*Photosynthesis Research* **80:** 333–343, 2004.



*Personal perspective*

# From chloroplasts to chaperones: how one thing led to another

R. John Ellis
*Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK (e-mail: jellis@bio.warwick.ac.uk; fax: +44-2476-523568)*

Received 16 April 2003; accepted in revised form 23 June 2003

*Key words:* Roger Barraclough, chaperonin, fraction I protein, Sean Hemmingsen, isolated intact chloroplasts, molecular chaperone, protein synthesis, Rubisco subunit binding protein, Rubisco synthesis

**Abstract**

Two lessons I have learned during my research career are the importance of following up unexpected observations and realizing that the most obvious interpretation of such observations can be rational but wrong. When you carry out an experiment there is usually an expectation that the result will fall within a range of predictable outcomes, and it is natural to feel pleased when this turns out to be the case. In my view this response is a mistake. What you should be hoping for is a puzzling result that was not anticipated since with persistence and luck further experiments may uncover something new. In this article I give a personal account of how studies of the synthesis of proteins by isolated intact chloroplasts from pea leaves eventually led to the discovery of the chaperonins and the formulation of the general concept of the molecular chaperone function that is now seen to be a fundamental aspect of how all cells work.

*Abbreviations:* GroEL – bacterially encoded protein required for phages lambda and T4 to replicate inside *E. coli*; Rubisco – ribulose-1,5-bisphosphate carboxylase-oxygenase

## Chloroplast research at Warwick

My interest in chloroplasts was stimulated by a person, a meeting and a book. The person was Rachel Leech, whom I met when I joined the Plant Physiology unit of the Agricultural Research Council at Imperial College, London, in 1959. Rachel had an extraordinary and infectious enthusiasm for research and worked for W.O. James, head of the adjacent Botany Department. Rachel introduced me to the method of James and Das (1957) for isolating chloroplasts from bean leaves. She was concerned that the usual methods of differential centrifugation employed at that time for isolating chloroplasts resulted in significant contamination of the chloroplast pellet by mitochondria. The James and Das method uses a sucrose–glycerol density gradient that removes this mitochondrial contamination. Since aspartate was identified as one of the products of labeled carbon dioxide fixation by spinach chloroplasts in pioneer experiments by F.R. (Bob) Whatley and Daniel Arnon (Arnon et al. 1956), it was concluded that chloroplasts contain glutamate-oxaloacetate aminotransferase (now termed aspartate aminotransferase). As it happened I had spent my PhD research studying this enzyme and I was aware that it is present in mitochondria from animal cells and non-green plant tissues. So Rachel prepared the chloroplasts and I assayed the pellets for transaminase. We found that the usual crude pellets of chloroplasts contain this enzyme even after washing by low speed centrifugation but that chloroplasts prepared by the James and Das method do not. Our first paper in *Nature* alerted researchers to the dangers of mitochondrial contamination in chloroplast preparations and concluded that chloroplasts do not contain this transaminase (Leech and Ellis 1961). Some years later, when the methodology for isolating intact chloroplasts capable of *in vivo* rates of carbon dioxide fixation

was being developed by David Walker and others (see Walker 2003), we realized that this conclusion is wrong. Electron microscopy by Rachel showed that the James and Das method removes the envelope and stroma so that the chloroplast preparations that we and many other laboratories used at that time consisted in fact of thylakoids and not intact chloroplasts. With the 20/20 vision of hindsight, it now seems obvious that much more attention should have been paid in the 1950s and early 1960s to the actual components present in the preparations of isolated organelles that were so popular. David Hall published a very useful summary in *Nature* of the properties of at least six different states in which chloroplasts can be isolated, ranging from whole intact organelles to subchloroplast particles (Hall 1972).

In 1969 I was a lecturer in the Biochemistry Department at Aberdeen in Scotland when I was invited to apply for the second appointment at the Department of Biological Sciences at the newly founded University of Warwick. The head of department, Derek Burke, was a virologist who was wedded to the idea, novel in the UK then, that to thrive at that time of economic difficulty it was essential to create large research groups with several permanent appointments centered around a leader who had identified a research topic ripe for exploitation. Since 1961, when I left Imperial College to work in the department of Hans Krebs in Oxford, my research had concerned the control of the pathway of sulphate reduction in bacteria and plants, and my chloroplast publications were confined to clarifying the specificity of antibiotic inhibitors of protein synthesis that were popular at that time. There were many excellent laboratories studying the mechanism of photosynthesis, so joining this competitive arena did not seem sensible for someone with my background. But in 1969 the Society for Experimental Biology organized in London a symposium on the control of organelle biogenesis where there was much discussion, but little hard data, about the respective roles of chloroplast and cytosolic protein synthesis (Miller 1970). The fact that chloroplasts contain significant quantities of DNA and ribosomes was established by 1962. The discovery of such quantities of DNA and ribosomes brought the idea of chloroplast autonomy into vogue to the extent that several attempts were made to grow isolated chloroplasts in culture despite genetic evidence indicating that many genes determining chloroplast structure and function behave in a Mendelian fashion, indicating a nuclear location.

The best general source of the evidence concerning chloroplast genetics at that time was the first edition of a book entitled 'The Plastids' written by Kirk and Tilney-Bassett (1967). John Kirk was writing this book while he and I were at Oxford (UK), and I soon came to admire it for its eclectic and highly critical evaluation of the evidence available about the structure, function, evolutionary origin and biogenesis of this class of plant organelle. This book helped me to develop a more critical approach to the conclusions claimed in papers by stressing the importance of looking hard at the data and considering alternative explanations to the ones favored by the authors.

In the year following this symposium the first use of sodium dodecyl sulfate (SDS) polyacrylamide gels to separate protein chains with high resolution was published (Laemmli 1970). I realized that an area with potential mileage was to combine the methodology devised by David Walker to isolate rapidly from pea leaves intact chloroplasts capable of carbon dioxide fixation with the use of polyacrylamide gels to analyze the products of protein and RNA synthesis by the chloroplast genetic system. Why are so many (up to 50%) of the ribosomes in a leaf cell located in the chloroplast? Is it to make all the proteins of the chloroplast, or only some of them, and if so, which? Do any proteins enter the chloroplast from the cytosol and if so, how? In short, what is the function of the chloroplast genetic system and how does it interact with the nucleo-cytosolic genetic system? The proposal that I form a Chloroplast Research Group at Warwick to study not photosynthesis but biogenesis was accepted, and thus began the most productive part of my research career.

## Light-driven chloroplast protein synthesis

My first PhD student at Warwick was Eric Blair who was stimulated by a paper from the laboratory of Daniel Arnon to devise a system in which isolated intact chloroplasts use light to incorporate labeled amino acids into polypeptides separated on SDS gels. The chloroplasts are isolated very rapidly from young developing pea leaves as recommended by David Walker, so the crude pellets used contain both intact and broken chloroplasts as well as mitochondria and nuclei. This heterogeneity does not matter since the incorporation of amino acids is strictly light-dependent and thus proceeds only within intact chloro-

plasts capable of carrying out photophosphorylation in the absence of added cofactors. This neat trick means that the labeled proteins are the result of run-off by polysomes active within the stroma of intact chloroplasts rather than by polysomes exposed to the resususpension buffer; it is thus more likely that correct elongation, termination and release of the proteins takes place, rendering their identification easier. We found that our isolated chloroplasts make two heavily labeled proteins visible on SDS gels, one present in the stromal fraction and the other in the thylakoid fraction, as well as many less well-labeled products. The thylakoid product we subsequently called peak D because it is the fourth labeled product on the gel but despite much effort we never managed to identify it (Eaglesham and Ellis 1974); we now know it is one of the reaction center proteins of Photosystem II, which by pure coincidence is often referred to as the D1 protein, where D stands for 'diffuse' (see Satoh 2003). The soluble product was identified by Eric Blair using peptide fingerprint analysis as the large subunit of Rubisco, called at that time fraction I protein (see Wildman 2002). This was the first definitive identification of a protein made inside the chloroplast (Blair and Ellis 1973).

There are calculations suggesting that Rubisco is the world's most abundant protein because it constitutes up to 50% of the total soluble leaf protein (Ellis 1979), so our discovery that the Rubisco large subunit is made by chloroplast ribosomes immediately explained why so many of the ribosomes in a leaf are present inside these organelles. Later work by the Warwick Group, and by the group of Nam-Hai Chua at the Rockefeller University, showed that the Rubisco small subunit is made by cytosolic ribosomes with an N-terminal extension that targets it to the chloroplast (Ellis 1981; Bartlett et al. 1982).

We expected that the labeled Rubisco large subunit would be incorporated into the abundant pool of Rubisco holoenzyme, but analysis on polyacrylamide gels lacking denaturing agent showed this not to be the case, unless the chloroplast resuspension buffer is based on sorbitol or sucrose as the osmoticum rather than the KCl that we used routinely (Ellis 1977). Instead we noted that the radioactive large subunit comigrates exactly with another prominently staining band of protein visible on non-denaturing gels of stroma; this protein had an estimated mass of about 700 kDa. The conclusion seemed obvious; the staining band of protein represents some soluble oligomeric form of large subunit which the chloroplasts accumulate before small subunits are imported to assemble the Rubisco holoenzyme. This conclusion we 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 made by isolated chloroplasts (Hartley and Ellis 1973).

## Discovery of the Rubisco large subunit binding protein

In 1979 a postdoc named Roger Barraclough (Figure 1) walked into my office to report the surprising observation that the staining band of protein comigrating with labeled Rubisco large subunit was in fact a different protein. He had done what I should have done in 1973, that is, excised the staining band from the non-denaturing gel and run it on an SDS gel. This analysis clearly showed that the staining protein has a subunit mass of 60 kDa and is unlabeled, while the radioactive large subunit migrates just below at about 52 kDa but is accompanied by no visible stain.

*Figure 1.* Roger Barraclough (left) listening to John Allen (editor of this paper) at a party to celebrate the 10-year anniversary of the Warwick Chloroplast Research Group in 1980.

Thus the staining band is not an oligomeric form of large subunit as I had believed since 1973 but a different protein that is not made by chloroplast ribosomes and which binds newly synthesized large subunits in a non-covalent fashion. The stoichiometry of binding is so low that it does not allow the complex of binding protein with labeled large subunits to be resolved from the bulk of the unlabeled 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. We know now of course that the staining band is the chaperonin 60 component of chloroplasts, so we can reinterpret these early observations in terms of one molecule of labeled Rubisco large subunit (52 kDa) bound to one oligomer of unlabeled chaperonin 60 (804 kDa). Figure 2 shows a subsequent repetition by Richard Johnson, a PhD student at Warwick, of the experiment that revealed the existence of the Rubisco large subunit binding protein.

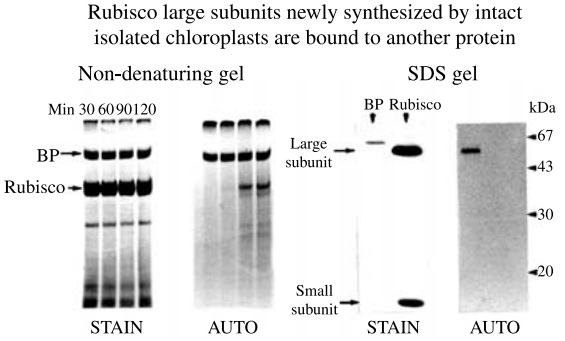


*Figure 2.* Discovery of the chloroplast binding protein (BP). Intact chloroplasts were isolated from young seedlings of *Pisum sativum* and illuminated at 20 °C in a medium containing sorbitol as osmoticum and $^{35}$S-methionine as labeled precursor. Samples were removed at intervals, the chloroplasts centrifuged down, lysed in hypotonic buffer and the soluble fraction electrophoresed on a 5% non-denaturing polyacrylamide gel. The gel was stained in Coomassie Blue (STAIN) and an autoradiograph made (AUTO). The stained bands of Rubisco and BP were excised from the 30 min track and analyzed separately on a 15% SDS polyacrylamide gel. The SDS gel was stained (STAIN) and an autoradiograph made (AUTO). Note that the labeled large subunits comigrate exactly with the staining band of the binding protein (compare the precise shapes of the staining and labeled bands, especially for the 120 min time point). These large subunits can be visualized by their radioactivity but not by their staining since the chemical amount made in this system is very small. The binding protein oligomer (BP on nondenaturing gel) is visible as a prominent stained band, as are its 60 kDa subunits (BP on SDS gel), but these are not radioactive since they are made in the cytosol. BP: Rubisco large subunit binding protein; Rubisco: ribulose bisphosphate carboxylase holoenzyme; large and small, large and small subunits of Rubisco. Experiment performed by Richard Johnson in 1986.

**Proposed role of the binding protein**

The identification of the Rubisco large sunit-binding protein (Barraclough and Ellis 1980) turned to be the first report in any system of a protein that binds the newly synthesized form of another protein, the next example being the BiP protein of the endoplasmic reticulum that binds heavy chains of immunoglobulin (Haas and Wabl 1983). Roger and I decided to concentrate on the binding protein because its surprising ability to bind newly synthesized Rubisco large subunits implied a role in Rubisco assembly.

Roger found that the binding protein has a subunit mass of about 60 kDa, which implied that the oligomer is composed of 12 subunits; today we know the true value is 14 subunits, characteristic of chaperonin 60 oligomers from organelles and prokaryotes. He also noted that labeled Rubisco large subunits bound to the binding protein are not precipitable by antisera to large subunits, unlike labeled 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 newly synthesized polypeptides bind to the top of the central cavity of chaperonin 60.

Roger decided to carry out time-course experiments in which isolated chloroplasts were resuspended in a buffer in which the osmoticum is sorbitol rather than KCl. In such a buffer labeled large subunits appear first in the staining band and later in the Rubisco holoenzyme; he found that large subunit synthesis stops after about 30 min but that the amount of labeled large subunit bound to the binding protein then declines while the amount migrating with the holoenzyme rises, implying a precursor-product relationship (Figure 3). Rubisco was known to be stable until leaves senesce and the chloroplasts were isolated from young growing leaves, so the binding seemed more likely to be part of a Rubisco assembly process rather than a Rubisco degradation mechanism. But why should the assembly of Rubisco require another protein?

It was discovered by John Gray working in Roger Kekwick's laboratory (Gray and Kekwick 1974) and confirmed later (Voordouw et al. 1984) 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; small subunits do not show this behavior. To this day it has proved impossible to reconstitute enzymically active plant Rubisco in a classic

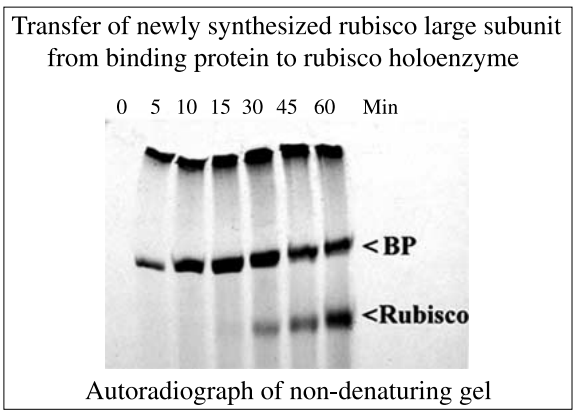


*Figure 3.* Transfer of labeled large subunits from binding protein to Rubisco holoenzyme. Pea chloroplasts were illuminated in a medium containing sorbitol as osmoticum and $^{35}$S-methionine. Samples withdrawn at intervals, the chloroplasts centrifuged down and lysed in hypotonic buffer. The soluble fraction was electrophoresed in a 5% non-denaturing polyacrylamide gel. The figure shows an autoradiograph of the gel. BP: binding protein complexed with labeled large subunits; Rubisco: ribulose bisphosphate carboxylase holoenzyme. Modified from Barraclough and Ellis (1980).

Anfinsen experiment, that is, by removing or diluting the denaturant from a denatured sample. This problem is peculiar to plant Rubisco, since successful renaturation is possible with Rubisco from prokaryote cells. For these reasons it seemed plausible to me to propose that the complex of Rubisco large subunit with the binding protein is an intermediate step in the assembly of Rubisco that functions to keep the large subunits from aggregating with one another until they bind to imported small subunits to form the holoenzyme. This suggestion anticipated the current view that a major function of the chaperonins is to prevent newly synthesized polypeptide chains from aggregating (Hartl and Hayer-Hartl 2002). In our first paper on the binding protein we proposed that the complex of binding protein and large subunit might be an obligatory intermediate in the assembly of ribulose bisphosphate carboxylase (Barraclough and Ellis 1980). Eight years were to elapse before this suggestion excited much interest.

## Polite skepticism

It remains my belief that scientists should resist the natural tendency to ignore unexpected observations that do not fit into the existing paradigm but take the risk of pursuing them in the hope that they will lead to new concepts and discoveries. In 1980 I decided to concentrate on the binding protein. During the next decade several postdocs and PhD students worked on this protein from pea, wheat, spinach and maize. Only one other laboratory pursued this topic in this time. We were greatly encouraged by the confirmation of our initial findings by Harry Roy, and by his demonstration that transfer of labeled large subunits from the binding protein to the Rubisco holoenzyme in isolated chloroplasts requires ATP (Bloom et al. 1983).

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 we did purify, characterize and clone the protein. Sean Hemmingsen (Figure 4) joined my laboratory in 1981 and proved to be an enthusiastic and productive postdoc. He discovered that the purified protein consists of equal amounts of two closely migrating subunits called alpha and beta (Hemmingsen and Ellis 1986) which have different aminoterminal sequences (Musgrove et al. 1987). The binding protein was found to occur in all plastids examined, including colorless plastids from tissues such as seed endosperm, which are not photosynthetic but nevertheless contain large amounts of Rubisco. An important observation made by Sean was that antisera to the binding protein detect a protein of about 60 kDa in extracts of not just photosynthetic bacteria but also of *Escherichia coli.* I recall that I was initially dubious of the significance of this finding because I was then wedded to the idea that the binding protein is nature's way

*Figure 4.* Sean Hemmingsen in 1984.

of overcoming the aggregation tendency of Rubisco large subunits.

During the period 1980–1987 I talked about our studies at many institutes and at conferences on plant molecular biology and photosynthesis. Few animal or microbial scientists attend such conferences, so our studies did not influence these fields. The response to the idea that the binding protein is required for Rubisco assembly by keeping the large subunits in a soluble form was generally polite skepticism, which I attribute to the acceptance of the principle of protein self-assembly established by the pioneering work of Christian Anfinsen, Donald Caspar and Aaron Klug. It was often pointed out to me that the folding of newly synthesized polypeptide chains, and any subsequent association into oligomers, is a spontaneous self-assembly process requiring no additional macromolecules, and so it said in the texbooks. A more pertinent objection was raised by John Gray who pointed out correctly that what I was observing was the binding of a very small amount of a highly labeled protein to a very large amount of a different unlabeled protein – ideal stoichiometry for a binding artefact. So I was desperate for a precedent and was referred by a staff colleague, Alan Colman, who later directed research that led to the cloning of Dolly the sheep, to a paper on the assembly of nucleosomes in extracts of *Xenopus* eggs (Laskey et al. 1978). This paper introduced the term 'molecular chaperone' to describe the properties of a nuclear protein called nucleoplasmin required for the correct assembly of nucleosomes from histones and DNA in egg extracts.

## Origin of the term molecular chaperone

Ron Laskey was interested in the rapid formation of nucleosomes that occurs when amphibian eggs are fertilized. 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, and the principle of self-assembly predicts that nucleosomes should reform when these structural components are mixed together in buffer at intracellular ionic strengths. This experiment is a spectacular failure; addition of monomeric histones to DNA at physiological ionic strength results in the rapid appearance of unspecific aggregates rather than nucleosomes. Further experiments in Ron Laskey's laboratory showed that addition of small amounts of *Xenopus* egg homogenate prevents this aggregation and allows nucleosome formation. The active factor was purified, characterized as an abundant acidic nuclear protein, and called nucleoplasmin. This protein binds to histones and reduces their strong positive charge; addition of the negatively charged DNA at physiological ionic strength then results 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 nucleosomes themselves. Secondly, nucleosomes can be assembled from histones and DNA in the absence of nucleoplasmin provided that the high salt concentration is reduced slowly by dialysis. Thus the role of nucleosomes is not to provide steric information for nucleosome assembly, but to reduce the positive charge of the histone monomers. This transient inhibitory role of nucleoplasmin allows the intrinsic self-assembly properties of the histones with DNA to predominate over the competing incorrect interactions made possible by the high density of opposite charges. The molecular details of how this is achieved are still obscure. But it is clear that the role of nucleoplasmin in nucleosome assembly is transitory and does not involve the formation or breakage of covalent bonds. It can thus be detected only by the use of non-denaturing techniques applied to 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' (Laskey et al. 1978).

This term was coined because the properties of nucleoplasmin are a precise molecular analogy to the role of human chaperones. The traditional role of the latter is to prevent incorrect interactions between pairs of human beings without either providing steric information necessary for their correct interaction or being present during their married life – but sometimes reappearing during divorce and remarriage!

## Extension of the term to other proteins

Since 1978 the term molecular chaperone did not appear in any other publication and the term was not used

to describe any other protein. I read the Laskey paper in 1985 and realized that since unassembled Rubisco large subunits also have a strong tendency to undergo incorrect interactions, perhaps the role of the binding protein could be thought of as similar to that of nucleoplasmin, that is, preventing aggregation by masking the interactive surfaces involved. For Rubisco large subunits the interactive surfaces are hydrophobic in nature whereas for histones the interactive surfaces are charged, but the principle is the same. The suggestion that the binding protein could be regarded as the second example of a molecular chaperone was made at a Royal Society discussion meeting on Rubisco that I organized with John Gray in 1985 and subsequently published in the proceedings (Musgrove and Ellis 1986).

I thought at first that nucleoplasmin and the chloroplast binding protein were special cases evolved to deal with certain oligomeric proteins whose assembly presented particular difficulty because of the propensity of their subunits to aggregate. What prompted me to extend the chaperone idea further was a speculative paper by Pelham (1986) on the functions of the heat shock protein (hsp) 70 and 90 families. This paper does not discuss either nucleoplasmin or the chloroplast binding protein or use the term molecular chaperone, but suggested that hsp70 and hsp90 proteins may be involved in a variety of protein assembly and disassembly processes. Pelham proposed that these heat shock proteins play such a role in normal unstressed cells and are required in increased amounts when proteins have been damaged by stress, both to unscramble aggregates and to prevent further damage by binding to exposed hydrophobic surfaces. It occurred to me that all these ideas could be gathered together under the umbrella of a more fundamental chaperone concept. Perhaps many different types of molecular chaperone exist because the problem of incorrect interactions is not confined to a few proteins but is widespread.

I presented this idea at the North Atlantic Treaty Organization (NATO) Advanced Study Institute meeting on plant molecular biology organized by Dieter von Wettstein in Copenhagen (Denmark) in June 1987. A representative from *Nature* was at that meeting who encouraged me to write a News and Views article describing this more generalized molecular chaperone concept. This article appeared later that year (Ellis 1987) 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 polypeptide chains and their assembly into oligomeric structures occur correctly*.' Thus the general concept of the molecular chaperone function was born from my realization that several apparently unrelated discoveries could be regarded as particular examples of a general, but hitherto unrecognized, intracellular process. This article marked the start of the now widespread use of the term molecular chaperone in the biochemical literature.

## 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 binding protein could be determined. This work reached fruition in 1987 when he determined the sequence of the alpha subunit of the binding protein found in the colorless plastids of castor bean endosperm. Checking with the sequence databases revealed a high amino acid sequence similarity to the *ams* gene of *E. coli*, a gene implicated in the control of mRNA stability (Chanda et al. 1985). Sean contacted H.F. Kung, the senior author on the Chanda et al. paper, who told him of a related protein sequenced by the immunologist Rick Young; this protein is the 65 kDa common bacterial antigen of *Mycobacterium leprae.* The common bacterial antigen, as the name indicates, is an antigen found in all bacterial species examined and is the dominant antigen in human bacterial diseases such as tuberculosis and leprosy. This link explained Sean's earlier finding that antisera to the chloroplast binding protein detect a 65 kDa protein in bacterial extracts. However enquiries to immunologists about the function of the common bacterial antigen revealed that nothing was known, so this similarity puzzled us greatly for a time.

The breakthrough came at the end of 1987 when Rick Young informed Sean about a protein sequence determined by the microbiologist Roger Hendrix; this protein was GroEL from *E. coli*. The resulting phone conversation between Sean and Roger left them both excited about the high similarity (about 50% identity) between two proteins both implicated in the assembly of other proteins. Roger had sequenced the GroEL protein some years before but had not put the sequence into the database. He had the draft of a manuscript describing the GroE operon on his desk. Sean relayed this information to me in the autumn of 1987 and I

then looked up the literature on GroEL, a protein I was unaware of until then.

This was my first eureka moment in science. I still recall the growing excitement I felt as I realized that the properties of GroEL are very similar to that of the chloroplast binding protein. Not only is GroEL an oligomer of fourteen 65 kDa subunits (Hendrix 1979) but it is also implicated in the assembly of another protein. GroEL was identified in several laboratories in the early 1970s as a bacterially-encoded protein required for phages lambda and T4 to replicate inside *E. coli* (Georgopolous et al. 1983). ('Gro' refers to phage growth and the suffix E refers to the observation that the failure of phage lambda to grow when GroE is mutated is overcome when the phage carries a mutation in the phage head gene E. 'L' refers to the larger protein encoded in the GroE operon while the smaller protein is called GroES.) There was evidence that the GroEL oligomer binds transiently and non-covalently to subunits of phage lambda B protein; 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. A mutation in GroEL results in the head proteins of phage T4 forming insoluble aggregates that associate with the bacterial cell membrane. 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 phage assembly 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 mature PhD student in my laboratory from 1986 to 1989, had isolated and sequenced a cDNA clone for the alpha subunit of the binding protein from wheat chloroplasts; the amino acid sequence is about 46% identical with that of GroEL and about 80% identical to that of the alpha subunit from castor bean. Papers from Richard Hallberg's laboratory reported that antisera to GroEL detect a protein in mitochondria from *Tetrahymena*, maize, yeast and human cells (McMullin and Hallberg 1988). Sean and I realized that there was now evidence from both bacteria and chloroplasts that linked the involvement of highly similar pre-existing proteins to the assembly of newly synthesized other proteins in a manner that fitted the general concept of molecular chaperones outlined in the 1987 *Nature* paper (Ellis 1987). Sean was inspired to call this family of proteins the 'chaperonins' and was able to convince Roger Hendrix of the merits of pooling our sequence data with his and to present the package to *Nature*. This paper appeared in May 1988 (Hemmingsen et al. 1988) and sparked the continuing wave of research in many laboratories on the structure and function of the chaperonins in both prokaryotic and eukaryotic cells. It was only 1 year later that two key papers appeared in *Nature* that clarified the site of action of the chaperonins at preventing aggregation (Goloubinoff et al. 1989; Ostermann et al. 1989).

A seminal paper that followed the identification of the chaperonins came from the laboratory of George Lorimer who studied the renaturation of the simpler dimeric form of Rubisco found in some bacteria. Elegant experiments with the Rubisco from *Rhodospirillum rubrum* showed that GroEL does indeed prevent incorrect interactions between the large subunits (Goloubinoff et al. 1989). The key observation is that denatured large subunits self-aggregate into inactive forms when the denaturant is diluted into renaturing buffer, but if GroEL is present in this buffer stable complexes form between the large subunits and the chaperonin; these complexes can be dissociated by adding a related chaperonin called GroES and ATP, with the appearance of enzymically active Rubisco. The chaperonins do not convey steric information for the assembly of bacterial Rubisco because large subunits will reassemble correctly in the absence of chaperonins provided the temperature is kept below 15 °C; low temperatures inhibit hydrophobic interactions.

Enzymatic activity by bacterial Rubisco is shown only by the dimeric molecule, so these renaturation experiments were unable to distinguish whether the chaperonin acts to assist the association of monomers to form the enzymically active dimers or acts to assist the folding of monomers which then associate into dimers. The first report to establish that the chaperonins act at the folding stage appeared from the laboratory of Ulrich Hartl just before the Lorimer report (Ostermann et al. 1989). This conclusion was based on studies of the folding of the monomeric enzyme dihydrofolate reductase after transport into mitochondria isolated from *Neurospora*. In the following decade a stream of high quality papers from the laboratories of Ulrich Hartl and Arthur Horwich established the current model for the mechanism of action of the chaperonins (reviewed in Hartl and Hayer-Hartl 2002). The essence of this mechanism is that the central cavity of the GroEl oligomer acts as an 'Anfinsen cage' within which a single partly folded protein chain can continue to fold in the absence of similar

chains with which it might otherwise aggregate (Ellis 2001a, b).

## Impact of the chaperone concept

Since 1987 the number of published papers with molecular chaperone (or chaperone) in the title or abstract has risen from one to the current total of 6315. Use of the term spread rapidly to the extent that regular conferences devoted to this topic have been run by the European Science Foundation, the European Molecular Biology Organisation and the Cold Spring Harbor Laboratory. The Molecular Chaperone Club has organized annual informal meetings in the UK since 1990 and regularly attracts about 100 participants (see www.bio.warwick.ac.uk/molchap). The journal *Cell Stress and Chaperones* (www.cellstress.uconn.edu) was launched in 1996 and is edited by Lawrence Hightower, one of the founders of the related stress protein field. The ultimate mark of respectability is the use of the term in student texts of biochemistry and molecular biology. But what actually is the molecular chaperone concept?

The first comprehensive statement of the general chaperone concept appeared in *Trends in Biochemical Sciences* (Ellis and Hemmingsen 1989). The suggestions made in this account have so far stood the test of time (Ellis and Hartl 2003). Molecular chaperones are defined as a large and diverse group of unrelated proteins that share the functional property of assisting the non-covalent assembly/disassembly of other macromolecular structures but which are not permanent components of these structures when they are performing their normal biological functions. Assembly is used here in a broad sense and includes several intracellular processes: the folding of nascent polypeptides both during their synthesis and after release from ribosomes, the unfolding and refolding of polypeptides during their transport 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 disassembly processes such as the partial unfolding and dissociation of subunits when some proteins carry out their normal functions, and the resolubilization and/or degradation of proteins partially denatured and/or aggregated by mutation and exposure to environmental stresses such as high temperatures. Some, but not all, chaperones are also stress proteins because the requirement of 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 the above definition is functional not structural and contains no constraints on the mechanism by which different chaperones may act; this is the reason for the use of the imprecise term ‘assist’. Molecular chaperones are not defined by either a common mechanism or by sequence similarity. Thus only two criteria need be satisfied to designate a macromolecule a molecular chaperone: (1) it must in some sense assist the non-covalent assembly/disassembly of other macromolecular structures, the mechanism being irrelevant, but (2) it must not be a component of these structures when they carry out their normal biological functions. In all cases studied to date, chaperones bind to regions of macromolecular structures that are inaccessible when these structures are correctly assembled, but that are accessible at other times. The number of distinct chaperone families stands at over 25 to date; the chaperonins are just one of these families (Ellis 1999).

Given that most denatured pure proteins that have been examined can refold *in vitro* into their functional conformations on the removal of the denaturing agent, the question arises as to why molecular chaperones exist. Current evidence suggests that, with two possible exceptions (Ellis 1998), chaperones do not provide steric information required for proteins to fold correctly, but 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 hydrophobic or charged) on their surfaces. In aqueous environments hydrophobic surfaces stick together, while charged surfaces will bind to ones bearing the opposite charge, a problem acute in the nucleus where negatively charged nucleic acids are bound to positively charged proteins. This aggregation problem is accentuated by the crowded state of the intracellular media that greatly favors macromolecular association processes (Ellis 2001a).

These considerations can be summarized into a simple unifying principle. Cells need a molecular chaperone function to both prevent and reverse incorrect interactions that may occur when potentially interactive surfaces are exposed to the intracellular environment. Thus the basic biological principle of spontaneous self-assembly has been replaced by the principle of chaperone-assisted self-assembly.

*Figure 5.* John Ellis (author) in 1999.

As requested by the editors, I have included here a photograph of myself (Figure 5).

## Acknowledgments

I am very grateful to Robert Freedman for providing encouragement and office facilities during my retirement. This paper was edited by John F. Allen.

## References

Arnon DI, Allen MB and Whatley FR (1956) Photosynthesis by isolated chloroplasts. 4. General concept and comparison of 3 photochemical reactions. Biochim Biophys Acta 20: 449–461

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

Bartlett SG, Grossman AR and Chua NH (1982) *In vitro* synthesis and uptake of cytoplasmically-synthesized chloroplast proteins. In: Edelman M, Hallick RB and Chua NH (eds) Methods in Chloroplast Molecular Biology, pp 1081-1091. Elsevier Biomedical Press, Amsterdam

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

Eaglesham ARJ and Ellis RJ (1974) Protein synthesis in chloroplasts. II. Light-driven synthesis of membrane proteins by isolated pea chloroplasts. Biochim Biophys Acta 335: 396–407

Ellis RJ (1977) Protein synthesis by isolated chloroplasts. Biochim Biophys Acta 463: 186–215

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 (1998) Steric chaperones. Trends Biochem Sci 23: 43–45

Ellis RJ (1999) Molecular chaperone. In: Creighton TE (ed) Encyclopedia of Molecular Biology. Wiley, New York

Ellis RJ (2001a) Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci 26: 597–604

Ellis RJ (2001b) Inside and outside the Anfinsen cage. Curr Biol 11: R1038–R1040

Ellis RJ and Hartl FU (2003) Protein folding and chaperones. In: Cooper DN (ed) Nature Encyclopedia of the Human Genome, Vol 4, pp 806–810. Nature Publishing Group, London

Ellis RJ and Hemmingsen SM (1989) Molecular chaperones: proteins essential for the biogenesis of some macromolecular structures. Trends Biochem Sci 14: 339–342

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

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

Gray JC and Kekwick RGO (1974) An immunological investigation of the structure and function of ribulose 1.5-bisphosphate carboxylase. Eur J Biochem 44: 481–489

Haas IG and Wabl M (1983) Immunoglobulin heavy-chain binding-protein. Nature 306: 387–389

Hall DO (1972) Nomenclature for isolated chloroplasts. Nature New Biol 235: 125–126

Hartl FU and Hayer-Hartl M (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295: 1852–1858

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 the ribulose bisphosphate large subunit binding protein. Plant Physiol 80: 269–276

Hemmingsen SM, Woolford C, van der Vies SM, Tilly K, Dennis DT, Georgopoulos CP, Hendrix RW and Ellis RJ (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. Nature 333: 330–334

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

James WO and Das VSR (1957) The organisation of respiration in chlorophyllous cells. New Phytol 56: 325–343

Kirk JTO and Tilney-Bassett RAE (1967) The Plastids: Their Chemistry, Structure, Growth and Inheritance. Freeman, London

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

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685

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

Leech RM and Ellis RJ (1961) Coprecipitation of mitochondria and chloroplasts. Nature 190: 790–792

McMullin TW and Hallberg RL (1988) A highly evolutionarily conserved mitochondrial protein is structurally related to the protein encoded by the *Escherichia coli* GroEL gene. Mol Cell Biol 8: 371–380

Miller PTO (1970) Control of Organelle Development. Symposium No. XXIV of the Society for Experimental Biology. Cambridge University Press, Cambridge, UK

Musgrove JE and Ellis RJ (1986) The Rubisco large subunit binding protein. Phil Trans R Soc London Ser B 313: 419–428

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

Ostermann J, Horwich AL, Neupert W and Hartl FU (1989) Protein folding in mitochondria requires hsp60 and ATP hydrolysis. Nature 341: 125–130

Pelham HRB (1986) Speculations of the functions of the major heat shock and glucose-regulated proteins. Cell 46: 956–971

Satoh K (2003) The identification of the Photosystem II reaction center: a personal story. Photosynth Res 76: 233–240

Voordouw G, van der Vies SM and Bouwmeister PP (1984) Dissociation of ribulose bisphosphate carboxylase oxygenase from spinach by urea. Eur J Biochem 141: 313–318

Walker DA (2003) Chloroplasts in envelopes: $CO_2$ fixation by fully functional intact chloroplasts. Photosynth Res 76: 319–327

Wildman SG (2002) Along the trail from Fraction I protein to Rubisco (ribulose bisphosphate carboxylase-oxygenase). Photosynth Res 73: 243–250