SECTION 1

Molecular Chaperones and the Cell Stress Response

1

Chaperone Function: The Orthodox View

R. John Ellis

1.1. Introduction

The term molecular chaperone came into general use after the appearance of an article in *Nature* that suggested it was an appropriate phrase to describe a newly defined intracellular function – the ability of several unrelated protein families to assist the correct folding and assembly/disassembly of other proteins [1]. The identification of the chaperonin family of molecular chaperones in the following year [2] triggered a tidal wave of research in several laboratories aimed at unravelling how the GroEL/GroES chaperones, and later the DnaK/DnaJ chaperones, from *Escherichia coli* facilitate the folding of newly synthesised polypeptide chains and the refolding of denatured proteins. This wave continues to surge, with the result that much detailed information is available about the structure and function of those families of chaperone that assist protein folding [3].

It is now well established that a subset of proteins requires this chaperone function, not because chaperones provide steric information required for correct folding but because chaperones inhibit side reactions that would otherwise cause some of the chains to form non-functional aggregates. The number of different protein families described as chaperones is now more than 25 – some, but not all, of which are also stress proteins – and there is no slackening in the rate of discovery of new ones. The success of this wave of research has changed the paradigm of protein folding from the earlier view that it is a *spontaneous* self-assembly process to the current view that it is an *assisted* self-assembly process [4].

Is another paradigm shift in the offing? Other chapters in this volume discuss the evidence that some molecular chaperones may have extracellular roles as cell-cell signalling molecules in addition to their intracellular roles in protein folding. This view has not found general acceptance, partly because it is novel and partly because of the paucity of high-quality evidence compared with that

available in support of the protein folding paradigm. The purpose of this introductory chapter is to summarise the conventional view of chaperone function to provide a context for the cell-cell signalling hypothesis discussed in later chapters.

1.2. Origins

The current association of the term molecular chaperone with protein folding in the cytoplasm overlooks the fact that this term was used first to describe the properties of a nuclear protein in assisting the assembly of nucleosome cores from folded histone proteins and DNA [5]. This acidic protein is abundant in the soluble phase of the nuclei of eggs and oocytes of the amphibian *Xenopus laevis* and is thus called nucleoplasmin. These nuclei are unusual in that they contain large amounts of histone proteins stored in preparation for the rapid assembly of nucleosomes associated with the rapid replication of DNA triggered by fertilisation. Nucleosomes fail to conform to the principle of protein self-assembly established by the pioneering work of Anfinsen for refolding proteins and by Caspar and Klug for virus assembly. This principle is an important corollary of the Central Dogma of molecular biology and states that all the steric information necessary for a protein chain to reach its functional conformation is present in the amino acid sequence of the primary translation product. This principle also applies to the assembly of macromolecular complexes from more than one subunit.

In nucleosomes, histones are bound to DNA by electrostatic interactions; disruption of these requires high salt concentrations, but exposing mixtures of isolated DNA and histones to the salt concentrations found inside the nucleus results in the formation of insoluble aggregates rather than nucleosomes. Nucleoplasmin solves this aggregation problem by transiently binding its acidic groups to positively charged groups on the histones, thus lowering their overall surface charge and allowing the intrinsic self-assembly properties of the histones to predominate over the incorrect interactions favoured by the high density of opposite charges. Control experiments show that nucleoplasmin does not provide steric information essential for histones to bind correctly to DNA, nor is it a component of assembled nucleosomes. It is these two latter features that laid the foundation for our current general concept of the function of chaperones [6].

The term molecular chaperone was later extended to include an abundant chloroplast protein called the rubisco large subunit binding protein, which functions to keep newly synthesised rubisco large subunits from aggregating until they assemble into the rubisco holoenzyme [7]. These subunits are notoriously

prone to aggregation, not because of electrostatic interactions but because they expose highly hydrophobic surfaces to the aqueous environment. For a while the term was restricted to the two proteins that assist the assembly of amphibian nucleosomes and chloroplast rubisco. Its modern usage started when the author suggested that the term could be usefully extended to describe the function of a larger range of proteins that were postulated to assist folding and assembly/disassembly reactions in a wide range of cellular processes [1].

1.3. The general concept of chaperone function

The suggestions made in the first comprehensive description of the chaperone function have so far stood the test of time [8]. Molecular chaperones are defined as being a large and diverse group of proteins that share the property of assisting the non-covalent assembly/disassembly of other macromolecular structures but which are not permanent components of these structures when these are performing their normal biological functions. Assembly is used here in a broad sense and includes several universal intracellular processes: the folding of nascent polypeptide chains both during their synthesis and after release from ribosomes, the unfolding and refolding of polypeptides during their transfer across membranes, and the association of polypeptides with one another and with other macromolecules to form oligomeric complexes.

Molecular chaperones are also involved in macromolecular *dis*assembly processes, such as the partial unfolding and dissociation of subunits when some proteins carry out their normal functions, and the re-solubilisation and/or degradation of proteins partially denatured and/or aggregated by mutation or by exposure to environmental stresses, such as high temperatures and oxidative conditions. Some, but not all, chaperones are also stress or heat shock proteins as the requirement for chaperone function increases under stress conditions that cause proteins to unfold and aggregate. Conversely, some, but not all, stress proteins are molecular chaperones.

It is important to note that this definition is functional, not structural, and it contains no constraints on the mechanisms by which different chaperones may act; this is the reason for the use of the imprecise term ‘assist’. Thus, molecular chaperones are defined neither by a common mechanism nor by sequence similarity. Only two criteria need be satisfied to designate a macromolecule a molecular chaperone. Firstly, it must in some sense assist the non-covalent assembly/disassembly of other macromolecular structures, the mechanism being irrelevant, and secondly, it must not be a component of these structures when they are performing their normal biological functions. In all cases studied so far, chaperones bind non-covalently to regions of macromolecules that are

inaccessible when these structures are correctly assembled and functioning but that are accessible at other times.

The term non-covalent is used in this definition to exclude those proteins that catalyse co- or post-translational covalent modifications. These are often important for protein assembly, but are distinct from the proteins being considered here. Protein disulphide isomerase may appear to be an exception, but it is not. It is both a covalent modification enzyme and a molecular chaperone, but these activities lie in different parts of the molecule [9] and can be functionally separated by mutation. Other examples include peptidyl-prolyl isomerase, which possesses both enzymatic and chaperone activities in different regions of the molecule, and the α-crystallins, which in the lens of the eye combine two essential functions in the same molecule, contributing to the transparency and refractive index required for vision as well as to the chaperone function, which combats the loss of transparency as the protein chains aggregate with increasing age. The proteasome particle has a chaperone-like activity involved in unfolding proteins prior to their proteolysis. Thus, in principle, there is no reason why molecular chaperones should not possess additional functions, and the possibility that many possess cell-cell signalling functions is the central postulate argued in the other chapters of this volume.

The number of distinct chaperone families continues to rise, and examples occur in all types of cell and in most intracellular compartments. The families are defined on the basis that members within each family have high sequence similarity, whereas members in different families do not. Table 1.1 presents an incomplete list of proteins described as chaperones; however, it must be emphasised that in many cases this description rests on *in vitro* data only and needs confirmation by *in vivo* methods. There is evidence that some chaperones cooperate with each other in defined reaction sequences, but this, along with many other aspects of 'chaperonology', is beyond the scope of this chapter.

1.4. Common misconceptions

As with any new field, misconceptions abound. A common error is to use the term 'chaperonin' synonomously with the term 'chaperone', but it should be noted that the chaperonins are just one particular family of chaperone – i.e., the family that contains GroEL, Hsp60, and tailless complex polypeptide-1 (TCP-1) (see Chapters 5 and 6 for more details on chaperonins). The occasional use of the non-sense term 'molecular chaperonin' in some respectable journals suggests that some people use these terms casually without reference to either their meaning or their history. It should be obvious that the word 'molecular' is used to qualify 'chaperone' because in common usage 'chaperone' refers to a

Table 1.1. Proteins described as molecular chaperones

Family	Proposed roles
Non-steric chaperones	
Nucleoplasmins/nucleophosmins	Nucleosome and ribosome assembly/disassembly
Chaperonins	Folding of newly synthesised and denatured polypeptides
Hsp27/28	Prevention of stress-induced aggregation by adsorbing unfolded chains
Hsp40	Protein folding and transport, oligomer disassembly
Hsp47	Pro-collagen folding in the endoplasmic reticulum (ER)
Hsp70	Protein folding and transport, oligomer disassembly
Hsp90	Cell cycle, hormone activation, signal transduction
Hsp100	Dissolution of insoluble protein aggregates
Calnexin/calreticulin	Folding of glycoproteins in ER
SecB protein	Protein transport in bacteria
Lim protein	Folding of bacterial lipase
Syc protein	Secretion of toxic YOP proteins by bacteria
Protein disulphide isomerase	Prevention of misfolding in ER
ExbB proteins (may be structural rather than chaperones)	Folding of TonB protein in bacteria
Ubiquinated ribosomal proteins	Ribosome assembly in yeast
NAC complex	Folding of nascent proteins
Signal recognition particle	Arrest of translation and targeting to ER membrane
Trigger factor	Folding of nascent polypeptides in bacteria
Prefoldin	Cooperation with chaperonins in folding of newly synthesised polypeptides in *Archaea* and the eukaryotic cytosol
Tim9/Tim10 complex	Prevention of aggregation of hydrophobic proteins during import across mitochondrial intermembrane space
23S Ribosomal RNA	Folding of nascent polypeptides
PrsA protein	Secretion of proteins by *Bacillus subtilis*
Clusterin	Extracellular animal chaperone
Phosphatidylethanolamine	Folding of lactose permease
RNA binding proteins	Folding of RNA
P45	Protection against denaturation in halophilic *Archaea*
Steric chaperones	
PapD proteins	Assembly of bacterial pili
Propeptides (Class I)	Folding of some proteases

person. The term 'molecular chaperonin' is therefore as non-sensical as the term 'molecular immunoglobulin'.

Another common misconception is that molecular chaperones are necessarily promiscuous – i.e., that each assists the assembly of many different types of polypeptide chain. This is true for the Hsp70, Hsp40 and GroE chaperonin

families but is not true for Hsp90, PapD, Hsp47, Lim, Syc, ExbB, PrtM/PrsA and prosequences, which are specific for their substrates. Similarly, it is not a universal property of chaperones that they hydrolyse ATP; Hsp100, Hsp90, Hsp70 and the chaperonins hydrolyse ATP, whereas trigger factor, Hsp40, prefoldin, calnexin, protein disulfide isomerase and papD do not. It is not even necessary that the chaperone function resides in molecules separate from their substrates. Thus, some pro-sequences are required for the correct folding of the remainder of the molecule but are then removed [10]. Another example of such intramolecular chaperones is the terminal ubiquitin residues of three ribosomal proteins in yeast; these residues promote the assembly of these proteins into the ribosome but are then removed, thus fulfilling the criteria suggested earlier for the chaperone function [11].

The term 'chemical chaperone' has been proposed to describe small molecules such as glycerol, dimethylsulfoxide and trimethylamine N-oxide that act as protein stabilising agents [12]. This terminology is unfortunate because proteins are also chemicals. However, its usage persists.

Finally, experience suggests that the distinction between molecular chaperones and stress proteins cannot be restated too often. The often-made interpretation, that because a protein accumulates after stress it must be a molecular chaperone, is incorrect, as is the belief that all molecular chaperones are stress proteins. For example, many heat shock proteins are ubiquitin-conjugating enzymes, while the cytosolic chaperonin of eukaryotic cells is not a stress protein.

1.5. Why do molecular chaperones exist?

Given that most denatured proteins that have been examined can refold into their functional conformations on removal of the denaturing agent *in vitro*, the question arises as to why molecular chaperones exist at all. Current evidence suggests that, with two possible exceptions [10], chaperones do not provide steric information for proteins to assemble correctly; rather they either prevent or reverse aggregation processes that would otherwise reduce the yield of functional molecules. Aggregation results because some proteins fold and unfold via intermediate states that expose some interactive surfaces (either charged or hydrophobic) to the environment. In aqueous environments hydrophobic surfaces stick together, while charged surfaces bind to ones bearing the opposite charge, a problem acute in the nucleus where negatively charged nucleic acids are bound to positively charged proteins. Thus, the existence of molecular chaperones does not cast doubt on the validity of the self-assembly principle. Rather, chaperones are required because, to operate efficiently under intracellular

conditions, self-assembly needs assistance to avoid unproductive side reactions. This is why the term molecular chaperone is not an example of academic whimsy, but a precise description, because the role of the human chaperone is to improve the efficiency of 'assembly' processes between people without providing the steric information for these processes.

Protein aggregation has long been observed to occur during the *in vitro* refolding of many pure denatured proteins in dilute buffer solutions, but only recently has it been appreciated that the high degree of macromolecular crowding that characterises the intracellular environment makes the aggregation problem much more severe *in vivo*. Although the total concentrations of macromolecules inside cells are in the range 200–400 mg/ml, the properties of the isolated macromolecules are commonly studied *in vitro* at much lower concentrations in uncrowded buffers. The large thermodynamic effects of the high total concentrations of macromolecules inside cells are not generally appreciated and include increasing the association constants of protein aggregation reactions by one to two orders of magnitude [13]. Aggregation is a specific process involving identical or very similar chains and is driven by the interaction of both hydrophobic side chains and main-chain atoms in segments of unstructured backbone that are transiently exposed on the surface of partly folded chains; it is thus a high-order process that increases in rate as the concentration of similar chains or the temperature is raised. Refolding experiments suggest that large multi-domain proteins suffer from aggregation to a greater degree than small single-domain proteins because they fold more slowly via partly folded intermediate states. Thus proteins differ greatly in their propensity to aggregate, and it is likely that chaperones have evolved to combat this tendency of a particular subset of proteins.

These considerations can be reduced to a simple unifying principle. *All cells need a chaperone function to both prevent and reverse incorrect interactions that may occur when potentially interactive surfaces are exposed to the intracellular environment. Such surfaces occur on nascent and newly synthesised unfolded polypeptide chains, on mature proteins unfolded by stress or degradative mechanisms, and on folded proteins in near-native conformation*.

Thus, it is as valid to talk about the chaperone function as it is to talk about the transport function or the defence function of other proteins. We can view the chaperone function as one of the universal mechanisms that enable the crowded state of the cellular interior to be compatible with life.

The best understood chaperones are those involved in the folding of newly synthesised polypeptide chains in *E. coli*. The next section summarises what is known about their mechanisms of action.

1.6. Chaperones involved in *de novo* protein folding

The folding of newly synthesised proteins inside cells differs from the refolding of denatured proteins *in vitro* in two respects [14]. Firstly, protein chains fold inside cells in highly crowded macromolecular environments that favour aggregation. Secondly, protein chains are made vectorially inside cells at a rate slower than the rate of folding. It takes about 20 seconds for a cell of *E. coli* to synthesise a chain of 400 residues at 37 °C; however, *in vitro*, many denatured proteins will refold completely well within this time. Thus there is the possibility that the elongating nascent chain will either misfold because it is incomplete or aggregate with identical elongating chains on the same polysome. It is important to realise that misfolding is conceptually distinct from aggregation. Misfolding can be defined as the chain reaching a partly folded conformation from which it is unable to reach the final functional conformation on a biologically relevant time scale. Misfolded chains may or may not bind to one another to form non-functional aggregates that may be as small as a dimer or large enough to be insoluble. Thus all aggregates are, by definition, misfolded, but to what extent misfolded, but unaggregated, chains occur in cells is unclear.

Molecular chaperones assist the folding of both nascent chains bound to ribosomes and newly synthesised chains released from ribosomes – i.e., in both co-translational and post-translational modes. The chaperones working in these co-translational and post-translational modes are distinct and can be usefully termed small and large chaperones, respectively, because this is a case where size is important for function [15]. Small chaperones are less than 200 kDa in size and include trigger factor, nascent chain-associated complex, prefoldin, the Hsp70 and Hsp40 families and their associated co-chaperones. Co-chaperones are defined as proteins that bind to chaperones to modulate their activity; they may or may not also be chaperones in their own right. Large chaperones are more than 800 kDa in size and include the thermosome in *Archaea*, GroE proteins in *Eubacteria* and the eukaryotic organelles evolutionarily derived from them, and the TCP-1 or TRiC complexes and associated co-chaperones in the cytosol of *Eukarya*. The large chaperones are evolutionarily related and are collectively referred to as the chaperonins. There are no large chaperones in the endoplasmic reticulum lumen of eukaryotic cells, but small chaperones such as BiP (an Hsp70 homologue), calnexin, calreticulin, and protein disulphide isomerase that assist the folding of chains transported into the lumen after synthesis in the cytosol are present. Table 1.2 lists some of the chaperones that assist protein folding.

Table 1.2. Chaperones that assist protein folding

	Other names		
Family	Eukaryotes	Prokaryotes	Functions
Hsp100	Hsp104, 78	ClpA/B/X	Disassembly of oligomers and aggregates
Hsp90	Hsp82, Hsp83, Grp94	HtpG	Regulate assembly of steroid receptors and signal transduction proteins
Hsp70	Hsc70, Ssal-4, Ssb1-2, BiP, Grp75	DnaK, Hsc66, Absent from many *Archaea*	Prevent aggregation of unfolded protein chains
Chaperonins	Hsp60, TRiC, CCT, TCP-1, rubisco subunit binding protein	GroEL, GroES	Sequester partly folded chains inside central cage to allow completion of folding in absence of other folding chains
Hsp40	Ydj1, Sis1, Sec63p, auxilin, zuotin, Hdj2	DnaJ	Stimulate ATPase activity of Hsp70
Prefoldin	GimC	Absent from *Bacteria*, present in *Archaea*	Prevent aggregation of unfolded protein chains
Trigger factor	Absent from *Eukarya*	Present	Bind to nascent chains as they emerge from ribosome
Calnexin, calreticulin	Present	Absent from prokaryotes	Bind to partly folded glycoproteins; located in ER membrane and lumen, respectively
Nascent-chain associated complex (NAC)	Present	Absent from prokaryotes	Bind to nascent chains as they emerge from ribosome
PapD	Absent from *Eukarya*	Present in some	Prevent aggregation of subunits of pili

1.6.1. Small chaperones

Small chaperones bind transiently to small hydrophobic regions (typically seven or eight residues long) on both nascent and completed, newly synthesised chains and thus prevent aggregation both during and after chain elongation by shielding these regions from one another (Figure 1.1) [3]. Trigger factor (48 kDa) is the