and quality control in the ER. *Cold Spring Harb. Perspect. Biol.* 3:a007526
14. Dunker AK, Silman I, Uversky VN, Sussman, JL. 2008. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* 18:756–64
15. Kerner MJ, Naylor DJ, Ishihama Y, Maier T, Chang HC, et al. 2005. Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* 122:209–20
16. Eichner T, Radford SE. 2011. A diversity of assembly mechanisms of a generic amyloid fold. *Mol. Cell* 43:8–18
17. Hartl FU, Bracher A, Hayer-Hartl M. 2011. Molecular chaperones in protein folding and proteostasis. *Nature* 475:324–32
18. Bolognesi B, Kumita JR, Barros TP, Esbjorner EK, Luheshi LM, et al. 2010. ANS binding reveals common features of cytotoxic amyloid species. *ACS Chem. Biol.* 5:735–40
19. Ignatova Z, Gierasch LM. 2004. Monitoring protein stability and aggregation *in vivo* by real-time fluorescent labeling. *Proc. Natl. Acad. Sci. USA* 101:523–28
20. Gershenson A, Gierasch LM. 2011. Protein folding in the cell: challenges and progress. *Curr. Opin. Struct. Biol.* 21:32–41
21. McGuffee SR, Elcock AH. 2006. Atomically detailed simulations of concentrated protein solutions: the effects of salt, pH, point mutations, and protein concentration in simulations of 1000-molecule systems. *J. Am. Chem. Soc.* 128:12098–110
22. Hong J, Gierasch LM. 2010. Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J. Am. Chem. Soc.* 132:10445–52
23. Young JC, Barral JM, Hartl FU. 2003. More than folding: localized functions of cytosolic chaperones. *Trends Biochem. Sci.* 28:541–47

24. Ellis JP, Culviner PH, Cavagnero S. 2009. Confined dynamics of a ribosome-bound nascent globin: cone angle analysis of fluorescence depolarization decays in the presence of two local motions. *Protein Sci.* 18:2003–15
25. Fedyukina DV, Cavagnero S. 2011. Protein folding at the exit tunnel. *Annu. Rev. Biophys.* 40:337–59
26. Woolhead CA, McCormick PJ, Johnson AE. 2004. Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* 116:725–36
27. Lu JL, Deutsch C. 2005. Folding zones inside the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* 12: 1123–29
28. O'Brien EP, Hsu S-TD, Christodoulou J, Vendruscolo M, Dobson CM. 2010. Transient tertiary structure formation within the ribosome exit port. *J. Am. Chem. Soc.* 132:16928–37
29. Wilson DN, Beckmann R. 2011. The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Curr. Opin. Struct. Biol.* 21:274–82
30. Elcock AH. 2006. Molecular simulations of cotranslational protein folding: fragment stabilities, folding cooperativity, and trapping in the ribosome. *PLoS Comput. Biol.* 2:e98
31. Cabrita LD, Hsu ST, Launay H, Dobson CM, Christodoulou J. 2009. Probing ribosome-nascent chain complexes produced in vivo by NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* 106:22239–44
32. Eichmann C, Preissler S, Riek R, Deuerling E. 2010. Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* 107:9111–16
33. Ugrinov KG, Clark PL. 2010. Cotranslational folding increases GFP folding yield. *Biophys. J.* 98:1312–20
34. Zhang G, Ignatova Z. 2011. Folding at the birth of the nascent chain: coordinating translation with co-translational folding. *Curr. Opin. Struct. Biol.* 21:25–31
35. Kaiser CM, Goldman DH, Chodera JD, Tinoco I Jr, Bustamante C. 2011. The ribosome modulates nascent protein folding. *Science* 334:1723–27
36. Netzer WJ, Hartl FU. 1998. Protein folding in the cytosol: chaperonin-dependent and -independent mechanisms. *Trends Biochem. Sci.* 23:68–73
37. Agashe VR, Guha S, Chang HC, Genevaux P, Hayer-Hartl M, et al. 2004. Function of trigger factor and DnaK in multidomain protein folding: increase in yield at the expense of folding speed. *Cell* 117:199–209
38. Hartl FU, Hayer-Hartl M. 2009. Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.* 16:574–81
39. Clarke TF IV, Clark PL. 2008. Rare codons cluster. *PLoS ONE* 3:e3412
40. Li G-W, Oh E, Weissman JS. 2012. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* 484:538–41
41. O'Brien EP, Vendruscolo M, Dobson CM. 2012. Prediction of variable translation rate effects on cotranslational protein folding. *Nat. Commun.* 3:868
42. Spencer PS, Siller E, Anderson JF, Barral JM. 2012. Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. *J. Mol. Biol.* 422:328–35
43. Brandt F, Etchells SA, Ortiz JO, Elcock, AH, Hartl FU, Baumeister W. 2009. The native 3D organization of bacterial polysomes. *Cell* 136:261–71
44. Tokuriki N, Tawfik DS. 2009. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* 459:668–73
45. Taipale M, Jarosz DF, Lindquist S. 2010. Hsp90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* 11:515–28
46. Tartaglia GG, Dobson CM, Hartl FU, Vendruscolo M. 2010. Physicochemical determinants of chaperone requirements. *J. Mol. Biol.* 400:579–88
47. Albanese V, Yam AYW, Baughman J, Parnot C, Frydman J. 2006. Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. *Cell* 124:75–88
48. McClellan AJ, Xia Y, Deutschbauer AM, Davis RW, Gerstein M, Frydman J. 2007. Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches. *Cell* 131:121–35
49. Dekker C, Stirling PC, McCormack EA, Filmore H, Paul A, et al. 2008. The interaction network of the chaperonin CCT. *EMBO J.* 27:1827–39

50. Yam AY, Xia Y, Lin HT, Burlingame A, Gerstein M, Frydman J. 2008. Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nat. Struct. Mol. Biol.* 15:1255–62
51. Gong Y, Kakihara Y, Krogan N, Greenblatt J, Emili A, et al. 2009. An atlas of chaperone-protein interactions in *Saccharomyces cerevisiae*: implications to protein folding pathways in the cell. *Mol. Syst. Biol.* 5:275
52. Hirtreiter AM, Calloni G, Forner F, Scheibe B, Puype M, et al. 2009. Differential substrate specificity of group I and group II chaperonins in the archaeon *Methanosarcina mazei*. *Mol. Microbiol.* 74:1152–68
53. Fujiwara K, Ishihama Y, Nakahigashi K, Soga T, Taguchi H. 2010. A systematic survey of *in vivo* obligate chaperonin-dependent substrates. *EMBO J.* 29:1552–64
54. Del Alamo M, Hogan DJ, Pechmann S, Albanese V, Brown PO, Frydman J. 2011. Defining the specificity of cotranslationally acting chaperones by systematic analysis of mRNAs associated with ribosome-nascent chain complexes. *PLoS Biol.* 9:e1001100
55. Oh E, Becker AH, Sandikci A, Huber D, Chaba R, et al. 2011. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* 147:1295–308
56. Calloni G, Chen T, Schermann SM, Chang H-C, Genevaux P, et al. 2012. DnaK functions as a central hub in the *E. coli* chaperone network. *Cell Rep.* 1:251–64
57. Powers ET, Powers DL, Gierasch LM. 2012. FoldEco: a model for proteostasis in *E. coli*. *Cell Rep.* 1:265–76
58. Bukau B, Deuerling E, Pfund C, Craig EA. 2000. Getting newly synthesized proteins into shape. *Cell* 101:119–22
59. Preissler S, Deuerling E. 2012. Ribosome-associated chaperones as key players in proteostasis. *Trends Biochem. Sci.* 37:274–83
60. Kramer G, Rauch T, Rist W, Vorderwülbecke S, Patzelt H, et al. 2002. L23 protein functions as a chaperone docking site on the ribosome. *Nature* 419:171–74
61. Ferbitz L, Maier T, Patzelt H, Bukau B, Deuerling B, Ban N. 2004. Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* 431:590–96
62. Merz F, Boehringer D, Schaffitzel C, Preissler S, Hoffmann A, et al. 2008. Molecular mechanism and structure of trigger factor bound to the translating ribosome. *EMBO J.* 27:1622–32
63. Kaiser CM, Chang H-C, Agashe VR, Lakshmipathy SK, Etchells SA, et al. 2006. Real-time observation of trigger factor function on translating ribosomes. *Nature* 444:455–60
64. Eisner G, Koch HG, Beck K, Brunner J, Müller M. 2003. Ligand crowding at a nascent signal sequence. *J. Cell Biol.* 163:35–44
65. Ullers RS, Houben EN, Raine A, ten Hagen-Jongman CM, Ehrenberg M, et al. 2003. Interplay of signal recognition particle and trigger factor at L23 near the nascent chain exit site on the *Escherichia coli* ribosome. *J. Cell Biol.* 161:679–84
66. Bingel-Erlenmeyer R, Kohler R, Kramer G, Sandikci A, Antolić S, et al. 2008. A peptide deformylase-ribosome complex reveals mechanism of nascent chain processing. *Nature* 452:108–13
67. Genevaux P, Keppel F, Schwager F, Langendijk-Genevaux PS, Hartl FU, Georgopoulos C. 2004. *In vivo* analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* 5:195–200
68. Gautschi M, Mun A, Ross S, Rospert S. 2002. A functional chaperone triad on the yeast ribosome. *Proc. Natl. Acad. Sci. USA* 99:4209–14
69. Raue U, Oellerer S, Rospert S. 2007. Association of protein biogenesis factors at the yeast ribosomal tunnel exit is affected by the translational status and nascent polypeptide sequence. *J. Biol. Chem.* 282:7809–16
70. Peisker K, Braun D, Wölfle T, Hentschel J, Fünfschilling U, et al. 2008. Ribosome-associated complex binds to ribosomes in close proximity of Rpl31 at the exit of the polypeptide tunnel in yeast. *Mol. Biol. Cell* 19:5279–88
71. Koplin A, Preissler S, Ilina Y, Koch M, Scior A, et al. 2010. A dual function for chaperones SSB-RAC and the NAC nascent polypeptide-associated complex on ribosomes. *J. Cell Biol.* 189:57–68
72. Wegrzyn RD, Hofmann D, Merz F, Nikolay R, Rauch T, et al. 2006. A conserved motif is prerequisite for the interaction of NAC with ribosomal protein L23 and nascent chains. *J. Biol. Chem.* 281:2847–57
73. Pech M, Spreter T, Beckmann R, Beatrix B. 2010. Dual binding mode of the nascent polypeptide-associated complex reveals a novel universal adapter site on the ribosome. *J. Biol. Chem.* 285:19679–87

74. Niwa T, Kanamori T, Ueda T, Taguchi H. 2012. Global analysis of chaperone effects using a reconstituted cell-free translation system. *Proc. Natl. Acad. Sci. USA* 109:8937–42
75. Kampinga HH, Craig EA. 2010. The Hsp70 chaperone machinery: J proteins as drivers of functional specificity. *Nat. Rev. Mol. Cell Biol.* 11:579–92
76. Tang YC, Chang HC, Hayer-Hartl M, Hartl FU. 2007. Snapshot: molecular chaperones, part II. *Cell* 128:412
77. Horwich AL, Fenton WA, Chapman E, Farr GW. 2007. Two families of chaperonin: physiology and mechanism. *Annu. Rev. Cell Dev. Biol.* 23:115–45
78. Etchells SA, Meyer AS, Yam AY, Roobol A, Miao Y, et al. 2005. The cotranslational contacts between ribosome-bound nascent polypeptides and the subunits of the hetero-oligomeric chaperonin TRiC probed by photocross-linking. *J. Biol. Chem.* 280:28118–26
79. Cuellar J, Martin-Benito J, Scheres SH, Sousa R, Moro F, et al. 2008. The structure of CCT-Hsc70NBD suggests a mechanism for Hsp70 delivery of substrates to the chaperonin. *Nat. Struct. Mol. Biol.* 15:858–64
80. Zhao R, Houry WA. 2007. Molecular interaction network of the Hsp90 chaperone system. *Adv. Exp. Med. Biol.* 594:27–36
81. Young JC, Moarefi I, Hartl FU. 2001. Hsp90: a specialized but essential protein-folding tool. *J. Cell Biol.* 154:267–73
82. Young JC, Hoogenraad NJ, Hartl FU. 2003. Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. *Cell* 112:41–50
83. Sharma K, Vabulas RM, Macek B, Pinkert S, Cox J, et al. 2012. Quantitative proteomics reveals that Hsp90 inhibition preferentially targets kinases and the DNA damage response. *Mol. Cell. Proteomics* 11:M111.014654
84. Neckers L. 2007. Heat shock protein 90: the cancer chaperone. *J. Biosci.* 32:517–30
85. Haslberger T, Bukau B, Mogk A. 2010. Towards a unifying mechanism for ClpB/Hsp104-mediated protein disaggregation and prion propagation. *Biochem. Cell Biol.* 88:63–75
86. Arndt V, Rogon C, Höhfeld J. 2007. To be, or not to be—molecular chaperones in protein degradation. *Cell. Mol. Life Sci.* 64:2525–41
87. Lakshmipathy SK, Tomic S, Kaiser CM, Chang HC, Genevaux P, et al. 2007. Identification of nascent chain interaction sites on trigger factor. *J. Biol. Chem.* 282:12186–93
88. Lakshmipathy SK, Gupta R, Pinkert S, Etchells SA, Hartl FU. 2010. Versatility of trigger factor interactions with ribosome-nascent chain complexes. *J. Biol. Chem.* 285:27911–23
89. Baram D, Pyetan E, Sittner A, Auerbach-Nevo T, Bashan A, Yonath A. 2005. Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action. *Proc. Natl. Acad. Sci. USA* 102:12017–22
90. Schlünzen F, Wilson DN, Tian P, Harms JM, McInnes SJ, et al. 2005. The binding mode of the trigger factor on the ribosome: implications for protein folding and SRP interaction. *Structure* 13:1685–94
91. Kramer G, Rutkowska A, Wegrzyn RD, Patzelt H, Kurz TA, et al. 2004. Functional dissection of *Escherichia coli* trigger factor: unraveling the function of individual domains. *J. Bacteriol.* 186:3777–84
92. Jakob RP, Zoldak G, Aumüller T, Schmid FX. 2009. Chaperone domains convert prolyl isomerases into generic catalysts of protein folding. *Proc. Natl. Acad. Sci. USA* 106:20282–87
93. Merz F, Hoffmann A, Rutkowska A, Zachmann-Brand B, Bukau B, Deuerling E. 2006. The C-terminal domain of *Escherichia coli* trigger factor represents the central module of its chaperone activity. *J. Biol. Chem.* 281:31963–71
94. Patzelt H, Kramer G, Rauch T, Schönfeld HJ, Bukau B, Deuerling E. 2002. Three-state equilibrium of *Escherichia coli* trigger factor. *Biol. Chem.* 383:1611–19
95. Gupta R, Lakshmipathy SK, Chang HC, Etchells SA, Hartl FU. 2010. Trigger factor lacking the PPIase domain can enhance the folding of eukaryotic multi-domain proteins in *Escherichia coli*. *FEBS Lett.* 584:3620–24
96. Hoffmann A, Becker AH, Zachmann-Brand B, Deuerling E, Bukau B, Kramer G. 2012. Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Mol. Cell* 48:63–74

97. Liu C-P, Perrett S, Zhou J-M. 2005. Dimeric trigger factor stably binds folding-competent intermediates and cooperates with the DnaK-DnaJ-GrpE chaperone system to allow refolding. *J. Biol. Chem.* 280:13315–20
98. Martinez-Hackert E, Hendrickson WA. 2009. Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. *Cell* 138:923–34
99. Zuiderweg ER, Bertelsen EB, Rousaki A, Mayer MP, Gestwicki JE, Ahmad A. 2013. Allostery in the Hsp70 chaperone proteins. *Top. Curr. Chem.* 328:99–152
100. Mayer MP. 2010. Gymnastics of molecular chaperones. *Mol. Cell* 39:321–31
101. Kityk R, Kopp J, Sinning I, Mayer MP. 2012. Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones. *Mol. Cell* 48:863–74
102. Deleted in proof
103. Liu QL, Hendrickson WA. 2007. Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. *Cell* 131:106–20
104. Polier S, Dragovic Z, Hartl FU, Bracher A. 2008. Structural basis for the cooperation of Hsp70 and Hsp110 chaperones in protein folding. *Cell* 133:1068–79
105. Schuermann JP, Jiang JW, Cuellar J, Llorca O, Wang L, et al. 2008. Structure of the Hsp110:Hsc70 nucleotide exchange machine. *Mol. Cell* 31:232–43
106. Bertelsen EB, Chang L, Gestwicki JE, Zuiderweg ER. 2009. Solution conformation of wild-type *E. coli* Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proc. Natl. Acad. Sci. USA* 106:8471–76
107. Mapa K, Sikor M, Kudryavtsev V, Waegemann K, Kalinin S, et al. 2010. The conformational dynamics of the mitochondrial Hsp70 chaperone. *Mol. Cell* 38:89–100
108. Schweizer RS, Aponte RA, Zimmermann S, Weber A, Reinstein J. 2011. Fine tuning of a biological machine: DnaK gains improved chaperone activity by altered allosteric communication and substrate binding. *ChemBioChem* 12:1559–73
109. Swain JF, Dinler G, Sivendran R, Montgomery DL, Stotz M, Gierasch LM. 2007. Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Mol. Cell* 26:27–39
110. Smock RG, Rivoire O, Russ WP, Swain JF, Leibler S, et al. 2010. An interdomain sector mediating allostery in Hsp70 molecular chaperones. *Mol. Syst. Biol.* 6:414
111. Zhuravleva A, Gierasch LM. 2011. Allosteric signal transmission in the nucleotide-binding domain of 70-kDa heat shock protein (Hsp70) molecular chaperones. *Proc. Natl. Acad. Sci. USA* 108:6987–92
112. Jiang J, Maes EG, Taylor AB, Wang L, Hinck AP, et al. 2007. Structural basis of J cochaperone binding and regulation of Hsp70. *Mol. Cell* 28:422–33
113. Ahmad A, Bhattacharya A, McDonald RA, Cordes M, Ellington B, et al. 2011. Heat shock protein 70 kDa chaperone/DnaJ cochaperone complex employs an unusual dynamic interface. *Proc. Natl. Acad. Sci. USA* 108:18966–71
114. Svärd M, Biterova EI, Bourhis JM, Guy JE. 2011. The crystal structure of the human co-chaperone $p58^{IPK}$. *PLoS ONE* 6:e22337
115. Hagiwara M, Maegawa K-I, Suzuki M, Ushioda R, Araki K, et al. 2011. Structural basis of an ERAD pathway mediated by the ER-resident protein disulfide reductase ERdj5. *Mol. Cell* 41:432–44
116. Kabani M, Beckerich JM, Brodsky JL. 2002. Nucleotide exchange factor for the yeast Hsp70 molecular chaperone Ssa1p. *Mol. Cell Biol.* 22:4677–89
117. Steel GJ, Fullerton DM, Tyson JR, Stirling CJ. 2004. Coordinated activation of Hsp70 chaperones. *Science* 303:98–101
118. Dragovic Z, Broadley SA, Shomura Y, Bracher A, Hartl FU. 2006. Molecular chaperones of the Hsp110 family act as nucleotide exchange factors of Hsp70s. *EMBO J.* 25:2519–28
119. Raviol H, Sadlish H, Rodriguez F, Mayer MP, Bukau B. 2006. Chaperone network in the yeast cytosol: Hsp110 is revealed as an Hsp70 nucleotide exchange factor. *EMBO J.* 25:2510–18
120. Shomura Y, Dragovic Z, Chang HC, Tzvetkov N, Young JC, et al. 2005. Regulation of Hsp70 function by HspBP1: Structural analysis reveals an alternate mechanism for Hsp70 nucleotide exchange. *Mol. Cell* 17:367–79
121. Xu Z, Page RC, Gomes MM, Kohli E, Nix JC, et al. 2008. Structural basis of nucleotide exchange and client binding by the Hsp70 cochaperone Bag2. *Nat. Struct. Mol. Biol.* 15:1309–17

122. Arakawa A, Handa N, Ohsawa N, Shida M, Kigawa T, et al. 2010. The C-terminal BAG domain of BAG5 induces conformational changes of the Hsp70 nucleotide-binding domain for ADP-ATP exchange. *Structure* 18:309–19
123. Yan M, Li J, Sha B. 2011. Structural analysis of the Sil1-Bip complex reveals the mechanism for Sil1 to function as a nucleotide-exchange factor. *Biochem. J.* 438:447–55
124. Shorter J. 2011. The mammalian disaggregase machinery: Hsp110 synergizes with Hsp70 and Hsp40 to catalyze protein disaggregation and reactivation in a cell-free system. *PLoS ONE* 6:e26319
125. Rampelt H, Kirstein-Miles J, Nillegoda NB, Chi K, Scholz SR, et al. 2012. Metazoan Hsp70 machines use Hsp110 to power protein disaggregation. *EMBO J.* 31:4221–35
126. Anttonen AK, Mahjneh I, Hämäläinen RH, Lagier-Tourenne C, Kopra O, et al. 2005. The gene disrupted in Marinesco-Sjögren syndrome encodes SIL1, an HSPA5 cochaperone. *Nat. Genet.* 37:1309–11
127. Senderek J, Krieger M, Stendel C, Bergmann C, Moser M, et al. 2005. Mutations in SIL1 cause Marinesco-Sjögren syndrome, a cerebellar ataxia with cataract and myopathy. *Nat. Genet.* 37:1312–14
128. Elad N, Farr GW, Clare DK, Orlova EV, Horwich AL, Saibil HR. 2007. Topologies of a substrate protein bound to the chaperonin GroEL. *Mol. Cell* 26:415–26
129. Horst R, Bertelsen EB, Fiaux J, Wider G, Horwich AL, Wüthrich K. 2005. Direct NMR observation of a substrate protein bound to the chaperonin GroEL. *Proc. Natl. Acad. Sci. USA* 102:12748–53
130. Sharma S, Chakraborty K, Müller BK, Astola N, Tang Y-C, et al. 2008. Monitoring protein conformation along the pathway of chaperonin-assisted folding. *Cell* 133:142–53
131. Hillger F, Hanni D, Nettels D, Geister S, Grandin M, et al. 2008. Probing protein-chaperone interactions with single-molecule fluorescence spectroscopy. *Angew. Chem. Int. Ed.* 47:6184–88
132. Tang Y-C, Chang H-C, Roeben A, Wischnewski D, Wischnewski N, et al. 2006. Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* 125:903–14
133. Tang YC, Chang HC, Chakraborty K, Hartl FU, Hayer-Hartl M. 2008. Essential role of the chaperonin folding compartment *in vivo*. *EMBO J.* 27:1458–68
134. Clare DK, Bakkes PJ, van Heerikhuizen H, van der Vies SM, Saibil HR. 2009. Chaperonin complex with a newly folded protein encapsulated in the folding chamber. *Nature* 457:107–10
135. Motojima F, Yoshida M. 2010. Polypeptide in the chaperonin cage partly protrudes out and then folds inside or escapes outside. *EMBO J.* 29:4008–19
136. Horovitz A, Willison KR. 2005. Allosteric regulation of chaperonins. *Curr. Opin. Struct. Biol.* 15:646–51
137. Clare DK, Vasishtan D, Stagg S, Quispe J, Farr GW, et al. 2012. ATP-triggered conformational changes delineate substrate-binding and -folding mechanics of the GroEL chaperonin. *Cell* 149:113–23
138. Koike-Takeshita A, Yoshida M, Taguchi H. 2008. Revisiting the GroEL-GroES reaction cycle via the symmetric intermediate implied by novel aspects of the GroEL(D398A) mutant. *J. Biol. Chem.* 283:23774–81
139. Thirumalai D, Lorimer GH. 2001. Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 30:245–69
140. Lin Z, Madan D, Rye HS. 2008. GroEL stimulates protein folding through forced unfolding. *Nat. Struct. Mol. Biol.* 15:303–11
141. Chakraborty K, Chatila M, Sinha J, Shi Q, Poschner BC, et al. 2010. Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. *Cell* 142:112–22
142. Baumketner A, Jewett A, Shea JE. 2003. Effects of confinement in chaperonin assisted protein folding: rate enhancement by decreasing the roughness of the folding energy landscape. *J. Mol. Biol.* 332:701–13
143. Hayer-Hartl M, Minton AP. 2006. A simple semiempirical model for the effect of molecular confinement upon the rate of protein folding. *Biochemistry* 45:13356–60
144. England J, Lucent D, Pande V. 2008. Rattling the cage: computational models of chaperonin-mediated protein folding. *Curr. Opin. Struct. Biol.* 18:163–69
145. Tian J, Garcia AE. 2011. Simulation studies of protein folding/unfolding equilibrium under polar and nonpolar confinement. *J. Am. Chem. Soc.* 133:15157–64
146. Apetri AC, Horwich AL. 2008. Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. *Proc. Natl. Acad. Sci. USA* 105:17351–55
147. Lin Z, Rye HS. 2004. Expansion and compression of a protein folding intermediate by GroEL. *Mol. Cell* 16:23–34

148. Wang JD, Herman C, Tipton KA, Gross CA, Weissman JS. 2002. Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell* 111:1027–39
149. Raineri E, Ribeca P, Serrano L, Maier T. 2010. A more precise characterization of chaperonin substrates. *Bioinformatics* 26:1685–89
150. Mallam AL, Jackson SE. 2012. Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat. Chem. Biol.* 8:147–53
151. Shomura Y, Yoshida T, Iizuka R, Maruyama T, Yohda M, Miki K. 2004. Crystal structures of the group II chaperonin from *Thermococcus* strain KS-1: steric hindrance by the substituted amino acid, and inter-subunit rearrangement between two crystal forms. *J. Mol. Biol.* 335:1265–78
152. Huo Y, Hu Z, Zhang K, Wang L, Zhai Y, et al. 2010. Crystal structure of group II chaperonin in the open state. *Structure* 18:1270–79
153. Cong Y, Baker ML, Jakana J, Woolford D, Miller EJ, et al. 2010. 4.0-Å resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. *Proc. Natl. Acad. Sci. USA* 107:4967–72
154. Pereira JH, Ralston CY, Douglas NR, Meyer D, Knee KM, et al. 2010. Crystal structures of a group II chaperonin reveal the open and closed states associated with the protein folding cycle. *J. Biol. Chem.* 285:27958–66
155. Dekker C, Roe SM, McCormack EA, Beuron F, Pearl LH, Willison KR. 2011. The crystal structure of yeast CCT reveals intrinsic asymmetry of eukaryotic cytosolic chaperonins. *EMBO J.* 30:3078–90
156. Zhang J, Ma B, DiMaio F, Douglas NR, Joachimiak LA, et al. 2011. Cryo-EM structure of a group II chaperonin in the prehydrolysis ATP-bound state leading to lid closure. *Structure* 19:633–39
157. Munoz IG, Yebenes H, Zhou M, Mesa P, Serna M, et al. 2011. Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. *Nat. Struct. Mol. Biol.* 18:14–19
158. Cong Y, Schroder GF, Meyer AS, Jakana J, Ma B, et al. 2012. Symmetry-free cryo-EM structures of the chaperonin TRiC along its ATPase-driven conformational cycle. *EMBO J.* 31:720–30
159. Kalisman N, Adams CM, Levitt M. 2012. Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *Proc. Natl. Acad. Sci. USA* 109:2884–89
160. Leitner A, Joachimiak LA, Bracher A, Mönkemeyer L, Walzthoeni T, et al. 2012. The molecular architecture of the eukaryotic chaperonin TRiC/CCT. *Structure* 20:814–25
161. Douglas NR, Reissmann S, Zhang J, Chen B, Jakana J, et al. 2011. Dual action of ATP hydrolysis couples lid closure to substrate release into the group II chaperonin chamber. *Cell* 144:240–52
162. Meyer AS, Gillespie JR, Walther D, Millet IS, Doniach S, Frydman J. 2003. Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. *Cell* 113:369–81
163. Spiess C, Miller EJ, McClellan AJ, Frydman J. 2006. Identification of the TRiC/CCT substrate binding sites uncovers the function of subunit diversity in eukaryotic chaperonins. *Mol. Cell* 24:25–37
164. Reissmann S, Parnot C, Booth CR, Chiu W, Frydman J. 2007. Essential function of the built-in lid in the allosteric regulation of eukaryotic and archaeal chaperonins. *Nat. Struct. Mol. Biol.* 14:432–40
164a. Rüβmann F, Stemp MJ, Mönkemeyer L, Etchells SA, Bracher A, Hartl FU. 2012. Folding of large multidomain proteins by partial encapsulation in the chaperonin TRiC/CCT. *Proc. Natl. Acad. Sci. USA* 109:21208–15
165. Kitamura A, Kubota H, Pack CG, Matsumoto G, Hirayama S, et al. 2006. Cytosolic chaperonin prevents polyglutamine toxicity with altering the aggregation state. *Nat. Cell Biol.* 8:1163–70
166. Behrends C, Langer CA, Boteva R, Böttcher UM, Stemp MJ, et al. 2006. Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. *Mol. Cell* 23:887–97
167. Tam S, Geller R, Spiess C, Frydman J. 2006. The chaperonin TRiC controls polyglutamine aggregation and toxicity through subunit-specific interactions. *Nat. Cell Biol.* 8:1155–62
168. Tam S, Spiess C, Auyeung W, Joachimiak L, Chen B, et al. 2009. The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation. *Nat. Struct. Mol. Biol.* 16:1279–85
169. Pearl LH, Prodromou C. 2006. Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu. Rev. Biochem.* 75:271–94

170. Shiau AK, Harris SF, Southworth DR, Agard DA. 2006. Structural analysis of *E. coli* Hsp90 reveals dramatic nucleotide-dependent conformational rearrangements. *Cell* 127:329–40
171. Ali MM, Roe SM, Vaughan CK, Meyer P, Panaretou B, et al. 2006. Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex. *Nature* 440:1013–17
172. Dollins DE, Warren JJ, Immormino RM, Gewirth DT. 2007. Structures of Grp94-nucleotide complexes reveal mechanistic differences between the Hsp90 chaperones. *Mol. Cell* 28:41–56
173. Harris SF, Shiau AK, Agard DA. 2004. The crystal structure of the carboxy-terminal dimerization domain of htpG, the *Escherichia coli* Hsp90, reveals a potential substrate binding site. *Structure* 12:1087–97
174. Roe SM, Ali MM, Meyer P, Vaughan CK, Panaretou B, et al. 2004. The mechanism of Hsp90 regulation by the protein kinase-specific cochaperone $p50^{cdc37}$. *Cell* 116:87–98
175. Vaughan CK, Gohlke U, Sobott F, Good VM, Ali MMU, et al. 2006. Structure of an Hsp90-Cdc37-Cdk4 complex. *Mol. Cell* 23:697–707
176. Southworth DR, Agard DA. 2011. Client-loading conformation of the Hsp90 molecular chaperone revealed in the cryo-EM structure of the human Hsp90:HOP complex. *Mol. Cell* 42:771–81
177. Krukenberg KA, Forster F, Rice LM, Sali A, Agard DA. 2008. Multiple conformations of *E. coli* Hsp90 in solution: insights into the conformational dynamics of Hsp90. *Structure* 16:755–65
178. Hainzl O, Lapina MC, Buchner J, Richter K. 2009. The charged linker region is an important regulator of Hsp90 function. *J. Biol. Chem.* 284:22559–67
179. Tsutsumi S, Mollapour M, Graf C, Lee C-T, Scroggins BT, et al. 2009. Hsp90 charged-linker truncation reverses the functional consequences of weakened hydrophobic contacts in the N domain. *Nat. Struct. Mol. Biol.* 16:1141–47
179a. Meyer P, Prodromou C, Hu B, Vaughan C, Roe SM, et al. Structural and functional analysis of the middle segment of hsp90: implications for ATP hydrolysis and client protein and cochaperone interactions. *Mol. Cell* 11:647–58
180. Retzlaff M, Hagn F, Mitschke L, Hessling M, Gugel F, et al. 2010. Asymmetric activation of the Hsp90 dimer by its cochaperone Aha1. *Mol. Cell* 37:344–54
181. Koulov AV, Lapointe P, Lu B, Razvi A, Coppinger J, et al. 2010. Biological and structural basis for Aha1 regulation of Hsp90 ATPase activity in maintaining proteostasis in the human disease cystic fibrosis. *Mol. Biol. Cell* 21:871–84
182. Li J, Soroka J, Buchner J. 2012. The Hsp90 chaperone machinery: conformational dynamics and regulation by co-chaperones. *Biochim. Biophys. Acta* 1823:624–35
183. Mollapour M, Tsutsumi S, Truman AW, Xu W, Vaughan CK, et al. 2011. Threonine 22 phosphorylation attenuates Hsp90 interaction with cochaperones and affects its chaperone activity. *Mol. Cell* 41:672–81
184. Soroka J, Wandinger SK, Mäusbacher N, Schreiber T, Richter K, et al. 2012. Conformational switching of the molecular chaperone Hsp90 via regulated phosphorylation. *Mol. Cell* 45:517–28
185. Hessling M, Richter K, Buchner J. 2009. Dissection of the ATP-induced conformational cycle of the molecular chaperone Hsp90. *Nat. Struct. Mol. Biol.* 16:287–93
186. Southworth DR, Agard DA. 2008. Species-dependent ensembles of conserved conformational states define the Hsp90 chaperone ATPase cycle. *Mol. Cell* 32:631–40
187. Mickler M, Hessling M, Ratzke C, Buchner J, Hugel T. 2009. The large conformational changes of Hsp90 are only weakly coupled to ATP hydrolysis. *Nat. Struct. Mol. Biol.* 16:281–86
188. Schmid AB, Lagleder S, Gräwert MA, Röhl A, Hagn F, et al. 2012. The architecture of functional modules in the Hsp90 co-chaperone Sti1/HOP. *EMBO J.* 31:1506–17
189. Richter K, Walter S, Buchner J. 2004. The co-chaperone Sba1 connects the ATPase reaction of Hsp90 to the progression of the chaperone cycle. *J. Mol. Biol.* 342:1403–13
190. Panaretou B, Siligardi G, Meyer P, Maloney A, Sullivan JK, et al. 2002. Activation of the ATPase activity of Hsp90 by the stress-regulated cochaperone Aha1. *Mol. Cell* 10:1307–18
191. Anckar J, Sistonen L. 2011. Regulation of HSF1 function in the heat stress response: implications in aging and disease. *Annu. Rev. Biochem.* 80:1089–115
192. Walter P, Ron D. 2011. The unfolded protein response: from stress pathway to homeostatic regulation. *Science* 334:1081–86

193. Haynes CM, Ron D. 2010. The mitochondrial UPR—protecting organelle protein homeostasis. *J. Cell Sci.* 123:3849–55
194. Kern A, Ackermann B, Clement AM, Duerk H, Behl C. 2010. HSF1-controlled and age-associated chaperone capacity in neurons and muscle cells of *C. elegans*. *PLoS ONE* 5:e8568
195. Gupta R, Kasturi P, Bracher A, Loew C, Zheng M, et al. 2011. Firefly luciferase mutants as sensors of proteome stress. *Nat. Methods* 8:879–84
196. Hershko A, Ciechanover A. 1998. The ubiquitin system. *Annu. Rev. Biochem.* 67:425–79
197. Finley D. 2009. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* 78:477–513
198. Kettern N, Dreiseidler M, Tawo R, Höhfeld J. 2010. Chaperone-assisted degradation: multiple paths to destruction. *Biol. Chem.* 391:481–89
199. Rubinsztein DC, Marino G, Kroemer G. 2011. Autophagy and aging. *Cell* 146:682–95
200. Arias E, Cuervo AM. 2011. Chaperone-mediated autophagy in protein quality control. *Curr. Opin. Cell Biol.* 23:184–89
201. Prakash S, Matouschek A. 2004. Protein unfolding in the cell. *Trends Biochem. Sci.* 29:593–600
202. Zhang MH, Windheim M, Roe SM, Peggie M, Cohen P, et al. 2005. Chaperoned ubiquitylation—crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex. *Mol. Cell* 20:525–38
203. Xu Z, Kohli E, Devlin KI, Bold M, Nix JC, Misra S. 2008. Interactions between the quality control ubiquitin ligase CHIP and ubiquitin conjugating enzymes. *BMC Struct. Biol.* 8:26
204. Gamerdinger M, Hajieva P, Kaya AM, Wolfrum U, Hartl FU, Behl C. 2009. Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *EMBO J.* 28:889–901
205. Komatsu M, Waguri S, Chiba T, Murata S, Iwata J, et al. 2006. Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature* 441:880–84
206. Kopito RR. 2000. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* 10:524–30
207. Kaganovich D, Kopito R, Frydman J. 2008. Misfolded proteins partition between two distinct quality control compartments. *Nature* 454:1088–95
208. Tyedmers J, Mogk A, Bukau B. 2010. Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell Biol.* 11:777–88
209. Specht S, Miller SBM, Mogk A, Bukau B. 2011. Hsp42 is required for sequestration of protein aggregates into deposition sites in *Saccharomyces cerevisiae*. *J. Cell Biol.* 195:617–29
210. Iwata A, Riley BE, Johnston JA, Kopito RR. 2005. HDAC6 and microtubules are required for autophagic degradation of aggregated huntingtin. *J. Biol. Chem.* 280:40282–92
211. Iwata A, Christianson JC, Bucci M, Ellerby LM, Nukina N, et al. 2005. Increased susceptibility of cytoplasmic over nuclear polyglutamine aggregates to autophagic degradation. *Proc. Natl. Acad. Sci. USA* 102:13135–40
212. Kon M, Cuervo AM. 2010. Chaperone-mediated autophagy in health and disease. *FEBS Lett.* 584:1399–404
213. Arndt V, Dick N, Tawo R, Dreiseidler M, Wenzel D, et al. 2010. Chaperone-assisted selective autophagy is essential for muscle maintenance. *Curr. Biol.* 20:143–48
214. Morley JF, Brignull HR, Weyers JJ, Morimoto RI. 2002. The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* 99:10417–22
215. Morley JF, Morimoto RI. 2004. Regulation of longevity in *Caenorhabditis elegans* by heat shock factor and molecular chaperones. *Mol. Biol. Cell* 15:657–64
216. Cohen E, Bieschke J, Perciavalle RM, Kelly JW, Dillin A. 2006. Opposing activities protect against age-onset proteotoxicity. *Science* 313:1604–10
217. Ben-Zvi A, Miller EA, Morimoto RI. 2009. Collapse of proteostasis represents an early molecular event in *Caenorhabditis elegans* aging. *Proc. Natl. Acad. Sci. USA* 106:14914–19
218. Cohen E, Paulsson JF, Blinder P, Burstyn-Cohen T, Du D, et al. 2009. Reduced IGF-1 signaling delays age-associated proteotoxicity in mice. *Cell* 139:1157–69
219. David DC, Ollikainen N, Trinidad JC, Cary MP, Burlingame AL, Kenyon C. 2010. Widespread protein aggregation as an inherent part of aging in *C. elegans*. *PLoS Biol.* 8:e1000450

220. Demontis F, Perrimon N. 2010. FOXO/4E-BP signaling in *Drosophila* muscles regulates organism-wide proteostasis during aging. *Cell* 143:813–25
221. Olzscha H, Schermann SM, Woerner AC, Pinkert S, Hecht MH, et al. 2011. Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* 144:67–78
222. Orr HT, Zoghbi HY. 2007. Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* 30:575–621
223. Gidalevitz T, Ben-Zvi A, Ho KH, Brignull H, Morimoto RI. 2006. Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* 311:1471–74
224. Hipp MS, Patel CN, Bersuker K, Riley BE, Kaiser SE, et al. 2012. Indirect inhibition of 26S proteasome activity in a cellular model of Huntington's disease. *J. Cell Biol.* 196:573–87
225. Bulawa CE, Connelly S, Devit M, Wang L, Weigel C, et al. 2012. Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *Proc. Natl. Acad. Sci. USA* 109:9629–34
226. Tsaytler P, Harding HP, Ron D, Bertolotti A. 2011. Selective inhibition of a regulatory subunit of protein phosphatase 1 restores proteostasis. *Science* 332:91–94
227. Nagai Y, Fujikake N, Popiel HA, Wada K. 2010. Induction of molecular chaperones as a therapeutic strategy for the polyglutamine diseases. *Curr. Pharm. Biotechnol.* 11:188–97
228. Westerheide SD, Morimoto RI. 2005. Heat shock response modulators as therapeutic tools for diseases of protein conformation. *J. Biol. Chem.* 280:33097–100
229. Mu TW, Ong DS, Wang YJ, Balch WE, Yates JR, et al. 2008. Chemical and biological approaches synergize to ameliorate protein-folding diseases. *Cell* 134:769–81
230. Calamini B, Silva MC, Madoux F, Hutt DM, Khanna S, et al. 2012. Small-molecule proteostasis regulators for protein conformational diseases. *Nat. Chem. Biol.* 8:185–96
231. Schaffar G, Breuer P, Boteva R, Behrends C, Tzvetkov N, et al. 2004. Cellular toxicity of polyglutamine expansion proteins: mechanism of transcription factor deactivation. *Mol. Cell* 15:95–105
232. Lotz GP, Legleiter J, Aron R, Mitchell EJ, Huang SY, et al. 2010. Hsp70 and Hsp40 functionally interact with soluble mutant huntingtin oligomers in a classic ATP-dependent reaction cycle. *J. Biol. Chem.* 285:38183–93
233. Auluck PK, Chan HY, Trojanowski JQ, Lee VM, Bonini NM. 2002. Chaperone suppression of α-synuclein toxicity in a *Drosophila* model for Parkinson's disease. *Science* 295:865–68
234. Hageman J, Rujano MA, van Waarde MAWH, Kakkar V, Dirks RP, et al. 2010. A DNAJB chaperone subfamily with HDAC-dependent activities suppresses toxic protein aggregation. *Mol. Cell* 37:355–69
235. Sarkar S, Ravikumar B, Floto RA, Rubinsztein DC. 2009. Rapamycin and mTOR-independent autophagy inducers ameliorate toxicity of polyglutamine-expanded huntingtin and related proteinopathies. *Cell Death Differ.* 16:46–56
236. Lee B-H, Lee MJ, Park S, Oh DC, Elsasser S, et al. 2010. Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* 467:179–84
237. Jahn TR, Radford SE. 2005. The yin and yang of protein folding. *FEBS J.* 272:5962–70

Annual Review of Biochemistry
Volume 82, 2013

Contents

Prefatory

Christian Raetz: Scientist and Friend Extraordinaire
William Dowhan, Hiroshi Nikaido, JoAnne Stubbe, John W. Kozarich, William T. Wickner, David W. Russell, Teresa A. Garrett, Kathryn Brozek, and Paul Modrich 1

Recent Advances in Biochemistry

Mechanisms for Initiating Cellular DNA Replication
Alessandro Costa, Iris V. Hood, and James M. Berger 25

The Chromatin Response to DNA Breaks: Leaving a Mark on Genome Integrity
Godelieve Smeenk and Haico van Attikum 55

Readout of Epigenetic Modifications
Dinshaw J. Patel and Zhanxin Wang 81

Flap Endonuclease 1
Lata Balakrishnan and Robert A. Bambara 119

New Mechanistic and Functional Insights into DNA Topoisomerases
Stefanie Hartman Chen, Nei-Li Chan, and Tao-shih Hsieh 139

Arrest Peptides: *Cis*-Acting Modulators of Translation
Koreaki Ito and Shinobu Chiba 171

Structural Basis of the Translational Elongation Cycle
Rebecca M. Voorhees and V. Ramakrishnan 203

CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea
Rotem Sorek, C. Martin Lawrence, and Blake Wiedenheft 237

Correlating Structure and Energetics in Protein-Ligand Interactions: Paradigms and Paradoxes
Stephen F. Martin and John H. Clements 267

Extracellular Chaperones and Proteostasis
Amy R. Wyatt, Justin J. Yerbury, Heath Ecroyd, and Mark R. Wilson 295

Molecular Chaperone Functions in Protein Folding and Proteostasis
Yujin E. Kim, Mark S. Hipp, Andreas Bracher, Manajit Hayer-Hartl, and F. Ulrich Hartl 323

Sumoylation: A Regulatory Protein Modification in Health and Disease
Annette Flotho and Frauke Melchior 357

Ubiquitin Ligases and Cell Cycle Control
Leonardo K. Teixeira and Steven I. Reed 387

Molecular Architecture and Assembly of the Eukaryotic Proteasome
Robert J. Tomko Jr. and Mark Hochstrasser 415

Design of Protein Catalysts
Donald Hilvert 447

Hydrogen Tunneling Links Protein Dynamics to Enzyme Catalysis
Judith P. Klinman and Amnon Kohen 471

Methylerythritol Phosphate Pathway of Isoprenoid Biosynthesis
Lishan Zhao, Wei-chen Chang, Youli Xiao, Hung-wen Liu, and Pinghua Liu 497

Posttranslational Biosynthesis of the Protein-Derived Cofactor Tryptophan Tryptophylquinone
Victor L. Davidson and Carrie M. Wilmot 531

Mitochondrial Complex I
Judy Hirst 551

Photosystem II: The Reaction Center of Oxygenic Photosynthesis
David J. Vinyard, Gennady M. Ananyev, and G. Charles Dismukes 577

The Voltage-Gated Calcium Channel Functions as the Molecular Switch of Synaptic Transmission
Daphne Atlas 607

Sphingosine-1-Phosphate and Its Receptors: Structure, Signaling, and Influence
Hugh Rosen, Raymond C. Stevens, Michael Hanson, Edward Roberts, Michael B.A. Oldstone 637

Membrane Fission Reactions of the Mammalian ESCRT Pathway
John McCullough, Leremy A. Colf, and Wesley I. Sundquist 663

Signal Recognition Particle: An Essential Protein-Targeting Machine
David Akopian, Kuang Shen, Xin Zhang, and Shu-ou Shan 693

Peroxisome Formation and Maintenance Are Dependent on the Endoplasmic Reticulum
Henk F. Tabak, Ineke Braakman, and Adabella van der Zand 723

Systemic Amyloidoses
Luis M. Blancas-Mejía and Marina Ramirez-Alvarado 745

Nanobodies: Natural Single-Domain Antibodies
Serge Muyldermans 775

Indexes

Cumulative Index of Contributing Authors, Volumes 78–82 799

Cumulative Index of Article Titles, Volumes 78–82 803

Errata

An online log of corrections to *Annual Review of Biochemistry* articles may be found at http://biochem.annualreviews.org/errata.shtml