Assembly chaperones: a perspective

R. John Ellis

School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK

The historical origins and current interpretation of the molecular chaperone concept are presented, with the emphasis on the distinction between folding chaperones and assembly chaperones. Definitions of some basic terms in this field are offered and misconceptions pointed out. Two examples of assembly chaperone are discussed in more detail: the role of numerous histone chaperones in fundamental nuclear processes and the co-operation of assembly chaperones with folding chaperones in the production of the world’s most important enzyme.

1. Introduction

Perusal of the extensive literature about molecular chaperones might lead you to conclude that their primary role is to reduce the tendency of newly synthesized and stress-denatured protein chains to misfold and aggregate together by hydrophobic interactions in the cytosol and endoplasmic reticulum. Such an emphasis on the role of chaperones in protein folding ignores the fact that the term ‘molecular chaperone’ was coined originally to describe a protein that is required for the assembly of folded histone proteins into oligomeric structures by electrostatic interactions with DNA inside the nucleus. Thus the undoubted importance of chaperones in assisting protein folding [1] has obscured their equally vital roles in protein assembly processes. Confusion also persists about the precise relation between stress proteins and molecular chaperones—are all chaperones also stress proteins, and are all stress proteins also chaperones? How should we define the term ‘molecular chaperone’ today? In this review I offer an account of the historical origins of the chaperone concept, suggest definitions of the terms used in this field and present a list of some assembly chaperones. My aim is to offer a perspective for subsequent articles that describe the roles of selected chaperones in protein assembly.

2. Origin

The term ‘molecular chaperone’ appeared first in 1978 in a paper from the laboratory of Ron Laskey to describe a nuclear protein required for the correct assembly of nucleosomes from histones and DNA in extracts of amphibian eggs [2]. Nucleosomes are oligomers of eight basic histone monomers bound by electrostatic charge interactions to negatively charged eukaryotic DNA. Isolated nucleosomes can be dissociated into their histone and DNA components by exposure to high salt concentrations. The principle of protein self-assembly, established by the pioneer work of Anfinsen [3] and Caspar and Klug [4], states that all the information required for protein chains to fold and assemble correctly with other proteins and/or nucleic acids is located in the primary structure of those chains. Thus this principle predicts that a mixture of histones and DNA placed in physiological salt concentrations should result in the spontaneous assembly of nucleosomes. But when Laskey tried this experiment with nucleosomes from Xenopus eggs, it was a spectacular failure; instead of nucleosomes, an insoluble aggregate formed [2].

Experiments showed that addition of small amounts of Xenopus egg homogenate prevents aggregate formation and allows nucleosome assembly. The active factor was purified, found to be an abundant nuclear acidic protein and named nucleoplasm. This protein binds electrostatically to folded histones, thereby reducing their basic charge, and permitting their inherent
self-assembly properties to predominate over their tendency to aggregate non-specifically with negatively charged DNA [2].

Two characteristics of the mode of action of nucleoplasmin were important in the subsequent development of the general concept of chaperone function that prevails today. First, nucleoplasmin is required only transiently for nucleosome assembly— it is not a component of the 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 essential for nucleosome assembly, but to reduce the strong basic charge of the histones. The molecular details of how nucleoplasmin works are still being unravelled [5,6], but it is clear that it does not involve the formation or breakage of covalent bonds. The binding of nucleoplasmin can thus only be detected by the use of non-denaturing techniques in the early stages of nucleosome assembly. Later work revealed an additional role of nucleoplasmin in decondensing sperm chromatin on fertilization of the egg, resulting in the replacement of the protamine proteins of the sperm nucleosomes by the histone proteins of the zygote. These properties of nucleoplasmin led to the suggestion that "the role of the protein we have purified is that of a "molecular chaperone" which prevents incorrect interactions between histones and DNA" (p. 419 [2]).

Thus, the term 'molecular chaperone' was coined, not as a metaphor or a whim, but because the properties of nucleoplasmin are a precise molecular analogy of the role of human chaperones. The traditional role of human chaperones is to prevent incorrect interactions between pairs of human beings, without either providing the steric information necessary for their correct interaction or being present during married life.

3. Generalization

The extension of the concept of a chaperone function to other proteins had to wait until 1985. In that year I organized a Royal Society Discussion meeting on the chloroplast protein called Rubisco. This acronym stands for ribulose biphosphate carboxylase-oxygenase, the enzyme that brings carbon dioxide into organic combination during photosynthesis. It is thus arguably the world’s most important enzyme. It is also the world’s most abundant enzyme, because it has a very low turnover number; Rubisco is limiting to plant growth under most environmental conditions so the plant has to make large amounts to compete successfully [7]. It is also the world’s most inefficient enzyme, because it is not saturated at the current atmospheric level of carbon dioxide. Moreover, Rubisco catalyses an oxygenase reaction that leads to a loss of carbon dioxide during photosynthesis via the process of photorespiration. None of these attributes mattered when the enzyme first evolved, because at that time there was no oxygen in the atmosphere, and the level of carbon dioxide was much higher than it is today. Thus Rubisco has long been the target of attempts to improve its properties for today’s environment by genetic engineering [8]. All such attempts have failed so far, one reason being the complex involvement of chaperones in both its folding and assembly. Thus it has not proved possible as yet to express chloroplast Rubisco in active form in Escherichia coli [9].

Chloroplast Rubisco is a heterooligomer of eight catalytic large subunits, synthesized inside the chloroplast, bound to eight structural small subunits, made in the cytosol. Rubisco is one of several proteins that fail to refold and assemble into an active enzyme when diluted out from a denaturant in a classic Anfinsen refolding experiment [10]. This failure is because of the tendency of unfolded large subunits to aggregate with one another and become insoluble. So, I was surprised when work in my laboratory in 1980 revealed that the unassembled large subunits synthesized inside isolated, intact chloroplasts are soluble. We established that this is because they are bound non-covalently and transiently to another protein before assembling into the Rubisco holoenzyme. We called this protein the Rubisco large subunit-binding protein, and speculated that the complex of this protein with newly synthesized large subunits is an obligatory intermediate in the assembly of Rubisco [11]. This speculation was controversial because it implied that protein self-assembly is not always a spontaneous process, but in some cases required the transient involvement of pre-existing proteins.

I came across Laskey’s nucleoplasmin paper in 1985, and realized that the role of the Rubisco large subunit-binding protein could be thought of as similar to that of nucleoplasmin, that is, preventing aggregation by transiently masking the interactive surfaces involved. For unfolded Rubisco large subunits, these surfaces are hydrophobic in nature, whereas for the folded histones they are charged, but the principle is the same. At a Royal Society Rubisco meeting in 1985, the proceedings of which appeared in the following year [12], I made the suggestion that the Rubisco large subunit-binding protein could be regarded as the second example of a molecular chaperone.

I thought initially that nucleoplasmin and the Rubisco large subunit-binding protein were special cases, evolved to deal with proteins where aggregation is a particular problem. But in 1987, a speculative paper by Hugh Pelham prompted me to extend the chaperone idea to a much wider range of proteins [13]. This paper makes no reference to nucleoplasmin and the Rubisco large subunit-binding protein, nor uses the term ‘molecular chaperone’, but instead discusses the possible roles of the heat-shock protein (HSP) 70 and 90 families in a range of assembly and disassembly processes. Pelham proposed that heat shock proteins function in unstressed cells by binding to exposed hydrophobic surfaces, but are required in increased amounts in stressed cells to unscramble aggregates, and to prevent further aggregation.

It occurred to me that all these ideas could be brought together under the umbrella of a more fundamental chaperone concept. I presented this idea at the NATO Advanced Study Meeting in Copenhagen in 1987, where the Nature representative persuaded me to write a News and Views article. This article appeared later that year with the 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 polypeptides and their assembly into oligomeric structures occur correctly’ (p. 378 [14]). It was this article that sparked the widespread use of the term ‘molecular chaperone’ in the literature.

In 1988, my laboratory, together with that of the phage biochemist Roger Hendrix, published the finding that the Rubisco large subunit-binding protein is 50 per cent identical with the GroEL protein of E. coli [15]. GroEL was identified in several laboratories in the 1970s as a bacterial protein required for phages T4 and lambda to replicate. Its precise role was unclear, but thought to involve the assembly of the phage head because...
a mutation in GroEL results in the head proteins forming an insoluble aggregate [16], but what its role in uninfected cells might be was not addressed. In 1988 it was also reported that antiseras to GroEL identified a protein inside mitochondria from various plant and animal species [17].

My postdoc, Sean Hemmingsen and I realized that there was new evidence from both bacteria and chloroplasts that linked the involvement of highly similar pre-existing proteins in the assembly of newly synthesized proteins in a manner that fitted the general concept of molecular chaperones that I had proposed in 1987. Sean named these chloroplast and bacterial proteins the ‘chaperonins’, and he and I wrote the first comprehensive proposal for the existence of a widespread class of cellular proteins that acted as molecular chaperones [18]. Detailed histories of the discovery of the molecular chaperone function have been published [19,20].

4. The molecular chaperone concept today

Before discussing how molecular chaperones are currently viewed, I should like to offer some definitions of terms used in this field.

(a) Definitions

Protein folding is the collapse of an extended primary translation product into a stable compact monomer. Monomers possess primary, secondary and tertiary structure, but not quaternary structure.

Protein assembly is the binding of monomers to one another to produce a functional oligomer. Oligomers possess quaternary structure.

The folding of a given polypeptide chain is characterized by the formation of a stable fold specific to that sequence, whereas assembly is characterized by the association of two or more folded monomers into a biologically functional oligomer. The distinction between folding and assembly is not absolute but quantitative, because in both processes there are changes in the conformation of protein chains. These changes in conformation are usually smaller during assembly than during folding. Note also that, while folding is defined entirely in structural terms, the definition of assembly contains a biological criterion in addition to a chemical one. The word ‘functional’ is used to distinguish these oligomers from non-functional assemblies, and to make this explicit, non-functional oligomers are commonly called aggregates or misassemblies.

This brings me to a fundamental point about the difference between chemistry and biology. What is important in chemistry is structure, because structure determines the properties of molecules. But what is important in biology is not structure per se, but the function enabled by that structure. This is because it is function that is selected for in evolution—structure does not matter, provided it enables functions that help the organism to survive in a competitive world. For example, a variety of proteins called crystallins are found in the lens of eyes, where they all provide required properties of transparency and refractive index. But some crystallins are metabolic enzymes and these vary between species. Lactate dehydrogenase is a crystallin found in avian and repilian eyes, but in the guinea pig it is replaced by alcohol dehydrogenase [21]. So, the basic argument is that, because proteins are the products of natural selection, we should include biological criteria as well as chemical criteria in our definitions.

Protein misfolding is the formation of a conformation that cannot reach the native conformation on a biologically relevant timescale. The term ‘native’ refers to the conformation of the purified functional protein. The term ‘non-native’ is defined as a conformation that is unable to reach the functional conformation on any timescale.

This inclusion of a biological criterion distinguishes this definition from others that are in use, such as the suggestion that a misfolded conformation is one that has to unfold to some extent before it can reach the native conformation [22]. I propose this definition because the prefix ‘mis’ meaning ‘wrong’, when applied to proteins, can have a meaning only in a biological context—a misfolded structure must have a biological consequence. Note that, on my definition, ‘misfolded’ does not mean the same as ‘non-native’ because a misfolded state could be a native state, i.e. a normal intermediate on its way to a functional state, but too slowly for biological needs. Moreover, this definition aligns the meaning of the term ‘misfolding’ with that of ‘misassembly’ in a pleasing symmetry.

Protein misassembly is the association of two or more polypeptide chains to form non-functional oligomers. A widely used term to describe protein misassemblies is ‘aggregates’, especially to describe so-called amyloid structures, such as those associated with neurodegenerative diseases. A minor problem with this term is that functional amyloids have been discovered in both prokaryocytes and eukaryotes [23], so it is preferable to use the term ‘misassembly’. Note that misassembled states are always misfolded, by definition.

Molecular chaperones are a large and diverse group of proteins that share the property of assisting non-covalent folding and unfolding, and the assembly and disassembly, of other macromolecular structures, but are not permanent components of these structures when they are performing their normal biological functions [24]. Chaperone substrates include RNA as well as proteins, because nucleic acids have less information content than proteins and so suffer more often from misfolding and misassembly. The reason for this difference in information content is that there are only 4 bases but 20 amino acids [25].

It is important to note that this definition is functional, and not structural, and contains no constraints on how different molecular chaperones may act. The qualification ‘non-covalent’ is used to exclude those proteins that catalyse co- and post-translational modifications. Protein disulphide isomerase might appear to be an exception, but it is both a covalent modification enzyme and a molecular chaperone [26]. There is no reason in principle why molecular chaperones should not possess additional functions, such as cell–cell signalling [27]. It is for this reason that I suggest that it is more useful to think of a molecular chaperone as a function, rather than a molecule; on this basis, a chaperone function can be a property of a molecule that may have additional and quite different functions.

Chaperone function is to prevent and/or reverse incorrect interactions that may result when potentially interactive surfaces are exposed to the crowded intracellular or extracellular environments. These surfaces occur on nascent and newly synthesized protein and RNA chains, on mature proteins unfolded by environmental stress or covalent modification, and on folded proteins in native and near-native conformations. Incorrect interactions are defined as those that result in biologically
non-functional products; such products are often cytotoxic, e.g. those implicated in neurodegenerative disease.

**Cochaperones** are proteins that modulate chaperone activity. Some cochaperones are also chaperones, e.g. HSP40. There is an expanding literature on cochaperones, especially those that modulate the functions of HSP70 and HSP90 proteins. Cochaperones modulate the activity of their cognate chaperones in a variety of different ways [28].

**Table 1**

<table>
<thead>
<tr>
<th>Chaperone Name</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>GroEL</td>
<td>Folding</td>
</tr>
<tr>
<td>HSP70</td>
<td>Folding</td>
</tr>
<tr>
<td>HSP90</td>
<td>Folding</td>
</tr>
</tbody>
</table>

**Misconceptions**

Like many fields, the chaperone field is noted for the persistence of various myths, perhaps because people come across this term in casual use at conferences without researching its origins. Examples of misconceptions found in the literature include the following.

1. **All chaperones assist protein folding.** The falsity of the belief that ‘all chaperones assist protein folding’ is demonstrated, not only by the origins of the chaperone concept, but also by the subject matter of this theme issue.

2. **Chaperones fold proteins.** The related claim that ‘chaperones fold proteins’ could be taken to imply that some chaperones provide steric information necessary for some proteins to fold. If this were the case, the principle of protein self-assembly would be falsified. But what the chaperone concept proposed from its inception is that the principle of spontaneous self-assembly should be replaced by the principle of assisted self-assembly [18]. Thus the principle of self-assembly is retained, but modified to include the need for chaperones that reduce unproductive side reactions, particularly aggregation.

3. **All chaperones are promiscuous for their substrates.** The initial work on molecular chaperones was triggered by the identification of the chaperonins, and so concentrated on the mechanism of action of the chaperonin called GroEL, together with the chaperone HSP70, both found in *E. coli* [1]. These chaperones assist the folding of many different, newly synthesized protein chains, so the belief arose that chaperones are necessarily promiscuous. Now that over 100 families of chaperones have been identified, it is clear that this is not the case for many of them, such as PapD, HSP47, Lin Syc, ExbB, PrtM/PrsA and proselves. For example, PapD is a chaperone found in the periplasmic space of *E. coli*, where it is essential for the correct assembly of pili. The gene for PapD can be deleted without altering the viability of the cell—it just cannot make pili any more [29]. But GroEL is essential for the survival of *E. coli* under all conditions because it assists the folding of over 80 different proteins, including some essential ones.

4. **All chaperones hydrolyse ATP.** The initial emphasis on GroEL and HSP70 also led to the erroneous view that chaperones necessarily hydrolyse ATP. Many chaperones, such as nucleoplasmin, PapD, calnexin and protein disulphide isomerase, do not require ATP to function.

5. **Molecular chaperonins.** The occasional use of the term ‘molecular chaperonin’ in some journals suggests that some people use these terms casually without reference to either their meaning or their origin. The term ‘chaperonin’ is also sometimes wrongly used as synonymous with ‘chaperone’, but the chaperonins are just one particular family of chaperone—the family that contains GroEL. Families of chaperone are defined by sequence similarity, i.e. each family contains proteins with similar sequences.

6. **Chemical chaperones.** The term ‘chemical chaperone’ is used to describe small molecules such as glycerol, dimethylsulphoxide and trimethylamine N-oxide that act as protein-stabilizing agents. This terminology is unfortunate, and can confuse students who sometimes ask me whether that means that proteins are not chemicals! The term should be replaced by ‘pharmacological chaperone’ or ‘kosmotrope’, the term used by physical chemists to describe small molecules that stabilize proteins. Kosmotrope is the opposite of ‘chaotrope’, used to describe molecules that unfold proteins.

7. **Chaperones are another name for heat-shock proteins.** It is a mistake to think that chaperones such as HSP70 and HSP90 are present only under stress conditions. The first papers to present the chaperone concept made no suggestion that they were necessarily stress (or heat-shock) proteins, although some clearly were [14,15,18]. It was argued instead that some chaperones are stress proteins, because the need for the chaperone function increases as proteins unfold under stress conditions. Subsequent reviews emphasized that only a subset of chaperones are stress proteins, and thus that chaperones and stress proteins are overlapping sets, not identical sets [30,31].

Despite this, the idea persists. For example, a recent survey of 300 articles listed in PubMed in 2009 found that one-third of them stated that molecular chaperones are stress proteins and that stress proteins are molecular chaperones. A meta-analysis of the relations between the two sets revealed that the majority of chaperone genes (66% for humans and 72% for Arabidopsis) are not induced by heat [32]. These figures are probably too small, because this analysis considered only well-studied chaperone families, and not many other families of chaperone, including the growing number of assembly chaperones, the subject of this theme issue. So, most chaperones are not stress proteins, and most stress proteins are not chaperones.

## 5. Assembly chaperones

Because changes in protein conformation occur both when new protein chains are emerging from the ribosome and collapsing into monomers, and during the subsequent stages of oligomer assembly, the distinction between folding chaperones and assembly chaperones is not absolute. For example, the trigger factor is a folding chaperone that binds to many newly synthesized chains as they emerge from the exit tunnel of the prokaryotic ribosome, while nucleoplasmin is an assembly chaperone that binds to folded histone proteins before they bind to DNA. Between these two extreme examples, many of the intermediate stages in the formation of functional oligomers also require chaperonin to minimize unproductive side reactions. The same observation applies to mature proteins destined either for degradation or for refolding after stress-induced denaturation. The situation is complicated further by the fact that some members of the same protein family, as defined by sequence similarity, act as folding chaperones, while other members act as assembly chaperones.

I am unaware of any comprehensive list of assembly chaperones, but table 1 lists some of the best-studied examples. I shall discuss two of these in more detail.
Table 1. Examples of assembly chaperone.

<table>
<thead>
<tr>
<th>oligomer</th>
<th>chaperone(s)</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>nucleosome</td>
<td>nucleoplasmmin,</td>
<td>[33,34]</td>
</tr>
<tr>
<td></td>
<td>nucleophosmin, Asf1,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAF-1 etc.</td>
<td></td>
</tr>
<tr>
<td>Rubisco</td>
<td>RbcX</td>
<td>[35]</td>
</tr>
<tr>
<td>proteasome</td>
<td>PAC1/2/3, hLmp1</td>
<td>[36]</td>
</tr>
<tr>
<td>pilus</td>
<td>PapD</td>
<td>[29]</td>
</tr>
<tr>
<td>hemoglobin</td>
<td>AHSP</td>
<td>[37]</td>
</tr>
<tr>
<td>nuclear hormone</td>
<td>HSP90</td>
<td>[38]</td>
</tr>
<tr>
<td>receptors and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>protein kinases</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ribosome</td>
<td>RAC, NAC, Jj1</td>
<td>[39]</td>
</tr>
<tr>
<td>myosin</td>
<td>UNC-45</td>
<td>[40]</td>
</tr>
<tr>
<td>spliceosomal snRNPs</td>
<td>plCh</td>
<td>[41]</td>
</tr>
</tbody>
</table>

(a) Nuclear chaperones

The first molecular chaperone to be discovered, nucleoplasmmin, was characterized by its ability to bind histones [2], but subsequent research identified many other nuclear chaperones involved in all phases of the synthesis, transport, assembly and disassembly of histones [33]. The nuclei of eukaryotic cells are crowded with molecules presenting high densities of electrostatic charge to the intranuclear environment, i.e. basic proteins and nucleic acids. This presentation creates the potential for incorrect interactions between molecules bearing many charges of opposite sign. This potential is reduced by the activities of various nuclear assembly chaperones that function during the deposition of histones onto newly replicated DNA, the accumulation of histones in oogenesis, the replacement of sperm protamines by histones after fertilization, the repair of damage to DNA outside S phase and selective gene expression in non-dividing cells. Thus histone chaperones bind to histones during their synthesis in the cytosol, escort them into the nucleus and assist the removal and deposition of histones during transcription, replication and DNA repair. A more recent discovery is that histone chaperones are also involved in the post-translational modification of histones that act as epigenetic markers [34].

The number of known histone chaperones is over 25 and continues to grow. De Koning et al. [33] suggest that they can be divided into three main classes:

1. those that bind, transport or transfer histones without involving additional partners, such as Asf1;
2. multichaperone complexes that combine several histone chaperone subunits, such as the CAF1 complex; and
3. chaperones that provide histone binding capacity within large enzymatic complexes, such as Arp4.

Histone chaperones can also be distinguished by their specificity for different histones. Table 2 provides a snapshot of histone chaperones and the processes that they influence.

During DNA replication in S-phase, chaperones assist the disassembly of the existing nucleosomes, and recycle them to form new nucleosomes on the newly synthesized DNA. The extra histones required during DNA replication are synthesized during S phase, and histone chaperones mediate their transport into the nucleus and their deposition in new nucleosomes. The passage of RNA polymerase II during transcription requires local chromatin rearrangements, and this requires histone replacement and exchange. This exchange can involve acetylation of some histones, which weakens their binding to DNA and thus increases the accessibility of DNA locally.

It is becoming increasingly obvious that assembly chaperones are involved in many, if not all, of the fundamental processes that occur in the nucleus.

(b) Rubisco chaperones

Experiments show that chloroplast Rubisco is a major factor limiting plant growth, so many attempts have been made to try to improve its properties by genetic engineering [7,8]. All such attempts have failed. One problem is the strong tendency of unassembled large subunits to aggregate together. The discovery of the chaperonins [15] raised hopes that co-expression of chaperonin genes with chloroplast Rubisco genes in E. coli would allow correct assembly. However, the only success in this type of experiment has been some cyanobacterial Rubiscos (called form I), and with a simpler form of Rubisco (called form II) found in anaerobic bacteria, consisting of just two large subunits [42]. Rubiscos from these prokaryotes have no agricultural relevance.

The same limitation applies to attempts to renature Rubisco from the denatured state in vitro. The form II Rubisco from anaerobic bacteria can be successfully renatured in vitro, provided that the chaperonins from E. coli (GroEL and GroES) are present, but chloroplast Rubisco cannot, nor can the similar Rubisco found in some cyanobacteria (called form I). The simplest assumption is that additional chaperones may be required for the assembly of form I Rubisco in some species, including species of green plants, and recently, support has been gained for this idea from studies on the in vitro assembly of a cyanobacterial Rubisco after release of large subunits from GroEL. These studies suggest that an additional chaperone called RbcX is required for the assembly of cyanobacterial Rubisco in some species [35].

RbcX functions by inhibiting the tendency of partially folded large subunits to aggregate together, and allows them to interact correctly with small subunits. Rubisco large subunits released from ribosomes are enclosed one at a time inside the cavity in GroEL capped by GroES. Inside this cavity, the single protein chain folds without the danger of it aggregating with similar folding chains. The partly folded chain is then released from the cage, a process that requires ATP hydrolysis, but is still not entirely free from the aggregation hazard, because it exposes a disordered C-terminal peptide of about 60 residues. A conserved binding motif of seven residues in this flexible segment is recognized by a hydrophobic cleft in RbcX. In this way, dimers of large subunit are bound to dimers of RbcX, and so are prevented from aggregating. This binding motif is absent from form II Rubiscos, which comprise only two large subunits.

The RbcX-large subunit dimers assemble into octamers. Rubisco small subunits then displace RbcX from these octamers, forming the Rubisco holoenzyme. Like most assembly chaperones, but unlike many folding chaperones such as GroEL/GroES, RbcX is highly specific for its protein substrates and does not require ATP hydrolysis to function. Figure 1
illustrates the complementary roles of the chaperonins and RbcX in the folding and assembly of the form I Rubisco from cyanobacteria, as deduced from in vitro renaturation studies [35].

Homologues of RbcX are found in plants [43], so an obvious strategy for future research on the chloroplast Rubiscos is to see whether they can also be renatured in vitro by supplying cognate chaperonins and RbcX. Or could it be that even more chaperones are required to assemble the world’s most important enzyme?

I thank Ulrich Hartl for kindly commenting on the draft manuscript.

References


2. Laskey RA, Honda BM, Mills AD, Finch JT. 1978 Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. Nature 275, 416–420. (doi:10.1038/275416a0)


33. Ellis RJ. 1996 Discovery of molecular chaperones. Cell Stress Chaperones 1, 155–160. (doi:10.1379/jcb.146-12681996(001)<0155:DOMC>2.3.CO;2(0)