ANNUAL REVIEW OF CELL BIOLOGY

VOLUME 5, 1989

GEORGE E. PALADE, *Editor*
Yale University School of Medicine

BRUCE M. ALBERTS, *Associate Editor*
University of California, San Francisco

JAMES A. SPUDICH, *Associate Editor*
Stanford University School of Medicine

ANNUAL REVIEWS INC. 4139 EL CAMINO WAY P.O. BOX 10139 PALO ALTO, CALIFORNIA 94303-0897

ANNUAL REVIEWS INC.
Palo Alto, California, USA

International Standard Serial Number : 0743-4634
International Standard Book Number : 0-8243-3105-2

TYPESET BY AUP TYPESETTERS (GLASGOW) LTD., SCOTLAND
PRINTED AND BOUND IN THE UNITED STATES OF AMERICA

PREFACE

The recent history of Cell Biology has been characterized by rapid and functionally significant growth in a large number of unusually productive directions. The oldest and strongest among them has already led to the merging of the once distinct fields of Cell Biology and Molecular Biology into a continuous, common body of knowledge kept in active ferment by concepts and technologies derived from molecular genetics as well as from other sources.

The more we advance in developing this basic common body of knowledge, the more clearly we appreciate the remarkable versatility of the evolutionary process and the detailed and subtle richness of information contained within the cell's genome. Evolution has exploited minor changes as well as major rearrangements to create families or superfamilies of genes that encode for a variety of gene products with many common structural motifs yet diversified functions. The number of protein isoforms is continuously increasing, and the old notion that gene products are stable through the ontogeny of individual organisms is being revised, because of the established existence of multiple isoforms, products of single genes or gene families, differentially expressed at specific sites (organs or parts of organs), at specific times in ontogeny. It seems that interactions with the environment (even with microenvironments on homeostatic organisms) influence the expression of the genome to a larger extent than previously assumed. It is already well established that the genome includes instructions for the expression of specific sets of genes in specific cell types, but perhaps we should already search for instructions that control the expression of cell-type specific genes under specific environmental conditions in ontogeny or in pathological processes in the adult. In some of the latter, fetal gene products are known to be re-expressed.

For one reason or another, less attention is given to "junk DNA" at present than a few years ago, and less resistance is encountered by major projects, such as the sequencing of the human and other genomes of immediate interest. These genomes may prove to contain less "evolutionary garbage" than when the projects were first envisaged. Yet much more structural and chemical information is needed to understand in specific terms, rather than in general outline, how gene expression and genome replication are controlled. The cognate research areas are already, and will remain for some time, domains of critical importance in contemporary Cell Biology.

Periods of rapid scientific advances often generate illusions. Impressed somehow by the massive influx of new information, we may be inclined to

believe that we are close to a full understanding of how cells are put together and carry out their functions. But often we discover that only a few years ago we were soaring over large areas of significant structures and activities we had blissfully ignored. And suddenly such areas become the equivalent of gold mining fields with the expected concentration of eager search forces, competition, frictions, claims, and counter claims. Protein traffic control in all cells, targeting, sorting, membrane traffic control in eukaryotic cells, and — connected with such operations — quality control of gene products appear to be among these recently opened gold mining fields. Signal transduction mechanisms represent a similar area, since so much remains to be uncovered between activated receptors and signal amplifiers on the distal side of the relevant chain of reactions and activated gene transcription or genome replication on the proximal end of the same chain. Equally attractive areas can be recognized in research fields dealing with the locomotor apparatus of the cell or the cell's interactions with its immediate environment, that is, with other cells and macromolecular components of the extracellular matrix. Each of these new and often unexpected developments is leading to spectacular advances in Cellular and Molecular Biology.

Yet there is still another area that may become in time even more fertile than those already mentioned. The genome and the basic organization at the cellular level is the same all the way from a fertilized egg to a fully mature organism that comprises a huge number of cells belonging to many highly differentiated cell types. All this complex but exquisitely controlled developmental process is guided by programmed genetic instructions that call for the expression of specific genes at specific sites and specific times in ontogeny. The basic mechanisms involving messengers, receptors, signal transduction, and other components of the relevant chains of reactants appear to be similar to those found in the cells of adult organisms, but the specific macromolecules that operate along these chains appear to be themselves differentiated. So far the best examples are found in terminal cell differentiation, especially in the immune system, that appears to have its own set of primary messengers — the interleukins — and its own set of receptors. We know little about the chain of reactants that connects activated receptors to the relevant programs of specific genes, but we know that the cognate gene products are cell-type specific and highly diversified. How many genes, we may ask, are involved in the construction of those elements of the immune system we have so far identified? And how many genes and gene products remain to be uncovered in the main cell types of the system or in their supporting cast represented by stromal cells, endothelial cells, and specialized epithelial cells?

It is worthwhile pointing out that work on the developmental history of

the immune system has advanced faster than work on the development of any other complex organismic counterpart. This situation may be explained by some of the characteristics of the immune system: it consists of sortable populations of essentially free living cells, free because they have to patrol the organism in search of targets and interacting partners. And it relies on diffusible primary messengers for most cells interactions. Each of these characteristics has been used to generate tools and assays that have greatly facilitated research on the development of the entire system. The complexity of developmental interactions may be higher in other cases, especially in the central nervous system. Moreover, the uncovering of the guiding mechanisms may prove to be a much more arduous operation: the cells in case often form compact, cohesive parenchyma, and development seems to rely primarily on direct cell-cell interactions, bypassing the need for diffusible primary messengers. In this case (and undoubtedly many others), messengers as well as receptors appear to be integral membrane proteins, some of them related to cell receptors of the immune system as members of a large gene superfamily.

Starting from such considerations, we may ask: how many genes are involved in the construction of a sensory neuron that connects the periphery to the cerebellum? How many genes are needed to build up a cerebellum and how many to generate an entire central nervous system? The anticipated answers are: very large numbers. This may become, in fact, a recurrent theme: how many genes contribute to the development of a nephron and how many to that of an entire kidney, its specialized vasculature included? Perhaps, after all, a good fraction if not most of that DNA in the genome is functionally needed.

The Editorial Committee of the Annual Review of Cell Biology has tried to provide in this volume a useful sampling of current activities in many of the interesting and challenging research areas mentioned above. As the table of contents shows, the sampling includes topics belonging to developmental biology, cellular immunology, and plant cell biology, in addition to topics of general interest to cellular and molecular biologists. The readers will tell us, we hope, if they consider our attempts successful. Comments and suggestions coming from the reading public, including the membership of the American Society for Cell Biology, will be most welcome.

GEORGE E. PALADE
EDITOR

Annual Review of Cell Biology
Volume 5, 1989

CONTENTS

CONTROL OF PROTEIN EXIT FROM THE ENDOPLASMIC RETICULUM, *Hugh R. B. Pelham* 1

ORIGIN AND EVOLUTION OF MITOCHONDRIAL DNA, *Michael W. Gray* 25

THE CHLOROPLAST CHROMOSOMES IN LAND PLANTS, *Masahiro Sugiura* 51

TENASCIN: AN EXTRACELLULAR MATRIX PROTEIN PROMINENT IN SPECIALIZED EMBRYONIC TISSUES AND TUMORS, *Harold P. Erickson and Mario A. Bourdon* 71

GROWTH FACTORS IN EARLY EMBRYOGENESIS, *D. A. Melton and Malcolm Whitman* 93

DYNEIN STRUCTURE AND FUNCTION, *Mary E. Porter and Kenneth A. Johnson* 119

COMMUNICATION BETWEEN MITOCHONDRIA AND THE NUCLEUS IN THE REGULATION OF CYTOCHROME GENES IN THE YEAST *SACCHAROMYCES CEREVISIAE*, *Susan L. Forsburg and Leonard Guarente* 153

MOLECULAR GENETICS OF DEVELOPMENT STUDIED IN THE TRANSGENIC MOUSE, *Heiner Westphal and Peter Gruss* 181

INITIATION OF EUKARYOTIC DNA REPLICATION IN VITRO, *Bruce Stillman* 197

LIPID TRAFFIC IN ANIMAL CELLS, *Gerrit van Meer* 247

PROTEIN OLIGOMERIZATION IN THE ENDOPLASMIC RETICULUM, *Stella M. Hurtley and Ari Helenius* 277

DROSOPHILA EXTRACELLULAR MATRIX, *J. H. Fessler and L. I. Fessler* 309

SIMPLE AND COMPLEX CELL CYCLES, *Fred Cross, James Roberts, and Harold Weintraub* 341

THE INTERLEUKIN 2 RECEPTOR, *Kendall A. Smith* 397

BIOGENESIS OF THE RED BLOOD CELL MEMBRANE-SKELETON AND THE CONTROL OF ERYTHROID MORPHOGENESIS, *Elias Lazarides and Catherine Woods* 427

(continued)

MEMBRANE TRAFFIC IN ENDOCYTOSIS: INSIGHTS FROM CELL-FREE ASSAYS, *Jean Gruenberg and Kathryn E. Howell* 453

THE BIOGENESIS OF LYSOSOMES, *Stuart Kornfeld and Ira Mellman* 483

INDEXES

Subject Index 527

Cumulative Index of Contributing Authors, Volumes 1–5 536

Cumulative Index of Chapter Titles, Volumes 1–5 538

Annu. Rev. Cell Biol. 1989. 5. 1–23

CONTROL OF PROTEIN EXIT FROM THE ENDOPLASMIC RETICULUM

Hugh R. B. Pelham

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

CONTENTS

INTRODUCTION 1
Outline of the Problem 2
PROTEIN EXPORT FROM THE ER 3
Nonselective Bulk Flow from the ER 3
The Rate of Bulk Flow 4
Factors Affecting the Rate of Protein Export 4
PROTEIN FOLDING: THE ROLE OF RESIDENT ER PROTEINS 5
Binding of Proteins to BiP 5
Action of BiP 6
Binding of Proteins to Other ER Residents 7
SORTING OF RESIDENT ER PROTEINS FROM SECRETED PROTEINS 8
Membrane Versus Soluble Proteins 8
Existence and Function of Luminal ER Proteins 9
Solubility of Luminal ER Proteins 9
A Common Signal on Soluble ER Proteins 11
Function of KDEL: A Retrieval Mechanism 12
Likely Site of Sorting: The Salvage Compartment 13
Retention of Soluble ER Proteins in Saccharomyces Cerevisiae 15
Mutants 16
IMPLICATIONS OF BULK FLOW AND RETRIEVAL 16
PROSPECTS 18

INTRODUCTION

The endoplasmic reticulum (ER) is the largest membrane-bound organelle in a typical eukaryotic cell. It consists of a continuous network of tubules and cisternae extending throughout the cytoplasm, with a total surface

0743-4634/89/1115-0001$02.00

area at least six times that of the plasma membrane. Most of the ER membrane is studded with ribosomes (thus forming the rough endoplasmic reticulum, or RER). Membrane and secretory proteins are synthesized on these ribosomes and enter the ER, which is the starting point for the secretory pathway; such proteins spend only a short time in this organelle before being transported, by a process of vesicle budding and fusion, to the Golgi apparatus and thence to the cell surface (Palade 1975). The ER membrane and the lumen that it encloses also contain a characteristic set of resident proteins that are involved in the processing of secretory proteins and in other metabolic functions such as phospholipid biosynthesis. Thus, export of newly synthesized polypeptides from the ER involves the separation of secretory proteins from resident proteins. In this article I discuss the mechanisms involved in this sorting process. Several other recent reviews have also covered this area (Pfeffer & Rothman 1987; Rothman 1987; Burgess & Kelly 1987; Rose & Doms 1988; Lodish 1988; Cutler 1988).

Outline of the Problem

Secretory and membrane proteins leave the ER in transport vesicles, which may bud from specialized ribosome-free regions of the ER membrane adjacent to the Golgi, termed transitional elements. They are delivered to the *cis* face of the Golgi apparatus and then continue their journey through the Golgi cisternae by means of repeated cycles of vesicle budding and fusion. The ER and the various regions of the Golgi differ in their content of oligosaccharide modifying enzymes, which allow the transport of glycoproteins through these compartments to be monitored by the sequential changes that occur to their oligosaccharide side chains (for reviews see Farquhar 1985; Kornfeld & Kornfeld 1985; Roth 1987). In general, glycoproteins found in the ER do not exhibit Golgi-specific modifications. This observation suggests that vesicular transport from ER to Golgi is both unidirectional and selective: most resident ER proteins do not reach the Golgi and once a secretory protein has reached the Golgi, it does not return to the ER.

Selective transport implies that individual proteins bear signals that control their movement. In principle, each step along the secretory pathway could be mediated by a specific transport signal and the destination of a protein determined by its lack of a signal for the next step. Alternatively, secretion could be the default pathway, driven by a nonselective process in which the contents of each compartment are transferred to the next; diversion from this pathway or retention in a specific compartment would then be signal-mediated (Pfeffer & Rothman 1987). There is now considerable evidence that a nonselective bulk flow to the cell surface does

exist, and those signals that have been most clearly identified divert proteins from this flow: the mannose-6-phosphate marker directs soluble lysosomal enzymes to lysosomes (reviewed by von Figura & Hasilik 1986; Kornfeld & Mellman, this volume) and, as discussed below, the *C*-terminal tetrapeptide KDEL causes the retention of resident soluble proteins in the lumen of the ER (Munro & Pelham 1987).

PROTEIN EXPORT FROM THE ER

Nonselective Bulk Flow from the ER

The selectivity of export from the ER could be explained if this step was receptor-mediated. The most extreme version of this hypothesis predicts that every exported protein contains a transport signal and that disruption of this signal by mutation would prevent transport. A number of workers have searched for such signals by mutating the genes encoding secretory or membrane proteins and measuring the rate of transport of the altered proteins. Many mutant proteins do fail to be secreted, but it seems that in most cases this is a nonspecific effect: the proteins do not fold correctly, and malfolded proteins, in general, are slow to leave the ER, often because they form insoluble aggregates (reviewed by Rose & Doms 1988).

On the other hand, there are several cases where proteins that do not normally leave the ER are induced to do so by the deletion of certain amino acid sequences. Thus, removal of the *C* terminus from the resident ER Protein BiP (binding protein) (Munro & Pelham 1987), of the *N* terminus of the rotavirus protein VP7 (Poruchynsky et al 1985), or of the cytoplasmic tail of the adenovirus E19 protein, a transmembrane protein (Paabo et al 1987), causes each of them to enter the secretory pathway, even though they do not normally do so and therefore would not be expected to possess a transport signal. Similarly, expression of the prokaryotic protein β-lactamase in *Xenopus* oocytes also results in secretion although this enzyme is not expected to have signals for transport to the Golgi complex (Wiedmann et al 1984). Most strikingly, Wieland et al (1987) have shown that a synthetic tripeptide consisting of the glycosylation sequence Asn-Tyr-Thr, esterified to make it membrane-permeable, can enter cells, reach the ER, be glycosylated (thus making it membrane-impermeant), and then be rapidly secreted. The existence of appropriate modifications to the attached oligosaccharide demonstrates that at least some of the tripeptide passes through the Golgi complex on the normal secretory pathway. These results argue strongly in favor of a bulk flow model in which secretion is the default fate for a protein containing no specific signals.

The Rate of Bulk Flow

The halftime for exit of individual secretory or membrane proteins from the ER varies from about 15 min for viral glycoproteins (Quinn et al 1984; Copeland et al 1988) and some serum proteins to 2 h or more for other serum proteins (Fries et al 1984; Lodish et al 1983; Yeo et al 1985). The most rapidly secreted proteins reach the cell surface with a halftime of about 30 min. Is bulk flow sufficiently rapid to account for the export of these proteins? Wieland et al (1987) argued that it is: they estimated the halftime for secretion of the tripeptide from tissue culture cells to be about 10 min. However, this rate is not easy to measure, and the validity of this estimate has been questioned (Rose & Doms 1988). One problem is that the kinetics of glycosylation of the added tripeptide are distinctly biphasic, for unknown reasons, and examination of the published data suggests that halftimes of 25–50 min are equally possible.

There are other uncertainties: it is not clear whether the glycosylated tripeptide is truly a bulk marker, or whether it can bind to proteins within the secretory pathway. Also, its nonphysiological nature, while desirable for a bulk phase marker, raises doubts about the precise way in which it is metabolized by cells and the effects it might have on them. Further measurements using a different marker would clearly be useful. Nevertheless, the current data suggest that the bulk flow rate is close to the rate of transport of the most rapidly secreted proteins. If this is so, there is no compelling reason to postulate the existence of a positive transport signal for any protein.

Factors Affecting the Rate of Protein Export

Many proteins leave the ER more slowly than expected from the presumed bulk flow rate. The most likely explanation for this is that they interact with resident proteins in the ER, and are retarded in a manner analogous to chromatography (Pfeffer & Rothman 1987). A rapidly secreted protein would then be characterized by an absence of interactions rather than a dominant signal for transport. In principle, this can be tested by constructing fusions between rapidly and slowly exported proteins. For example, the soluble ER protein BiP, when its *C* terminus is removed, is secreted very slowly from COS cells ($t_{1/2}$ approx. 3 h; Munro & Pelham 1987); cathepsin D leaves the ER much more rapidly (Pelham 1988). A fusion protein containing both BiP and cathepsin D sequences leaves the ER at the slow rate characteristic of the truncated BiP (H. Pelham, unpublished observations), which is consistent with the idea that the low rate of export of truncated BiP is due to interactions with ER proteins rather than the lack of a "rapid transport" signal.

Membrane proteins seem to be particularly prone to hindrance A large number of mutations have been introduced into such proteins as the VSV G protein and influenza hemagglutin (HA), and in many cases transport is impaired Changes in the luminal, cytoplasmic, and transmembrane domains can all slow export and no general signal has been identified for either retention or transport (for extensive discussion of these experiments see Rose & Doms 1988) It seems that each protein is a special case For example, VSV G, influenza HA, and the adenovirus E19 protein all have a single transmembrane segment and a small cytoplasmic tail Removal of this tail has no effect on the rapid transport of HA (Doyle et al 1986), greatly slows VSV G export (Doms et al 1988), and promotes the export of E19, which normally resides in the ER (Paabo et al 1987)

PROTEIN FOLDING THE ROLE OF RESIDENT ER PROTEINS

It has become apparent that the correct folding and assembly of secretory proteins is necessary for their efficient transport, presumably because unfolded proteins tend to bind to other proteins in the ER or form aggregates that are unable to enter transport vesicles (for discussion see Rose & Doms 1988, Lodish 1988, Helenius & Hurt, this volume) Folding of proteins in vitro is often a slow and inefficient process, but in vivo most proteins achieve their final conformation within a few minutes This efficiency is due at least in part, to the presence of several resident proteins of the ER that actively promote the folding process In addition, these resident proteins probably contribute to the selectivity of transport by binding (even weakly) to incompletely or incorrectly assembled proteins they may prevent their export

Binding of Proteins to BiP

A major ER resident involved in protein assembly is the soluble protein BiP (binding protein, also known as the glucose-regulated protein GRP78). Increased synthesis of BiP is induced when abnormal proteins accumulate in the ER (Lee 1987, Kozutsumi et al 1988), and BiP preferentially associates with such proteins Thus, it binds to mutant or malfolded forms of influenza HA (Gething et al 1986), to a derivative of SV40 T antigen that is fused to a signal sequence and thus enters the ER (Sharma et al 1985) and, in an in vitro translation-translocation system, to unoxidized (but not mature, disulphide-linked) prolactin and to unglycosylated, but not glycosylated, invertase (Kassenbrock et al 1988) Many proteins fold poorly or aggregate when their glycosylation is inhibited (e g Leavitt et al 1977), and in such cases association with BiP is frequently observed (Bole et al

1986; Gething et al 1986, Dorner et al 1987; Hendershot et al 1988). Binding of BiP to all these substrates is hydrophobic in nature, which suggests that they are recognized, at least in part, by the presence of exposed hydrophobic residues that are buried in the mature, properly folded protein.

Although many of the known substrates for BiP are aberrant proteins that remain in the ER, secretory proteins can also interact with BiP before they achieve their mature state. For example, BiP binds transiently to at least some immunoglobulin heavy chains prior to their association with light chains (Haas & Wabl 1983, Bole et al 1986), and to several human serum glycoproteins expressed in CHO cells (Dorner et al 1987). With the recent cloning of the yeast BiP gene, it has become possible to test the importance of BiP in the normal process of protein transport to the Golgi complex. Preliminary results indicate that BiP is essential for viability, and that within 15 min of warming a temperature sensitive BiP mutant to the nonpermissive temperature, import of proteins into the ER ceases (Rose & Misra 1989, L. Moran, personal communication, M. Rose, personal communication). These results suggest that BiP interacts with a variety of nascent proteins and may be required for some of them to complete their translocation into the ER.

Action of BiP

BiP is closely related to the cytoplasmic heat shock protein hsp70 and is likely to act in a similar way (Munro & Pelham 1986). Like hsp70, BiP binds tightly to ATP and ATP hydrolysis allows the protein to be released from at least some of the substrates to which it binds (Munro & Pelham 1986, Dorner et al 1987, Kassenbrock et al 1988). A simple working model for the function of these proteins is that they act as reversible detergents: they bind to hydrophobic surfaces, but at intervals use the energy of ATP hydrolysis to change their conformation to a nonbinding state (Lewis & Pelham 1985, Pelham 1986). This would have the effect of maintaining unfolded proteins in solution without sequestering them permanently in a nonfunctional complex, the proteins could thus avoid aggregation or precipitation but still be able to achieve their final tertiary and quaternary structures.

That such a detergent-like role could be useful is shown by studies of the effects of detergents on protein folding in vitro. For example, the enzyme rhodanese is composed of two domains with hydrophobic surfaces that normally interact with each other, but during refolding they tend to bind to other molecules and form nonproductive aggregates, so that very little active enzyme is obtained. Addition of the detergent lauryl maltoside weakens these nonproductive interactions and greatly increases the yield of correctly refolded enzyme (Tandon & Horowitz 1986).

Although the main function of BiP is probably to promote protein assembly, it may also serve to prevent export of misfolded proteins from the ER; such proteins remain associated with BiP until they either fold correctly or are degraded. It is not clear whether BiP is required for the retention of aberrant proteins, because most of them would probably aggregate and thus fail to be secreted even if BiP were absent. However, Dorner et al (1988) reported that the secretion of an overexpressed derivative of human tissue plasminogen activator was improved when the level of BiP was specifically reduced by the expression of anti-sense RNA. In this case, at least, it seems that interaction with BiP can limit the export of a secretory protein from the ER.

Binding of Proteins to Other ER Residents

Another ER resident that interacts with unfolded proteins is the enzyme protein disulphide isomerase (P.D.I.), which catalyses thiol oxidation and disulphide exchange reactions (Freedman 1984). P.D.I. is a remarkable protein with several distinct roles. Besides existing as a free monomer, it is also an essential subunit of prolyl 4-hydroxylase, an enzyme that catalyses the modification of prolyl residues in newly synthesized collagen (Pihlajaniemi et al 1987). Furthermore, it binds to the Asn-X-Ser/Thr acceptor sequence for *N*-linked glycosylation and is an important component of the oligosaccharide transferase (Geetha-Habib et al 1988). All forms of P.D.I. are soluble, or only loosely associated with the ER membrane, and are released from isolated microsomes when these are ruptured by incubation at high pH. P.D.I.-depleted microsomes can still import γ-gliadin (a nonglycosylated wheat storage protein), but the protein cannot achieve its correct disulphide-bonded state: reconstitution of microsomes in the presence of purified P.D.I. restores this function (Bulleid & Freedman 1988). P.D.I. can also be cross-linked to nascent immunoglobulin chains in vivo (Roth & Pierce 1987). These results indicate that P.D.I. interacts directly with newly synthesized secretory proteins and is required for their correct folding.

Other proteins may also help to lubricate the process of protein assembly in the ER and, perhaps at the same time, prevent premature export. Possible examples include a membrane protein termed CD-ω or TRAP, which associates transiently with partially assembled T-cell receptor-CD3 complexes (Alarcon et al 1988; Bonifacino et al 1988); the collagen-binding protein colligin (Hughes et al 1987; Saga et al 1987); and the abundant luminal protein GRP94, which is related to the cytoplasmic heat shock protein hsp90 and, like BiP, is synthesized at higher rates when aberrant proteins accumulate in the ER (Sorger & Pelham 1987; Lee 1987).

SORTING OF RESIDENT ER PROTEINS FROM SECRETED PROTEINS

Membrane Versus Soluble Proteins

Resident proteins of the ER must avoid export by the bulk flow pathway. For membrane proteins it is easy to imagine that residence results from the same features that slow the export of mutant proteins, in particular the tendency to form large aggregates. They may also be held in position by interactions between their cytoplasmic domains and other cellular structures. For example, the "crystalloid" ER that forms in cells expressing high levels of the enzyme HMG-CoA reductase consists of membrane tubules that interact extensively with each other, presumably via the cytoplasmic domains of integral membrane proteins (Anderson et al 1983). Similarly, the inner membrane of the nuclear envelope contains glycoproteins that are associated with the nuclear lamina (Senior & Gerace 1988). Major components of the RER, including the ribophorins and the signal peptidase complex, form an aggregate that resists disruption by nonionic detergents (Crimaudo et al 1987). These components should be further cross-linked in vivo as a result of their association with polysomes and are probably unable to enter transport vesicles.

Whether all ER membrane proteins are associated with some large structure remains to be seen. Few interactions have been studied in detail, although some progress has been made. For example, in adenovirus-infected cells the class I histocompatibility antigens remain in the ER because they are bound to the adenovirus E19 protein (Burgert & Kvist 1985; Andersson et al 1985; Severinsson & Peterson 1985). Retention of E19 in the ER is mediated by its short cytoplasmic tail (Paabo et al 1987), but the protein that interacts with this tail has not been identified. Another well-studied example is the VP7 polypeptide of the rotavirus SA11. Retention of this protein in the ER is mediated by the first 60 amino acids of the mature protein, which can form an amphipathic helix (Poruchynsky & Atkinson 1988). Surprisingly, the VP7 signal sequence is also required, even though it is cleaved from the mature protein: replacement with the HA signal leads to cleavage at the identical site, but also to secretion of VP7 as a soluble protein (Stirzaker et al 1987; Stirzaker & Both 1989). It seems that the formation of a membrane anchor requires the interaction of the VP7 signal sequence with the *N* terminus of the mature protein. Although the precise way in which VP7 becomes associated with the membrane is obscure, it is clear that it is this association that is responsible for its ER location.

For truly soluble luminal proteins, retention represents a much greater conceptual problem. If such proteins spend at least part of their time as

diffusible entities of small size, they will inevitably tend to be transported by the bulk flow pathway. A special mechanism must therefore exist to prevent their loss from the cell. The following sections outline the evidence for the existence of soluble resident proteins in the ER and for a specific retention mechanism that keeps them there.

Existence and Function of Luminal ER Proteins

Table 1 lists some currently known luminal ER proteins, several of which have been independently discovered and named by different groups. Many of them are thought to bind to or act on newly synthesized secretory proteins: these include P.D.I., BiP, GRP94, and calligin, and also glucosidase II, which trims glucose residues from the oligosaccharide side chains of glycoproteins. There are a number of carboxyesterases that probably play a role in the detoxification of aromatic compounds; these include the gut esterase of *C. elegans*, the 60-kd liver esterases listed in Table 1, and other liver esterases such as egasyn (Brown et al 1987) and E1 (Harano et al 1988). Other proteins have less obvious functions: ERp72 is known only as an ER protein, while reticulin is a calcium-binding protein (as are the acidic proteins P.D.I., BiP and GRP94, at high calcium concentrations), and perhaps plays a role analogous to that of calsequestrin in the sarcoplasmic reticulum (Macer & Koch, 1988).

Solubility of the Luminal ER Proteins

The evidence that most of the proteins listed in Table 1 are not membrane-associated is of three main types. First, they are soluble in the absence of detergents, or in the presence of detergent concentrations that are too low to solubilize true integral membrane proteins (Strous & van Kerkhof 1989; Koch et al 1988; Bulleid & Freedman 1988; Bole et al 1986; Brown et al 1987). Second, glucosidase II, GRP94, P.D.I. and BiP have been localized by immunogold labeling at the EM level: they are not concentrated at the membrane but appear spread evenly across the lumen of the ER (Lucocq et al 1986; Koch et al 1988; Akagi et al 1988; J. Tooze, personal communication). Third, all of them apart from glucosidase II have been sequenced and they show no obvious transmembrane segments (see Table 1 for references). One report has suggested the presence of a single transmembrane segment in GRP94 (ERp99; Mazzarella & Green 1987). However, the 21-amino acid segment in question contains two charged residues and is part of an extended region of homology between GRP94 and the soluble protein hsp90. Given the other properties of GRP94, there seems no reason to believe that it spans the membrane (Koch et al 1988).

These arguments do not exclude the possibility that the proteins are attached to some immobile matrix in the ER lumen. This point has been

Table 1 Some luminal ER proteins

Protein	Other names	Species	*C* terminus	Refs.[e]
P.D.I.		rat	QKAV KDEL	1
	ERp59	mouse	QKAV KDEL	2
	Prolyl 4-hydroxylase[a]	human	QKAV KDEL	3
	T3BP[b]	cow	QKAV KDEL	4
	GSBP[c]	chick	QKAM KDEL	5
BiP	GRP78	rat	DTSE KDEL	6
		mouse	DTSE KDEL	7
		hamster	DTSE KDEL	8
		human	DTAE KDEL	9
		chick	EAAE KDEL	10
		Drosophila	DADL KDEL	11
		C. elegans[d]	DLDD KDEL	12
		C. elegans[d]	PSED HDEL	13
		tomato	EDDS HDEL	14
		S. cerevisiae	DYFE HDEL	15
		P. falciparum	EDVD SDEL	16
GRP94		hamster	STAE KDEL	17
	ERp99, endoplasmin	mouse	ESTE KDEL	18, 19
	hsp108	chick	STDV KDEL	20
Calligin	47 kd protein, gp46	rat	KM RDEL	21–23
ERp72		mouse	RSRT KEEL	24
Gut esterase		*C. elegans*	HSSN KDEL	25
60 kd esterase 1		rabbit	RETE HIEL	26
60 kd esterase 2		rabbit	HTEL	27
Auxin binding protein		*Zea mays*	FEAA KDEL	28
Reticulin	calregulin	rabbit	RRQA KDEL	29, 30
	CRP55	mouse	PAQA KDEL	31
Glucosidase II		mammals		32

[a] Prolyl 4-hydroxylase is a soluble $\alpha_2\beta_2$ tetramer; the β subunit is P.D.I.
[b] Thyroid hormone binding protein.
[c] Glycosylation site binding protein.
[d] *Caenorhabditis elegans* has two genes that encode BiP-like proteins.
[e] References: 1. Edman et al 1985. 2. Srinivasan et al 1988. 3. Pihlajaniemi et al 1987. 4. Yamauchi et al 1987. 5. Geetha-Habib et al 1988. 6. Munro & Pelham 1986. 7. Haas & Meo 1988. 8. Ting et al 1987. 9. Ting & Lee 1988. 10. Stoeckle et al 1988. 11. D. Rubin & K. Palter pers. commun.. 12. Heschl & Baillie 1989. 13. M. Heschl & D. Baillie pers. commun.. 14. A. Bennett pers. commun.. 15. Rose & Misra 1989. 16. Kumar et al 1988. 17. Sorger & Pelham 1987. 18. Mazzarella & Green 1987. 19. Smith & Koch 1987. 20. Kuloman et al 1986. 21. Kurkinen et al 1984. 22. Saga et al 1987. 23. D. Nandan. E. Zgunet pers. commun.. 24. M. Green pers. commun.. 25. J. McGhee pers. commun.. 26. Korza & Ozols 1988. 27. Ozols 1988. 28. D. Klambt pers. commun.. 29. Opas et al 1988. 30. Khanna et al 1987. 31. Macer & Koch 1988. M. Smith & G. Koch pers. commun.. 32. Strous et al 1987.