root nodule formation, and be of general significance for our insight into the development of plants.

REFERENCES AND NOTES

1. F. J. Bergersen, *Root Nodules of Legumes: Structure and Functions* (Research Studies Press, Wiley, Chichester, 1982).
2. K. D. White, *Roman Farming* (Thames and Hudson, London, 1970).
3. J. M. Vincent, in *Nitrogen Fixation*, W. E. Newton and W. H. Orme-Johnson, Eds. (University Park Press, Baltimore, 1980), vol. 2, pp. 103–129.
4. D. P. S. Verma and S. R. Long, *Int. Rev. Cytol. Suppl.* **14**, 211 (1983).
5. K. R. Schubert, *Annu. Rev. Plant Physiol.* **37**, 539 (1986).
6. B. G. Rolfe and P. Gresshoff, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **39**, 297 (1988).
7. S. R. Long, *Cell* **56**, 203 (1989).
8. J. P. Nap and T. Bisseling, in *Molecular Biology of Symbiotic Nitrogen Fixation*, P. M. Gresshoff, Ed. (CRC Press, Boca Raton, FL, 1990), pp. 181–229.
9. A. van Kammen, *Plant Mol. Biol. Rep.* **2**, 43 (1984).
10. F. Govers, J. P. Nap, A. van Kammen, T. Bisseling, *Plant Physiol. Biochem.* **25**, 309 (1987).
11. A. Kondorosi, in *Plant-Microbe Interactions*, T. Kosuge and E. W. Nester, Eds. (McGraw-Hill, New York, 1989), vol. 3, pp. 383–420.
12. E. Appelbaum, in *Molecular Biology of Symbiotic Nitrogen Fixation*, P. M. Gresshoff, Ed. (CRC Press, Boca Raton, FL, 1990), pp. 131–158.
13. G. N. Gussin, C. W. Ronson, F. M. Ausubel, *Annu. Rev. Genet.* **20**, 567 (1986); G. Ditta, in *Plant-Microbe Interactions*, T. Kosuge and E. W. Nester, Eds. (McGraw-Hill, New York, 1989), vol. 3, pp. 11–24.
14. J. Kijne, in *Biological Nitrogen Fixation*, G. Stacey, R. Burris, H. Evans, Eds. (Chapman and Hall, New York, in press).
15. G. Caetano-Anolés, D. K. Crist-Estes, W. D. Bauer, *J. Bacteriol.* **170**, 3164 (1988).
16. T. Gloudemans *et al.*, *Plant Mol. Biol.* **12**, 157 (1989).
17. P. Y. Yao and J. M. Vincent, *Plant Soil* **45**, 1 (1969).
18. P. Lerouge *et al.*, *Nature* **344**, 781 (1990).
19. P. Lerouge *et al.*, in *Proceedings of the International Symposium on Nitrogen Fixation*, P. M. Gresshoff and G. Stacey, Eds. (in press).
20. K. A. Vanden Bosch *et al.*, *EMBO J.* **8**, 335 (1989).
21. B. Scheres *et al.*, *Plant Cell*, in press.
22. B. Scheres *et al.*, *Cell* **60**, 281 (1990).
23. J. P. Knox, S. Day, K. Roberts, *Development (Cambridge)* **106**, 47 (1989).
24. C. L. Diaz *et al.*, *Nature* **338**, 579 (1989).
25. B. Horvath and T. Bisseling, unpublished results.
26. L. D. Kuykendall, *Int. Rev. Cytol. Suppl.* **13**, 299 (1981).
27. W. Newcomb, D. Sippel, R. L. Peterson, *Can. J. Bot.* **57**, 2603 (1979).
28. W. Newcomb, *Int. Rev. Cytol. Suppl.* **13**, 247 (1981).
29. K. Tran Thanh Van *et al.*, *Nature* **314**, 615 (1985).
30. J. F. Witty, F. R. Minchin, L. Skøt, J. E. Sheehy, *Oxf. Surv. Plant Mol. Cell Biol.* **3**, 275 (1986).
31. C. van de Wiel *et al.*, *EMBO J.* **9**, 1 (1990).
32. H. J. Franssen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4495 (1987).
33. A. Delauney and D. P. S. Verma, *Plant Mol. Biol. Rep.* **6**, 279 (1988).
34. C. A. Appleby, *Annu. Rev. Plant Physiol.* **35**, 433 (1984).
35. J. P. Nap and T. Bisseling, *Physiol. Plant.* **79**, 407 (1990).
36. J. G. Robertson and K. J. F. Farnden, in *The Biochemistry of Plants*, B. J. Miflin, Ed. (Academic Press, New York, 1980), vol. 5, pp. 85–113.
37. M. E. Baker and M. H. Saier, *Cell* **60**, 185 (1990).
38. D. P. S. Verma *et al.*, in: *Abstracts of VI NATO Advanced Study Institute on Plant Molecular Biology* (1990), p. 188.
39. J. Stougaard, K. A. Marcker, L. Otten, J. Schell, *Nature* **321**, 669 (1986); E. Ø. Jensen, K. A. Marcker, J. Schell, F. J. De Bruijn, *EMBO J.* **7**, 1265 (1988).
40. K. Jacobsen *et al.*, *Plant Cell* **2**, 85 (1990).
41. W. C. Yang and T. Bisseling, unpublished results.
42. M. E. Dudley, T. W. Jacobs, S. R. Long, *Planta* **171**, 289 (1987).
43. K. R. Libbenga and R. J. Bogers, in *The Biology of Nitrogen Fixation*, A. Quispel, Ed. (North-Holland, Amsterdam, 1974), pp. 430–472.
44. A. M. Hirsch, T. V. Bhuvaneswari, J. G. Torrey, T. Bisseling, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1244 (1989).
45. S. R. Long and J. Cooper, in *Molecular Genetics of Plant-Microbe Interactions*, R. Palacios and D. P. S. Verma, Eds. (APS Press, St. Paul, 1988), pp. 163–178.
46. M. A. Hall, in *Developmental Control in Animals and Plants*, C. F. Graham and P. F. Wareing, Eds. (Blackwell, Oxford, ed. 2, 1984), pp. 313–330.

47. We thank H. J. Franssen, F. Govers, B. Horvath, and A. van Kammen for critical reading of the manuscript, M. J. van Iersel, G. Heitkönig, and L. Kesaulya-Monster for secretarial assistance, P. Madern for drawings, and C. van de Wiel for cytological help in preparing Fig. 3.

Molecular Chaperones: The Plant Connection

R. JOHN ELLIS

Molecular chaperones are a family of unrelated proteins found in all types of cell. They mediate the correct assembly of other polypeptides, but are not components of the mature assembled structures. Chaperones function by binding specifically to interactive protein surfaces that are exposed transiently during many cellular processes and so prevent them from undergoing incorrect interactions that might produce nonfunctional structures. The concept of molecular chaperones originated largely from studies of the chloroplast enzyme rubisco, which fixes carbon dioxide in plant photosynthesis; the function of chaperones forces a rethinking of the principle of protein self-assembly.

THE STUDY OF THE MOLECULAR BIOLOGY OF PLANTS IS undergoing a rapid expansion, fueled both by technical advances and by the realization that there may be economic advantages if plants can be manipulated in new ways. One of the few basic concepts in molecular biology that has originated from research with plants is that of molecular chaperones. This concept developed from studies on the biogenesis of the chloroplast enzyme that fixes carbon dioxide in photosynthesis (ribulose bisphosphate carboxylase-oxygenase or rubisco), but the field now encompasses animal and microbial cells and has medical and biotechnological aspects (*1–10*). In this article I describe how the chaperone concept developed from studies on rubisco and discuss the implications of this concept for future plant research.

Rubisco Biogenesis

Rubisco indirectly or directly plays a vital role in the metabolism of all cells, since it is the principal catalyst that brings carbon into organic combination from atmospheric carbon dioxide during photosynthesis. Its biogenesis is unusually complex and involves the interaction of light as a developmental trigger with two distinct genetic systems, one located within the chloroplast and the other within the nucleus (*11*). Despite its vital role, rubisco is a poor catalyst and has both a low affinity for carbon dioxide and a small turnover number; thus autotrophic organisms devote a major part

The author is professor in the Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom.

of their synthetic effort to making many molecules of this enzyme. Indeed, rubisco is thought to be the most abundant single protein in the biosphere (*11*). In addition, this enzyme catalyzes an oxygenase reaction, apparently as an inevitable consequence of its mechanism of action as a carboxylase (*12*). This oxygenase reaction is the first step in photorespiration, a process peculiar to plants, which salvages the products of the oxygenase reaction. This salvage pathway is not totally efficient; some carbon is lost as carbon dioxide during its operation and this loss is a major factor that limits plant productivity. Thus, the challenge for genetic engineers is to try to improve the ability of rubisco to catalyze carbon dioxide fixation (*12*). All attempts to assemble enzymically active molecules of crop plant rubisco in *Escherichia coli* have failed, so that in vitro mutagenesis techniques cannot be used to improve the properties of this vital catalyst. This assembly problem has provided a spur for studies on the mechanism of assembly of rubisco (*13*).

The rubisco in the majority of photosynthetic organisms is an oligomeric molecule consisting of eight large subunits (molecular weight, ~52,000) and eight small subunits (molecular weight, ~14,000). The large subunits each carry an active site but each site includes residues from an adjacent large subunit so that dimers must assemble for activity. Small subunits are necessary for full activity, since octamers of large subunits have only 1% of the activity of the hexadecamer. In most organisms, the large subunits are encoded in the chloroplast genome and are synthesized by chloroplast ribosomes, while the small subunits are encoded in nuclear genes and synthesized in the cytosol as larger precursors that are subsequently imported into chloroplasts (*12, 14–16*). During investigations of rubisco large subunits synthesized by isolated chloroplasts, a protein was discovered (which turned out to be a molecular chaperone) that mediates the rubisco assembly process inside the chloroplast (*17, 18*).

The large subunit of rubisco is the major soluble product of light-driven protein synthesis by isolated intact chloroplasts (*19*). Analysis of this product by nondenaturing polyacrylamide gel electrophoresis reveals that the majority of large subunits synthesized in vitro does not migrate with the existing rubisco oligomer made by the plant before chloroplast isolation, indicating that rubisco assembly is slow compared to synthesis. The rubisco large subunits migrate instead with another abundant chloroplast protein of lower mobility. Initially we thought that this other protein was an aggregate of rubisco large subunits since it comigrated exactly with rubisco large subunits synthesized in vitro; however, analysis by two-dimensional gel electrophoresis revealed that this conclusion was incorrect (*17*). Instead this other protein is an oligomer of 14 subunits, each with a molecular weight of about 60,000. This oligomer binds noncovalently to rubisco large subunits synthesized in vitro (Fig. 1); the stoichiometry of binding is one large subunit bound to one oligomer. This abundant chloroplast protein was called the rubisco large subunit binding protein or binding protein. The binding protein is composed of two related types of 60-kD subunit, α and β, which are encoded by nuclear genes and are imported into the chloroplast after synthesis in the cytosol (*18, 20*).

Protein synthesis by isolated intact chloroplasts ceases after about 30 minutes, but at longer times the amount of large subunits associated with the binding protein declines, while the amount in the rubisco oligomer increases. These observations led to the proposal that the transient binding of the large subunits to another protein might be an obligatory step in the rubisco assembly process (*17*). This notion was indirectly supported by reports that large subunits prepared from purified rubisco by the use of denaturing agents form insoluble aggregates when the denaturant is removed, even in the presence of soluble small subunits (*21*). These results imply that rubisco large subunits tend to undergo incorrect interactions with one another, that is, interactions that produce structures that are not functional in their normal biological context.

The proposal that the assembly of an oligomeric protein might require the assistance of another protein had few precedents in 1980. The paradigm for protein folding and oligomerization was the notion of self-assembly, which argues that all the information for the structure and function of a protein resides within the amino acid sequence of its polypeptide chains. There was some evidence for "assembly proteins" in certain phage systems, but the prevailing view was that these were exceptions to the general rule; many purified proteins had been successfully renatured in vitro so there was no need to invoke special assembly factors in vivo (*22*). Nevertheless the idea was pursued by us and by Harry Roy (*23*). To rule out the possibility that the binding phenomenon was an artifact, we attempted to develop an in vitro system where the assembly of enzymically active rubisco oligomers could be demonstrated, so that we could test for the possible involvement of other proteins in the assembly process. After much effort this goal eluded us, but progress came later from a different direction. Meanwhile I came across the term "molecular chaperone."

Origin of the Term "Molecular Chaperone"

The term molecular chaperone was first used by R. A. Laskey and his colleagues to describe the properties of nucleoplasmin, an abundant acidic nuclear protein required for the assembly of nucleosomes from histones and DNA in *Xenopus* oocyte extracts (*24*). Addition of monomeric histones to DNA under physiological conditions of ionic strength results in the rapid appearance of nonspecific aggregates rather than nucleosomes. If the histones are first incubated with nucleoplasmin before addition of DNA, however, nucleosome cores form instead. Nucleoplasmin is required only for nucleosome assembly and does not form part of the nucleosomes themselves. The steric information for nucleosome assembly resides in the histones and not in nucleoplasmin. The role of the nucleoplasmin is thus not to provide steric information for

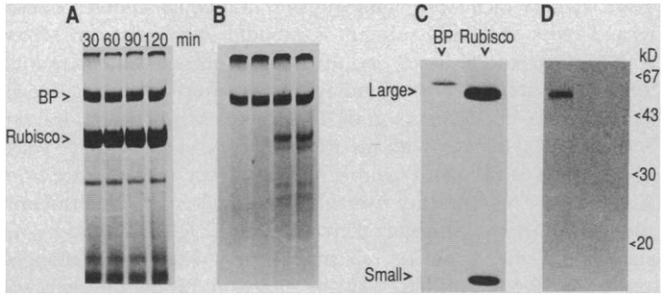

Fig. 1. Binding of rubisco large subunits synthesized in vitro to another chloroplast protein [binding protein (BP)]. Intact chloroplasts were isolated from young leaves of pea (*Pisum sativum*) and illuminated at 20°C in a medium containing sorbitol as osmoticum and [^{35}S]methionine as labeled precursor. Samples were removed at intervals, the chloroplasts lysed in hypotonic buffer, and the soluble fraction electrophoresed on a (**A** and **B**) 5% nondenaturing polyacrylamide gel. The gel was stained in Coomassie Blue (A) and an autoradiograph made (B). The stained bands of rubisco and BP were excised from the 30-min track and analyzed separately on a (**C** and **D**) 15% polyacrylamide gel containing SDS as a denaturing agent. The SDS gel was stained (C) and autoradiographed (D). Labeled large subunits comigrate exactly with the staining band of the binding protein (A and B); these large subunits can be visualized by their radioactivity but not by their staining (C and D); the binding protein is visible as a stained band on the SDS gel but is not labeled (C and D). BP, rubisco large subunit binding protein; rubisco, holoenzyme of ribulose bisphosphate carboxylase-oxygenase; large and small, large and small subunits of rubisco. These results are from (*52*) with permission.

nucleosome assembly but to reduce the positive charges of the heavily charged histone monomers and so allow correct interactions with DNA to predominate over incorrect interactions. In the words of Laskey and his colleagues (*24*), "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."

The work on nucleoplasmin suggested 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. This suggestion that the binding protein might be a molecular chaperone required for correct rubisco assembly was made at an international symposium on rubisco in December 1985 and subsequently published (*1*). It soon became clear that the need for chaperone function might be much more widespread. H. R. B. Pelham argued that members of the major heat shock family of proteins (hsp70) in animal and microbial cells are involved in the assembly and disassembly of proteins in the nucleus, cytosol, and endoplasmic reticulum (*25*). Some members of the hsp70 family are present in unstressed cells, and they can bind to denatured or abnormal proteins in a manner reversible by adenosine triphosphate (ATP). Pelham suggested that these proteins play a role in protein folding and assembly in unstressed cells and are required in increased amount when proteins have been damaged by stress, both to unscramble aggregates and to prevent further damage by binding to exposed hydrophobic surfaces. Subsequently the existence in all types of cell of a family of unrelated groups of proteins that act as molecular chaperones in a variety of cellular processes was proposed (*2, 3*) and developed (*4–6*). The basic problem remained: Is the rubisco subunit binding protein really a molecular chaperone and how can this be demonstrated?

Table 1. Proteins regarded as molecular chaperones. [Reprinted from (*6*) with permission, © 1990 W. B. Saunders]

Name	Proposed roles
Nucleoplasmins	Nucleosome assembly Transcription? Ribosome assembly and transport?
Chaperonins	Protein folding Oligomer assembly Protein transport DNA replication mRNA turnover Stress protection
Heat shock protein 70 class	Protein folding Oligomer assembly Protein transport Oligomer disassembly?
Heat shock protein 90 class	Masking of hormone receptor
Signal recognition particle	Protein transport
Prosequence of subtilisin	Subtilisin folding
Prosequence of α-lytic protease	α-lytic protease folding
Ubiquitinated ribosomal proteins	Ribosome assembly in eukaryotes
Trigger factor	Protein transpot
Sec B protein	Protein tranport
Pap D protein	Pilus assembly in *E. coli*

Breakthrough

After 1980 my laboratory characterized the binding protein. The protein was purified from pea, barley, and wheat leaves and found to consist of oligomers of about 840 kD, which could be reversibly dissociated by incubation with MgATP into equal amounts of two related types of 60-kD subunit. Antibodies were raised to screen cDNA expression libraries, and the DNA sequence was determined for the α subunit of the binding protein from *Ricinus communis* (castor bean endosperm) and *Triticum aestivum* (wheat leaves). Searching of databases with the *Ricinus* α subunit sequence revealed a high amino acid identity with a nucleotide-derived sequence from *E. coli* (*26*). This sequence was one-third the size of the α subunit and was within the *ams* gene (altered messenger RNA stability). This gene complements an *E. coli* mutant that shows a prolonged messenger RNA half-life, but the function of the protein was unclear. The *ams* protein also shows a high sequence identity to a mycobacterial protein called the 65-kD antigen or common bacterial antigen (*27*). This protein is well studied immunologically since it is the major immunogen in human bacterial infections, but there was no information about its function. The *ams* gene was then found by R. W. Hendrix (*28*) to be identical to part of an *E. coli* gene encoding a well-known polypeptide called the groEL protein. There are striking similarities between the properties and proposed function of the the groEL protein and the rubisco subunit binding protein. The groEL protein was identified in the early 1970s as a bacterial protein required to assemble several phages, including phage λ (*29*). It is an oligomer of 14 60-kD subunits that binds transiently and noncovalently to monomers of phage λ protein B; the complex is stable and can be studied on nondenaturing polyacrylamide gels. The complex is believed to be a necessary intermediate in the formation of an oligomeric structure called the preconnector made of 12 phage B subunits. Phage head proteins assemble upon the preconnector (*30*). In the absence of groEL, the head proteins of phage T4 form insoluble aggregates that associate with the cell membrane (*31*). Rubisco also aggregates into inactive lumps in the absence of the binding protein (*21*). We realized that the binding protein is a member of a ubiquitous family of proteins involved in assembly processes, a view reinforced when another homolog was found in mitochondria (*32, 33*). A joint paper introduced the term "chaperonin" to describe this particular class of molecular chaperone (*34*).

Besides supporting the general concept of molecular chaperones, this paper (*34*) stimulated work on the role of the chaperonins in bacterial rubisco assembly by others. Rubisco from *Rhodospirillum rubrum* consists only of dimers of large subunits. Final proof of the correctness of the chaperone concept with respect to bacterial rubisco was provided by an elegant paper (*35*), which described the use of a purified in vitro assembly system to show that chaperonin prevents incorrect interactions of rubisco large subunits. The key observation is that denatured large subunits self-aggregate into inactive complexes when the denaturant is diluted, but if the appropriate chaperonin is present in the diluting buffer, stable complexes form between the large subunits and the chaperonin; these complexes can be dissociated by adding a related chaperonin and ATP, with the appearance of enzymically active rubisco dimers (*35*). The chaperonins do not convey steric information for rubisco folding since large subunits will assemble correctly in the absence of chaperonins below 15°C (*36*). These two papers (*35, 36*) thus support the basic tenets of the molecular chaperone concept.

There are two messages from this story. The first is technical; researchers studying the synthesis of a particular protein should include nondenaturing techniques in their armory in case the assembly of their protein involves transient noncovalent binding to a chaperone. Secondly, plant molecular biologists need to fight the temptation to confine their attention to the plant literature, a temptation that is increasing as this literature and its attendant conferences expand.

Table 2. Properties of the chaperonins. [Reprinted from (*13*) with permission, © 1990 *Annual Reviews*]

Preferred name	Altenative name	Subunit MW	Number of subunits
Bacterial chaperonin 60 (cpn60)	GroEL 65-kD antigen	57,259 (*E. coli*) 56,686 (*M. leprae*)	14
Bacterial chaperonin 10 (cpn10)	GroES	10,368 (*E. coli*)	7
Mitochondrial chaperonin 60 (cpn60)	hsp60 Mitonin HuCha60	60, 830 (yeast) 57,939 (*H. sapiens*)	14
Plastid chaperonin 60 (cpn60)	Rubisco subunit binding protein	57,393 (*T. aestivum*) 56,453 (*B. napus*)	14

The Molecular Chaperone Concept Today

Molecular chaperones are currently defined as a family of unrelated cellular proteins that mediate the correct assembly of other polypeptides, but are not themselves components of the final structures (*4, 6*). Assembly includes chain folding, oligomerization, and changes in the degree of folding or oligomerization that may take place during the transport and function of the protein. The essential function of molecular chaperones is likely to prevent the formation of incorrect structures that are nonfunctional biologically. Chaperones inhibit unproductive assembly pathways that result from incorrect interactions between parts of polypeptide chains and other molecules, which can include the same or other polypeptides, nucleic acids, or even small metabolites (*37*). This function is required because many cellular processes that involve proteins carry an inherent risk of malfunction due to the intrinsic reactive nature of protein.

In a number of basic cellular processes, interactive protein surfaces are transiently exposed to the intracellular environment. The term "interactive surfaces" refers to any region of intramolecular or intermolecular contact that is significant in maintaining the structure. Such interactive surfaces are important in maintaining both the conformation of monomeric proteins and the quaternary structure of oligomeric proteins, in some cases in combination with other types of molecule such as nucleic acids. Interactive surfaces might be transiently exposed during protein synthesis and protein transport, during reactions such as DNA replication and clathrin cage recycling in which subunit-subunit interactions change, and during the assembly of oligomeric complexes inside organelles from subunits made in more than one subcellular compartment, where the binding propensities of the subunits might need to be reduced before they are all in the same compartment.

The hypothesis of protein self-assembly (*38*) supposes that all the interactions between surfaces exposed in such processes are correct, that is they are both necessary and sufficient to produce the functional conformation. However, the existence of molecular chaperones argues that it is more likely that in any given assembly process there is a certain probability that incorrect interactions will occur that lead to the formation of incorrect structures that are nonfunctional. Where this probability is low, molecular chaperones may not be required, but where it is high they may be essential to allow the formation of sufficient functional structures for cellular needs. The chaperone concept does not contradict the principle of self-assembly; but suggests that in many cases this process needs assistance.

Chaperones also limit damage caused by stresses such as heat, because of their ability to inhibit incorrect interactions that might otherwise occur between partially denatured proteins. Thus some, and perhaps all, heat shock proteins are a subset of molecular chaperones (*39*); however not all chaperones are heat shock proteins.

The chaperones involved in the assembly of nucleosomes (*40*) and bacterial rubisco (*38*) recognize structural features of the protein that are accessible only during early stages of assembly (*35, 36, 41*). The chaperone binds noncovalently to these features and so inhibits incorrect assembly pathways that would act as kinetic dead-end traps. The binding is reversed under circumstances that favor the formation of the correct structure; in the case of some chaperones, this reversal requires ATP hydrolysis.

Table 1 presents a list of molecular chaperones. The family of chaperones is defined in functional terms, while the separate classes within this family are defined on the basis of sequence similarity.

The Chaperonins

Chaperonins are a group of sequence-related chaperones found in all bacteria examined (including the eubacteria, archaebacteria, cyanobacteria, and rickettsiae), in all mitochondria examined (including those from yeast, *Tetrahymena, Drosophila, Xenopus*, maize

Table 3. Proteins related to the chaperonins. Eleven chaperonin 60 protein sequences were used in computer program FASTP, which employs the algorithm of Lipman and Pearson (*53*), to screen for similar proteins. The statistical significance of these scores is determined by the Z value, where $Z > 3$ means possibly significant, $Z > 6$ probably significant, and $Z > 10$ significant. M1, human mitochondria; M2, yeast mitochondria; P1, *Mycobacterium leprae*; P2, *Mycobacterium tuberculosis*; P3, *E. coli*; P4, *Coxiella burnetti*; P5, *Anacystis nidulans*; C1, *Ricinus communis* plastid α; C2, *Triticum aestivum* plastid α; C3, *Brassica napus* plastid α; C4, *Brassica napus* plastid β. Taken from (*54*) with permission.

Protein	Z value										
	M1	M2	P1	P2	P3	P4	P5	C1	C2	C3	C4
DNA-directed RNA polymerase					5	5	4		7	7	4
Myosin heavy chain				7		7		9		12	9
Insulin receptor precursor			3	3	2	3					
TCP-1			16	16		13	11			15	
30S ribosomal protein A	6	7			9	8			4		
Apolipoprotein B100 precursor		6			6				8		7
Cytochrome b			4					5	6	5	
Fusion glycoprotein			9	10			10				7
Heat shock p70 (*Trypanosoma*)		4								6	3
DNA K protein		6			4	5					

leaves, and human cells), and in all plastids examined (including chloroplasts, chromoplasts, and leucoplasts). Table 2 lists some of the properties of the chaperonins. They are all abundant, constitutive proteins that increase to varying extents after stresses such as heat shock or an increase in the cellular content of unfolded protein. In *E. coli* and *Saccharomyces cerevisiae*, these proteins are essential for cell viability at all temperatures (*42, 43*). Bacterial chaperonins are major immunogens during human bacterial infections because of their accumulation during the stress of infection. Since chaperonins occur in human mitochondria they may be involved in autoimmune disease triggered by bacterial infections, a possibility of current medical interest (*44*).

There are two types of chaperonin that are now called chaperonin 60 (cpn60) and chaperonin 10 (cpn10). Cpn60 has a subunit molecular weight in the range 56,000 to 61,000 with amino acid sequence identities of 43 to 54%. The cpn10 sequences show similarity with a region near the NH_2-terminus of the cpn60 sequences (*45*). The cpn10 type (molecular weight, ~10,000) occurs in bacteria and mitochondria (*47*), but not so far in plastids.

The plastid cpn60, like that from mitochondria and bacteria, occurs as a 14-subunit oligomer in crude extracts. When the plastid chaperonin is purified and analyzed on denaturing polyacrylamide gels, two polypeptides, α and β, of very similar mobility can be resolved. The α and β polypeptides occur in equal amounts and show 50% amino acid identity to each other, in the case of *Brassica napus*, but their precise arrangement in the oligomer is unknown (*45*). There is no evidence that α and β subunits occur in the bacterial and mitochondrial cpn60. The oligomeric form of cpn60 is composed of two stacked rings of seven subunits each, with a central hole; this appearance is unusual for an oligomeric protein and makes cpn60 easy to identify by electron microscopy. The oligomeric form of cpn10 is a single ring of seven subunits. The bacterial and plastid cpn60 oligomers respond to added MgATP by hydrolyzing it and dissociating into smaller forms; the cpn10 oligomer binds to the cpn60 oligomer in the presence of MgATP. These responses to MgATP are involved in the mechanism of action of the chaperonins, which is best characterized in the case of rubisco assembly (*13*). The chaperonins mediate the assembly of many proteins in bacteria, mitochondria, and plastids (*6, 33, 39, 42, 47, 48*).

Implications for Plant Research

Microbial cells are widely used to study the function of molecular chaperones because of the ease of analysis and experimentation both genetically and biochemically. However there are several areas of chaperone research where plant systems could be profitably employed.

1) *Determining the structural basis for chaperone binding.* Any member of the chaperonin and hsp70 classes of chaperone can recognize and bind to a feature that is present in a wide but limited range of unrelated proteins, but which is accessible only in the early stages of folding. For example, the plastid chaperonin binds to many but not all chloroplast proteins that are imported from the cytosol (*47*). Elucidating the nature of this feature is a problem similar to that posed by the signal recognition particle, which also shows specific but sequence-independent recognition of signal peptides. The abundance of the plastid chaperonin and rubisco suggests that the complex of plastid cpn60 with the rubisco large subunit should be easy to prepare in large amounts so that attempts can be made to produce crystals for analysis by x-ray diffraction. Such studies would also provide the prelude to another area of plant research—the genetic engineering of rubisco for agricultural purposes.

2) *Improving crop plant rubisco.* Although rubisco from photosynthetic bacteria and cyanobacteria has been expressed in *E. coli* and mutant forms produced, no progress has been made with respect to crop plant rubisco, which seems unable to use the endogenous chaperonins of *E. coli* for correct assembly (*13*). It might be possible to coexpress the plastid chaperonins in the same *E. coli* cells that are making rubisco subunits from the same species as the chaperonins. It may be necessary to coexpress both plastid cpn60 and cpn10 chaperonins for this approach to be successful, but as yet a plastid version of cpn10 has not been reported. A search for this chaperonin should be a high priority, possibly in the context of attempting to extend to crop plant rubisco the chaperonin-mediated renaturation of denatured polypeptides in vitro that has been so successful with the simpler bacterial rubisco (*35, 36, 46*).

3) *Searching for chaperonins outside mitochondria and plastids.* The fact that the bacterial chaperonins play an essential role in mediating the assembly of many cytosolic proteins makes it likely that chaperonins exist in the cytosol of eukaryotic cells to carry out the same basic function. A sequence similarity search of the PIR (NBRF) database with the FASTP program and the sequences of 11 cpn60 proteins shows several proteins with some similarity to the chaperonins (Table 3). The highest similarity is shown by a protein called *t* complex polypeptide 1 (TCP-1); this polypeptide is a product of a gene carried in the *t* locus on chromosome 17 of mice and is associated with changes in spermatogenesis. The TCP-1 protein occurs in the cytosol of all cells of mice but is especially abundant in testis (*49*). A monoclonal antibody against TCP-1 recognizes a protein composed of ~62-kD subunits that occurs in crude soluble extracts of pea leaves but not of chloroplasts or mitochondria (*50*). This protein may be a cytosolic version of the chaperonin. Independent work based on similar sequence studies has reached the same conclusion (*51*). These observations suggest that chaperonins may occur in all compartments where protein assembly needs to be assisted.

4) *Maximizing the expression of foreign proteins in transgenic plants.* There is much interest in using transgenic plants to produce foreign proteins of economic value. In cases where the foreign protein is produced, but lacks the required biological properties due to a failure to assemble correctly in the alien environment, it may be useful to coexpress the appropriate chaperone in the same plant to improve the yield of active product.

REFERENCES AND NOTES

1. J. E. Musgrove and R. J. Ellis, *Philos. Trans. R. Soc. London Ser. B* **313**, 419 (1986).
2. R. J. Ellis, S. M. van der Vies, S. M. Hemmingsen, in *Plant Molecular Biology*, NATO-ASI series A (Plenum, New York, 1987), vol. 140, p. 33.
3. R. J. Ellis, *Nature* **328**, 378 (1987).
4. _____ and S. M. Hemmingsen, *Trends Biochem. Sci.* **14**, 339 (1989).
5. R. J. Ellis, S. M. van der Vies, S. M. Hemmingsen, *Biochem. Soc. Symp.* **55**, 145 (1989).
6. R. J. Ellis, *Semin. Cell Biol.* **1**, 1 (1990).
7. S. Lecker *et al.*, *EMBO J.* **8**, 2703 (1989).
8. J. E. Rothman, *Cell* **59**, 591 (1989).
9. R. P. Beckmann, L. A. Mizzen, W. J. Welch, *Science* **248**, 850 (1990).
10. A. L. Horwich, W. Neupert, F.-U. Hartl, *Trends Biotechnol.* **8**, 126 (1990).
11. R. J. Ellis, *Trends Biochem. Sci.* **4**, 241 (1979).
12. T. J. Andrews and G. C. Lorimer, in *The Biochemistry of Plants* (Academic Press, New York, 1987), vol. 10, p. 131.
13. A. A. Gatenby and R. J. Ellis, *Annu. Rev. Cell Biol.* **6**, 125 (1990).
14. R. J. Ellis and J. C. Gray, *Phil. Trans. R. Soc. London Ser. B* **313**, 303 (1986).
15. K. Keegstra, L. J. Olsen, S. M. Theg, *Annu. Rev. Plant Physiol. Mol. Biol.* **40**, 471 (1989).
16. C. Dean, E. Pichersky, P. Dunsmuir, *ibid.* **40**, 415 (1989).
17. R. Barraclough and R. J. Ellis, *Biochim. Biophys. Acta* **608**, 19 (1980).
18. R. J. Ellis and S. M. van der Vies, *Photosyn. Res.* **16**, 101 (1988).
19. G. E. Blair and R. J. Ellis, *Biochim. Biophys. Acta* **319**, 223 (1973).
20. S. M. Hemmingsen, *Semin. Cell Biol.* **1**, 47 (1990).
21. G. Voordouw, S. M. van der Vies, P. P. Bouwmeister, *Eur. J. Biochem.* **141**, 131 (1984).
22. T. E. Creighton, *The Proteins* (Freeman, New York, 1984), p. 312.
23. H. Roy, *Plant Cell* **1**, 1035 (1989).
24. R. A. Laskey, B. M. Honda, A. D. Mills, J. T. Finch, *Nature* **275**, 416 (1978).
25. H. R. B. Pelham, *Cell* **46**, 989 (1986).

26. P. K. Chanda, M. Ono, M. Kuwano, H. F. Kung, *J. Bacteriol.* **161**, 446 (1985).
27. T. M. Shinnick *et al.*, *Infect. Immun.* **56**, 446 (1988).
28. R. W. Hendrick, unpublished data.
29. C. P. Georgopoulos *et al.*, *J. Mol. Biol.* **76**, 645 (1973); M. Sternberg, *ibid.*, p. 1.
30. J. Kochan and H. Murialdo, *Virology* **131**, 100 (1983).
31. T. Takano and T. Kakefuda, *Nature New Biol.* **239**, 34 (1972).
32. T. W. McMullin and R. L. Hallberg, *Mol. Cell Biol.* **8**, 371 (1988).
33. R. L. Hallberg, *Semin. Cell Biol.* **1**, 37 (1990).
34. S. M. Hemmingsen *et al.*, *Nature* **333**, 330 (1988).
35. P. Goloubinoff *et al.*, *ibid.* **342**, 884 (1989).
36. P. V. Viitanen *et al.*, *Biochemistry* **29**, 5665 (1990).
37. R. J. Ellis, *Nature* **345**, 484 (1990).
38. C. B. Anfinsen, *Science* **181**, 223 (1973).
39. R. T. Morimoto, A. Tissieries, C. Georgopoulos, Eds., *Stress Proteins in Biology and Medicine* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1990).
40. S. M. Dilworth and C. Dingwall, *BioEssays* **2**, 44 (1988).
41. S. Lecker *et al.*, *EMBO J.* **8**, 2703 (1989).
42. C. Georgopoulos and D. Ang, *Semin. Cell Biol.* **1**, 19 (1990).
43. D. S. Reading *et al.*, *Nature* **337**, 655 (1989).
44. D. B. Young, *Semin. Cell Biol.* **1**, 27 (1990).
45. R. Martel, L. P. Cloney, L. E. Pelcher, S. M. Hemmingsen, *Gene*, in press.
46. T. H. Lubben *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
47. T. H. Lubben *et al.*, *Plant Cell* **1**, 1223 (1989).
48. M. Y. Cheng *et al.*, *Nature* **337**, 620 (1989); J. Ostermann *et al.*, *ibid.* **341**, 125 (1989); A. L. Horwich *et al.*, *Trends Biotechnol.* **8**, 126 (1990); F.-U. Hartl *et al.*, *Science* **247**, 930 (1990).
49. K. Willison *et al.*, *Cell* **57**, 621 (1989).
50. G. M. F. Watson, N. H. Mann, R. J. Ellis, unpublished data.
51. R. S. Gupta, *Biochem. Int.* **4**, 833 (1990).
52. R. A. Johnson, thesis, University of Warwick, United Kingdom (1987).
53. D. J. Lipman and W. R. Pearson, *Science* **227**, 1435 (1985).
54. S. M. van der Vies, thesis, University of Warwick, United Kingdom (1989).

The Cauliflower Mosaic Virus 35*S* Promoter: Combinatorial Regulation of Transcription in Plants

PHILIP N. BENFEY AND NAM-HAI CHUA

Appropriate regulation of transcription in higher plants requires specific cis elements in the regulatory regions of genes and their corresponding trans-acting proteins. Analysis of the cauliflower mosaic virus (CaMV) 35*S* promoter has contributed to the understanding of transcriptional regulatory mechanisms. The intact 35*S* promoter confers constitutive expression upon heterologous genes in most plants. Dissection into subdomains that are able to confer tissue-specific gene expression has demonstrated that the promoter has a modular organization. When selected subdomains are combined, they confer expression not detected from the isolated subdomains, suggesting that synergistic interactions occur among cis elements. The expression patterns conferred by specific combinations of 35*S* subdomains differ in tobacco and petunia. This indicates that a combinatorial code of cis-regulatory elements may be interpreted differently in different species.

RECENT STUDIES OF GENES EXPRESSED IN A TISSUE-SPECIFIC manner (*1*) and genes that control specific developmental pathways (*2*) have revealed the importance of transcriptional mechanisms in the control of development in higher plants. In some cases, the DNA sequence elements that are necessary for transcriptional regulation and the protein factors that interact with these sequences have been identified. What has emerged is a complex picture in which DNA sequence elements that are important for regulation are scattered over thousands of base pairs (bp), and these elements interact with factors that can be either ubiquitous or highly restricted in their distribution. A simple model in which transcriptional regulation is mediated solely by the presence or absence of a particular trans-acting factor now seems inadequate. Rather, transcriptional regulation may be accomplished through combinatorial mechanisms (*3, 4*) by which diverse expression patterns are achieved through different combinations of a limited number of regulatory elements and trans-acting factors.

Laboratory of Plant Molecular Biology, The Rockefeller University, New York, NY 10021.

If combinatorial processes control transcription, we should be able to identify the basic components and generate different transcription patterns when the components are combined in different ways. In order to observe the effects of different combinations of the basic components, gene expression must be monitored in a variety of tissues and throughout development. Plants are particularly well suited to this type of analysis, because transgenic plants can be rapidly and easily generated, and reporter gene expression can be monitored in most cells and at various stages of development. We have used a viral promoter that is able to confer expression in most plant cells as a model system to dissect some of the combinatorial properties of transcriptional control in plants. In this article, we review our current knowledge of the components within this promoter and the results of combining these components in different ways. We contrast expression patterns conferred by various combinations of cis elements in flowers from two different species. We also compare these findings with evidence for combinatorial mechanisms of gene regulation obtained from plants and other organisms.