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

Discovery of molecular chaperones

R. John Ellis

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

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

CHLOROPLAST RESEARCH

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

Received 25 March 1996; Accepted 15 April 1996

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

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

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

AN ERRONEOUS CONCLUSION

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

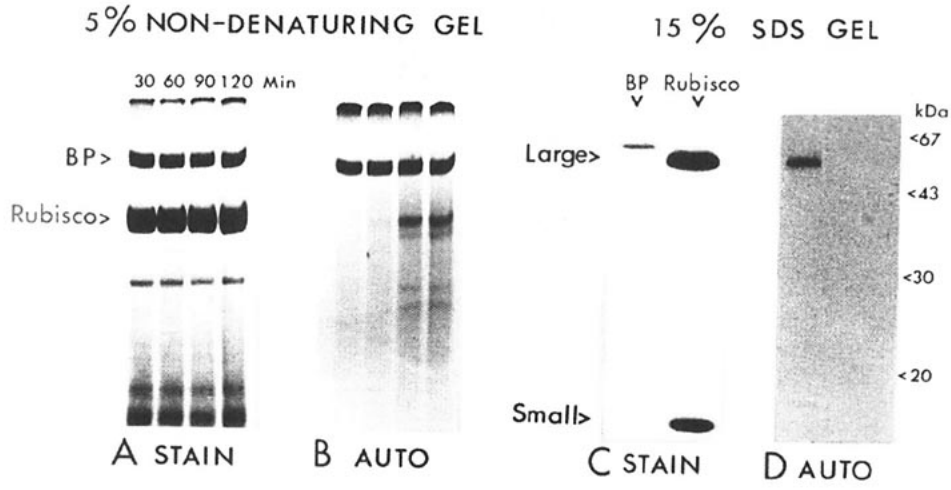

Figure Discovery of the chloroplast-binding protein (BP). Intact chloroplasts were isolated from young seedlings of *Pisum sativum* and illuminated at 20°C in a medium containing sorbitol as osmoticum and ^{35}S-methionine as labelled precursor; the illumination serves as an energy source for protein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

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

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

PROPOSED ROLE OF THE BINDING PROTEIN

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

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

A SPELL IN LIMBO

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

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

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

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

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

ORIGIN OF THE TERM MOLECULAR CHAPERONE

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

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

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

EXTENSION TO OTHER PROTEINS

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

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

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

DISCOVERY OF THE CHAPERONINS

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

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

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

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

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

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

Note added in proof

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

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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