Endoplasmic Reticulum Stress Signaling in Disease

STEFAN J. MARCINIAK AND DAVID RON

Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom; and
Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, New York

I. INTRODUCTION

The success of multicellular organisms owes much to the efficiency gained from the cooperation between specialized cell types. This cooperation requires communication between distant components frequently achieved by the binding of molecules from one cell to receptors on another. Such information transfer requires proteins to be synthesized that can withstand the harsh extracellular environment, both as soluble ligands (hormones and transmitters) and cell surface molecules (receptors and adhesion molecules).

The extracellular compartment differs sufficiently from the cytosol that proteins destined for secretion or insertion into the plasma membrane require modifications inappropriate to the cytosol, such as glycosylation and disulfide bond formation. Generation of these modifications necessitates a compartment topologically distinct from the cytosol, which is provided by a membranous network called the endoplasmic reticulum (ER). With the evolution of this compartment, eukaryotes internalized a portion of the extracellular space in which they modify, fold, and assemble secreted and membrane proteins. The evolution of eukaryotes also created the challenge of regulating a protein maturation machinery outside the confines of the cytosol. Failure of this machinery to fold newly synthesized endoplasmic reticulum “client” proteins presents unique dangers to the cell and is termed “ER stress.” Early in evolution, a homeostatic mechanism developed to maintain the balance between the demand for ER function and ER synthetic capacity; furthermore, as organisms became more complex, especially with the appearance of long-lived professional secretory cells, the importance and complexity of this machinery increased greatly.

This review addresses the mechanisms that enable higher metazoans to produce extracellular proteins in the face of varying demand and the consequences should these mechanisms fail. In the first section, a brief overview of the ER protein maturation machinery is pre-
sented, along with some of the consequences of protein misfolding. Thereafter, the response to short-term perturbations in ER function is discussed, with emphasis given to human diseases caused by dysregulation of this response, followed finally by an examination of the importance of ER stress signaling over the longer term to secretory cell differentiation and lipid metabolism.

A. Synthesis, Folding, and Misfolding of ER Protein

Translation of membrane and extracellular proteins is performed by ribosomes on the cytosolic surface of the ER (3, 94). A signal recognition particle within the cytosol cotranslationally recognizes a signal sequence within the nascent polypeptide chain and directs it to a proteinaceous pore in the ER membrane, the Sec61 complex (92, 127, 149).

Proteins achieve a specific folded conformation first by acquisition of secondary structure, e.g., helices, due to the intrinsic properties of the amino acid sequence (39). This is followed by a further search for native structure, which involves diffusion of these preformed segments and assembly into larger modules (84). These processes entail the burial of amino acid side chains into a close-packed structure excluding water from the protein’s core (25). “Misfolding,” in this context, indicates that a protein persistently maintains the potential for nonnative interactions with the solvent promoting illegitimate interactions with other cellular components. Another consequence of aberrant folding is aggregation of proteins into insoluble higher order structures. These may be of three varieties, all of which may cause disease: disordered aggregates (e.g., rhodopsin in autosomal retinitis pigmentosa, Ref. 163), amyloid fibrils (e.g., amyloid β-peptide/tau in Alzheimer’s disease, Refs. 89, 148), and nonamyloid fibrils (e.g., α1-antitrypsin in α1-antitrypsin deficiency, Ref. 110).

Disordered aggregates are complex insoluble accumulations of protein resistant to normal degradation, while both amyloid and nonamyloid fibrils are simple noncovalent polymers, one-dimensional crystals formed from repeating subunits. Nonamyloid fibrils are characteristic of the serpinopathies, human diseases caused by mutations within the serpin family of proteins (111); these mutations allow a specific intermolecular interaction resulting in polymerization. In contrast, amyloid can be formed through polymerization of a number of otherwise unrelated proteins, each causing a distinct disease (79, 148, 207). Despite intensive study of these aggregates, surprisingly little is certain about the mechanism of their toxicity; however, it is becoming apparent that even in the case of amyloidoses, in which extracellular deposits are characteristic, some of the deleterious effect comes from toxic gain of function due to protein oligomers within the cell (227).

The inappropriate interactions of misfolded proteins threaten to disrupt normal cellular function, so the biosynthetic machinery has evolved sophisticated mechanisms to minimize their occurrence. Within the ER, chaperones such as BiP bind to incompletely folded proteins shielding them from other molecules (13, 64, 65, 126, 200). In addition, other catalysts directly promote correct folding by the addition of carbohydrates (63), cis-trans isomerization about peptide bonds (166), and the creation and rearrangement of disulfide bonds (189).

B. ER-Associated Degradation

Some proteins never attain their correct conformation, perhaps due to a mutation impeding correct folding or because the cell lacks the energy to drive sufficient cycles of chaperone interaction. After a lag period of 30–90 min, misfolded ER client proteins are disposed of by ER-associated degradation (ERAD) (107). The timer that dictates this delay, at least for misfolded glycoproteins, is likely to be ER mannosidase I, which operates by trimming mannose residues from N-linked glycans (38, 76). Evidence that this regulates degradation comes from its inhibition, which retards ERAD (115, 186), and from its overexpression, which enhances ERAD (72). Recently, EDEM, a stress-inducible catalytically inactive mannosi-dase homolog, has been shown to interact with misfolded glycoproteins and effect their extraction from the folding cycle (43, 73, 122).

The path by which unfolded proteins are retrotranslocated into the cytosol is not clear. Some evidence supports a role for the original Sec61 translocon complex (150, 228), although a new complex containing derlin-1 and p97, a cytosolic ATPase, has been shown to be involved in retrotranslocation of major histocompatibility complex (MHC) class I molecules (106, 219). Once within the cytosol, ERAD substrates are degraded by the ubiquitin/proteosome pathway (49, 51, 151).

For these synthetic and quality assurance processes to proceed efficiently, there needs to be coordination between the input load of unfolded client proteins and the maturation machinery of the ER. If the client protein load is excessive compared with the reserve of ER chaperones, the cell is said to be experiencing “ER stress.” If unchecked, ER stress threatens to overwhelm the processing capacity of the ER, leading to the accumulation of unfolded proteins and collapse of the secretory pathway. Consequently, signaling pathways have evolved that respond to ER stress by regulating processes on both sides of the ER membrane through an adaptive mechanism termed the unfolded protein response (UPR).
C. The UPR

The nature of ER stress encountered by a cell dictates the nature of its UPR. During normal function a secretory cell will experience dramatic variations in the flux of new proteins through its ER in response to changes in demand. During a transition from low protein synthesis to high, there is a need to increase ER protein-folding capacity to avoid overloading of chaperones. Consequently, the UPR regulates transcription factors whose targets include genes for components of the ER protein maturation machinery. However, inherent in this response is a temporal delay, since new chaperones cannot be made instantly. To avoid the accumulation of misfolded client protein during this hiatus, the rate of secretory protein synthesis must also be under UPR-directed modulation through a more rapid transcription-independent mechanism. Furthermore, when protein-folding efficiency falls, for instance, in ischemic tissue when lack of energy impedes the proper function of ER chaperones and enzymes, the primary adaptation may be through reduced client synthesis rather than by increasing the levels of this machinery, which itself would be costly in energy. Therefore, both transcriptional and translational signals emanate from the stressed ER to allow a coordinated response.

Not all changes in ER protein flux are transitory nor can all perturbations be survived. In some cases the change in ER protein synthesis accompanies a change in cellular phenotype, such as during differentiation into a secretory cell. Here the prolonged activation of UPR signaling itself appears to play a role in the differentiation process, leading to dramatic alterations in ER structure. When the level of ER stress is too great to allow adaptation, the cell may die. It remains unclear to what extent the UPR plays a role directly in cell death, but studies of ER stress lethality have revealed novel potential therapeutic targets for intervention in human disease.

These concepts will be elaborated upon in more detail below with emphasis on the human diseases to which they are most relevant.

II. ENDOPLASMIC RETICULUM HOMEOSTASIS

A. Short-Term Perturbations

Tissues whose primary function is the secretion of protein should depend most strongly on the UPR, especially those liable to make large changes in ER client protein load. A good example is the pancreatic β-cell, which produces insulin in response to changes in circulating glucose. It constantly executes changes in ER synthetic capacity to track current glucose levels. Although the precise mechanisms remain unclear, it has been established that upswings in glucose, at least in the short term, promote proinsulin synthesis through enhanced protein translation (75, 203). This stimulation of proinsulin translation by glucose is specific to the β-cell and dependent on cis-acting elements in untranslated regions of proinsulin mRNA (206). The resultant large fluctuations in ER client protein load are compensated for by the UPR, but render the β-cell highly vulnerable to defects in ER stress signaling in animal models (54) and in human disease (34).

1. Regulating client protein load

Clues to the homeostatic importance of regulating new protein synthesis in response to changes in ER load have come from a rare human disease. Wolcott-Rallison syndrome is an autosomal recessive condition characterized by early development of diabetes mellitus with associated bone, liver, renal, and neuronal defects (209). It has been mapped to 2p12 where candidate gene analysis yielded mutations in EIF2AK3 (34), better known as PKR-like eukaryotic initiation factor 2 (eIF2α) kinase (PERK) (56) or pancreatic eIF2α kinase (PEK) (171).

PERK is an ER-resident transmembrane protein ubiquitously expressed, but highly enriched in professional secretory cells (56, 170, 171, 175). Its cytosolic portion is highly homologous to a yeast stress-responsive kinase, Gcn2p (56, 170). Unlike PERK, Gcn2p is a soluble protein that enables yeast to adapt their rate of new protein synthesis to the levels of available amino acids (125, 190, 191, 208). When amino acids are limiting, Gcn2p phosphorylates the α-subunit of eIF2α.

The eIF2 complex is essential in all eukaryotes for new protein synthesis, since it recruits the initiator methionyl tRNA to ribosomes about to begin translation (69). Phosphorylation of eIF2α inhibits this activity and thus globally reduces protein translation. In metazoans, the GCN2 gene has expanded into a family of related eIF2α kinases, all of which inhibit protein translation in response to stress. These kinases all possess homologous catalytic domains but have different stress-sensitive regulatory domains. GCN2 persists to respond to amino acid starvation, while new members sense disparate stresses including viral infection (PKR and PKZ) and iron deficiency (HRI) (9, 10, 23, 44, 159). The existence of PERK, an ER transmembrane eIF2α kinase, led to the appreciation that cells are able to adjust their level of new protein synthesis to match their ER folding capacity and prevent overload of the secretory pathway (55, 56).

To respond to ER stress, PERK must transduce a luminal signal across the ER membrane to its cytosolic kinase domain. The nature of this stimulus is intriguing. The ER is a site of secretory protein maturation and so will, by necessity, harbor incompletely folded and actively folding protein intermediates. The level of these is a reg-
ister of client protein throughput, rather than an indicator of the ER’s capacity to fold them. A homeostatic mechanism that responded solely to unfolded protein could not, therefore, measure the efficiency of ER protein biosynthesis, but, instead, only respond to the absolute client load. Evolution appears to have found a more subtle solution: PERK responds not to the level of misfolded protein within the ER, but instead to changes in the ER chaperone reserve, that is, to the unfolded/misfolded protein-to-chaperone ratio. BiP is an abundant chaperone within the ER that binds to folding proteins through interaction with exposed hydrophobic residues. BiP overexpression has long been known to suppress the UPR (41, 197). Two potentially complementary models have been proposed for the mechanism(s) involved. During unstressed conditions, BiP also binds to the luminal domain of PERK (12, 108). This interaction correlates with the inactive, monomeric state of PERK. When an increase in ER client load is experienced, BiP dissociates from PERK, perhaps through sequestration to unfolded clients or by more direct mechanisms. The dissociation of the PERK-BiP complex is hypothesized to allow PERK to cluster in the plane of the membrane, leading to activation of the cytosolic kinase domain through a process of trans-auto-phosphorylation and a dramatic increase in affinity towards eIF2α (12, 116).

Recently, crystal structure data on IRE1, whose luminal domain is homologous to that of PERK, raise the possibility that direct binding of sensing molecules to unfolded proteins might also take place (32). Dimerization of the luminal portion of IRE1 appears to generate a groove similar to that found in the peptide-binding pocket of MHC molecules. If the groove formed by dimerization of IRE1 can be shown to bind unfolded proteins, this would suggest an alternative model whereby unfolded domains that are unbuffered by chaperones signal directly in the unfolded protein response, perhaps combining direct elements (unfolded protein) and indirect elements (chaperone reserve).

Humans with this Wolcott-Rallison syndrome share some clinical features with PERK−/− mice, which are born with essentially normal islets that are capable of normal insulin synthesis and secretion (54). Indeed, when isolated islets from prediabetic PERK−/− mice are challenged with glucose, their secretion of insulin is greater than wild-type controls. During the first few weeks of life these pups develop overt diabetes due to progressive β-cell destruction. These findings indicate that PERK is essential for β-cells to cope with normal physiological ER stress arising from day-to-day insulin synthesis. This involves restraining the insulin synthetic response to glucose by attenuating new protein translation. In other words, protein synthesis is limited in normal animals by eIF2α phosphorylation to a level with which the existing ER machinery can cope; in the PERK−/− animals and Wolcott-Rallison patients, new proteins continue to enter the ER regardless of its ability to fold them. This manifests as the accumulation of misfolded products (109) and excessive stress signaling and culminates in cell death (54). Interestingly, individual mutations of the other known eIF2α kinases (PKR, HRI, and GCN2) do not result in mice with diabetes, indicating that only eIF2α phosphorylation in response to ER stress is involved in this phenotype (52, 217, 226).

The importance of translational regulation in response to ER stress may not be restricted to rare genetic diseases, as a recent study, using transgenic mice subtly impaired in their ability to phosphorylate eIF2α has revealed a phenotype remarkably similar to that of human type II diabetes. PERK and all eIF2α kinases phosphorylate a single residue on their substrate: serine-51. Mutation of this serine to an alanine (eIF2αS51A) prevents its phosphorylation in response to any stress. Heterozygous eIF2αS51A mice are capable of significant translational regulation by phosphorylation of the remaining population of eIF2α molecules and are born with normal pancreatic islets and under normal conditions do not develop diabetes (164). In contrast, eIF2αS51A homozygotes are born with severe β-cell deficiency by late embryonic stage (165). However, when heterozygous mice are fed a high-fat diet, a new and interesting phenotype is revealed (164). These mice are more prone to obesity because of a failure to increase energy expenditure in response to the excess calories. Remarkably, they more rapidly exhibit hyperleptinemia, mild hyperinsulinemia, and raised fasting glucose, features of the “metabolic syndrome” currently dominating the epidemic of type 2 diabetes in developed countries. When islets from high-fat diet-fed eIF2αS51A mice were analyzed in vitro, they demonstrated increased basal insulin secretion but reduced stimulated insulin secretion. Prolonged high-fat feeding led to distension of the ER by unfolded proteins, including proinsulin, and to profound glucose intolerance. This surprising finding highlights the exquisite sensitivity of pancreatic β-cells even to subtle dysregulation of eIF2α phosphorylation and raises the intriguing possibility that minor variations in efficiency of eIF2α phosphorylation could, in principle, contribute to significant morbidity. As yet, no studies have addressed the possible existence of such variations in eIF2α signaling between patient groups, for example, obese subjects with or without associated features of metabolic syndrome.

2. Integrated stress response

The effects of eIF2α phosphorylation are not restricted to attenuation of protein translation, but also include activation of a transcriptional program. This is illustrated in a family of white matter hypomyelination disorders, termed CACH/VWM leukodystrophies. These
are autosomal recessive conditions with a spectrum of severity from congenital to adult-onset disease, involving the progressive loss of mental and motor faculties due to deterioration of brain white matter. In all cases the causative mutations have been found within subunits of an eIF2α-interacting protein, eIF2B (46, 47, 100, 158). As part of its normal active cycle, eIF2 undergoes rounds of GTP binding, hydrolysis, and guanine nucleotide exchange. Only in the GTP-bound state is eIF2 capable of bringing methionyl tRNA to the ribosome. The guanine nucleotide exchange factor (GEF) for eIF2 is, in fact, eIF2B; however, once phosphorylated, eIF2 binds avidly to its GEF inhibiting further exchange (27). In this way, phosphorylation of a small pool of eIF2 can inhibit most GEF activity (Fig. 1). The mutations that cause the CACH/VWM disorders appear to interfere with eIF2B GEF activity, but recent studies with cells from patients with CACH/VWM have shown that, rather than affecting global translation rates during stress, these mutations increase signaling via a transcription factor known as ATF4 (82, 193).

Whilst the majority of RNA transcripts experience decreased translation during periods of increased eIF2α phosphorylation, a small and still poorly defined subset is translated more efficiently (53, 68, 121, 124). This paradoxical effect is due to the presence of multiple upstream open reading frames (uORFs) 5′ to the coding sequence initiation codon (1, 68, 112, 194). The best-characterized example in mammals is the stress-inducible transcription factor ATF4 (112, 194). This transcript has two uORFs, the second overlapping out of frame with the true ATF coding sequence. During resting unstressed conditions, ribosomes scan along the mRNA translating uORF1 and then recapacitate by re-binding eIF2/GTP/Met-tRNA ternary complex in time to translate inhibitory uORF2. During stress, limiting levels of eIF2 ternary complex lead to a delay in recapacitation of these scanning ribosomes, such that they fail to reinitiate at uORF2 but instead scan to the ATF4 initiation codon. By then, a proportion have reacquired ternary complex allowing translation of the active transcription factor.

In yeast, the single eIF2α kinase Gcn2p regulates a transcription program governed by the response to amino acid deprivation through induction of the transcription factor GCN4, by a mechanism analogous to that of ATF4 (37, 68, 125). As the family of eIF2α kinases has diversified to respond to numerous stresses, so too has the transcriptional program it regulates in metazoans. It continues to regulate genes essential for amino acid sufficiency, but now ATF4 also induces antioxidant genes and genes of the ER protein maturation machinery (57). Because eIF2α phosphorylation triggers a final common pathway in response to many stresses, this transcriptional program in combination with protein translation modulation has been termed the integrated stress response (ISR) (57) (Fig. 1).

Oxidative protein folding is inextricably linked to the generation of reactive oxygen species (ROS), in large part through the activity of ER oxidase 1 (ERO1), which generates much of the oxidizing potential of the organelle (48, 152, 187, 188). Without ER stress-regulated activation of ATF4, PERK−/− β-cells, and those from Wolcott-Rallison patients, lack the prosurvival effects of the ISR with its antioxidant expression program (54, 57). It is worth remarking, therefore, that recent data indicate the oxidation of cysteines to form disulfide bonds leads directly to the generation of ROS and cell death (58) and that ER stress-dependent ERO1α induction promotes ER oxidation (117). Insulin, with its three disulfide bonds per molecule, might therefore be expected to impose a considerable ROS load on the cell. The finding that targeted deletion of a widely expressed ER-associated reductase,
**Neb5or**, leads to a selective loss of β-cells, once more underlines their sensitivity to redox stress (212).

3. **Recovery of protein translation**

**A**) eIF2α PHOSPHATASES. The execution of an adaptive gene expression program like the ISR can only be performed if the attenuation of protein translation accompanying eIF2α phosphorylation is reversed (29, 114, 131). Indeed, persisting loss of protein synthetic activity would compromise cell survival, regardless of a need for adaptation to stress. For these reasons, phosphatase activity exists to relieve eIF2α phosphorylation, restoring normal translation initiation activity.

Viral infection leads to eIF2α phosphorylation through activation of PKR (10, 90). This is part of an innate antiviral response antagonizing viral protein translation. Frequently, the relationships between infective agents and host are characterized by the adaptations of one party being countered by adaptations within the other. For example, viruses have evolved proteins that directly inhibit the activity of PKR (21, 153); however, in the case of herpes simplex virus (HSV), viral protein translation can be maintained despite continued PKR activity. This is achieved through enhanced dephosphorylation of eIF2α by a virally encoded protein, ICP34.5, which binds to a host serine/threonine phosphatase, protein phosphatase 1 (PP1), and directs its activity towards eIF2α (61, 62).

It is now well appreciated that protein phosphatases, including PP1, are represented within the genome by relatively few genes and achieve their specificity through the formation of complexes with regulatory subunits (28). It was therefore gratifying when suppressor screens of ER stress revealed two mammalian genes, CReP and GADD34, with striking homology to ICP34.5 (80, 131). These proved to be responsible for reversing PERK-induced translational attenuation, as they too are regulatory subunits of PP1 that specifically promote eIF2α dephosphorylation (29, 80, 131). CReP is constitutively expressed, while GADD34 is strongly induced during ER stress. The induction of GADD34 in particular appears central to the reversal of stress-induced translational attenuation (18, 87, 114, 131, 132).

**B**) **ROLE OF CHOP.** Activation of ER stress signaling has been correlated with the induction of cell death in many models of ER stress. The mechanisms by which this might promote cell death remain unclear; however, CHOP, a transcription factor downstream of the PERK-ATF4 axis, was thought to be important to the process, since its deletion ameliorates tissue damage during ER stress (20, 136, 137, 174, 229). The downstream mediators of this effect remained unknown until recent work demonstrated that some of the toxic effects of CHOP are the result of GADD34 induction (117).

The CHOP dependence of ER stress toxicity is well illustrated in a murine model of diabetes, the Akita mouse. Mice, unlike humans, have two genes encoding insulin (Ins1 and Ins2). These appear functionally redundant, as deleting both alleles of either gene has no effect on glucose homeostasis (101). It is therefore remarkable that a spontaneous mutation in the Ins2 gene (Akita) leads to a severe diabetic phenotype inherited as a semi-dominant trait (196, 223).

Insulin is translated as a single proinsulin polypeptide that undergoes oxidation within the ER to form three intramolecular disulfide bonds. After excision of a short peptide from the prohormone, these three bonds hold the remaining two insulin chains together. The Akita mutation converts a conserved cysteine to tyrosine preventing formation of one of these bonds (196). This mutant insulin is not secreted, but instead degraded (8, 196). Mice harboring the Akita mutation have apparently normal pancreatic islets at birth, but go on to develop diabetes in the ensuing weeks due to selective β-cell loss in the absence of inflammation. This was initially attributed to nonspecific effects (196); however, when this mutation was introduced into mice with a targeted deletion of the CHOP gene, onset of diabetes was significantly delayed due to the preservation of β-cell numbers (136).

As indicated above, CHOP is a transcription factor highly induced during ER stress (197, 229). The protective effect of CHOP deletion appears not restricted to the Akita mouse, since it also affords protection against cytokine-induced β-cell death, in which nitric oxide triggers ER stress by the depletion of ER calcium stores (20, 137), nor is it restricted to the pancreas, as CHOP−/− animals are resistant to renal damage caused by the ER stress-inducing toxin tunicamycin (229). CHOP deletion also protects mice from dopaminergic neuron loss following 6-hydroxydopamine injection (174). This drug, which is commonly used as a model for Parkinson’s disease since it selectively kills dopaminergic neurons in vivo, causes ER stress in dopaminergic neurons in tissue culture (71, 162).

One popular hypothesis held that CHOP, a metazoan specific gene, evolved to promote the death of individual cells in response to insurmountable levels of ER stress, affording particular benefit to multicellular organisms (135). In accordance with this, CHOP overexpression has been shown to sensitize cells to the toxicity of ER stress (118). Analysis, however, of the CHOP transcriptional program failed to reveal an obvious connection to pathways that promote cell death. Instead, CHOP appears to defend ongoing protein secretion, in part through induction of GADD34 (117). In this light, it appears that the ER-stressed cell is challenged with balancing the need to defend its chaperone reserve, by limiting secretory protein synthesis, against the need of both cell and organism for ongoing ER synthetic function. According to this
model, the protection from death afforded by deletion of the CHOP gene in some models of ER stress reflects, therefore, an enforced shift of this balance from synthetic function towards the defense of chaperone reserve (Fig. 2). In accordance with this, GADD34 null animals have been shown to be protected from ER stress-induced nephrotoxicity equally well as CHOP−/− animals (117).

Of course, there exists the formal possibility that CHOP-dependent GADD34 induction evolved to promote cell death rather than ongoing protein synthesis; however, two pieces of evidence argue against this. First of all, in tissue culture GADD34 mutant cells are more prone to cell death in response to ER stress induced by thapsigargin (132). This agent causes ER stress through depletion of ER calcium stores and is particularly effective at shutting down protein synthesis, indicating that recovery of protein translation is necessary for the adaptive prosurvival effects of UPR signaling. Second, in murine models of Pelizaeus-Merzbacher leukodystrophy (due to a proteolipid protein mutation), CHOP expression appears to have an antia apoptotic effect, a strange phenotype for a prodeath signal (176).

It remains unclear whether the toxic effects of CHOP in models of ER stress are primarily mediated through GADD34. Although this may be true for tunicamycin-induced nephrotoxicity, it has yet to be proven for other models. For example, it is possible that other CHOP target genes, such as ERO1α, may mediate toxicity in a tissue-specific fashion. In cell culture, CHOP overexpression leads to ROS production (118), and CHOP-dependent ERO1α induction appears responsible for increased oxidation in the stressed ER (117). This may be relevant to tissues known to be sensitive to redox perturbation, such as the pancreatic β-cell.

Perhaps there are circumstances experienced by multicellular organisms during which the cost of worsening ER stress is compensated for by enhanced protein secretion, at least in a subset of cells. Because osteoblasts are replenishable secretory cells they may behave in this fashion and, consistent with this, recent data suggest that the bones of CHOP−/− animals demonstrate defects due to impaired osteoblast function (147). If the importance of CHOP in the defense of protein secretion is a general feature of replenishable secretory cells, it is tempting to speculate that terminally differentiated effector cells of the immune system might also behave thus. Challenging CHOP−/− and GADD34 mutant animals with infectious agents could test this hypothesis.

An exciting prediction arising from these findings is that careful titration of GADD34 inhibitors might be of therapeutic benefit in diseases involving ER stress, through the preservation of ER chaperone reserve. It is encouraging, therefore, that a recent screen for small molecules that protect cells from ER stress yielded a compound that enhances eIF2α phosphorylation and, in HSV-infected cells, appears to block eIF2α dephosphorylation by ICP34.5 (15). If compounds can be synthesized that have specificity for GADD34 over CREP, there exists the potential for selective effects on ER-stressed tissue, perhaps limiting systemic toxicity.

4. NF-κB and ER activity

A distinct antiviral response initiated by the accumulation of viral proteins in the ER has previously been postulated (140, 141, 143). This followed from observations that virally encoded proteins could activate NF-κB signaling in a cell autonomous fashion (119, 144, 145). This is in contrast to most canonical NF-κB stimuli, e.g., tumor necrosis factor (TNF)-α, which originate from outside the cell. This “ER overload response” (EOR), as it was termed, appeared not to be specific for viral proteins, since other membrane proteins, e.g., MHC class I, when overexpressed could also induce NF-κB signaling (145). It was subsequently extended to include NF-κB induction from ER accumulation of other nontransmembrane proteins, including α1-antitrypsin polymers (67, 95). The transduction mechanism for this pathway remained obscure, although release of ER calcium and ROS were both suggested (140, 143).

Recent discoveries have implicated PERK signaling in the activation of NF-κB in response to ER perturbations, challenging the notion that the EOR is distinct from
the UPR (36, 78). There exists a protein family of NF-κB inhibitors termed I-κB, which, in unstimulated cells, is expressed constitutively and sequesters NF-κB inactive within the cytoplasm (160). Classically, NF-κB is activated through the relief of this inhibition, for example, TNF receptor activation triggers I-κB phosphorylation, causing its ubiquitination and degradation by the proteosome (146). In the case of ER stress, at least when signaled by PERK, disinhibition of NF-κB is achieved once again by reducing IκB levels, but through a novel mechanism. I-κB has a far shorter half-life than NF-κB. Consequently, translocational attenuation preferentially lowers IκB levels, releasing NF-κB to execute its transcription program. A similar effect can be obtained by treatment of cells in culture with other agents that inhibit protein synthesis, e.g., cycloheximide (36, 78). This novel mechanism extends to NF-κB activation in response to other stresses, including ultraviolet (UV) irradiation and amino acid deprivation, where GCN2 appears to be the relevant kinase (77, 78, 211). Some controversy persists within this field, since the existence of a PERK-dependent NF-κB cascade does not preclude activation of NF-κB by other ER-originating signals. This will be clarified once EOR stressors have been more extensively studied in PERK mutant and eIF2αS51A cell lines.

The physiological significance of ER stress-induced NF-κB signaling likely reflects the importance of NF-κB signaling in other settings, namely, the coordination of inflammatory response and promotion of cellular survival. In bacterial infection, the induction of NF-κB by microbial components sensed by Toll-like receptors leads to upregulation of early effectors of the innate immune response, including chemokines, proinflammatory cytokines, and immune receptors (7, 91, 215). These early effectors then generate a further cycle of NF-κB induction and generation of an antimicrobial response (104). The induction of such a cascade by detection of viral components within the ER, and subsequent PERK-dependent NF-κB activation, might represent an analogous antiviral response.

Developmentally, NF-κB is a prosurvival signal in some cell types, for example, in the maturation of B and T lymphocytes and in the development and regeneration of the liver (6, 103, 183, 210). While many stimuli that induce NF-κB also promote cell death, paradoxically NF-κB tends to oppose signal-induced cell death by induction of antiapoptotic genes including c-IAP1, c-IAP2, and FLIP (120, 195). Indeed, NF-κB can have proliferative effects through its targets c-myc and cyclinD1 (142, 155). It is therefore possible that NF-κB induction may also link secretory activity to trophic signals.

**B. Long-Term Change**

In contrast to the continuous modulation of ER synthetic activity, differentiation of a cell into a highly secretory phenotype requires a dramatic expansion of ER. This is exemplified in the development of B-lymphocytes into plasma cells. In this regard, defects in plasma cell differentiation have helped reveal an additional role for the UPR in the regulation of absolute ER mass within the cell.

1. **Linking ER mass to demand**

   Multiple myeloma is a hematological neoplasm characterized by presence in the blood of a monoclonal immunoglobulin or Bence Jones protein (free monoclonal κ- or λ-light chains). These are produced by myeloma cells within the bone marrow, which over time expand to displace normal marrow leading to anemia and immunological deficiency. The plasma concentration of the monoclonal paraprotein can reach exceedingly high levels (>100 g/l), even leading to complications through increased blood viscosity. The causative myeloma cells are monoclonal expansions of plasma cells, immune cells normally charged with large-scale immunoglobulin manufacture.

   To understand the genesis of plasma cells and myelomas, the transcription factors necessary for their differentiation have been determined. Early work demonstrated that myeloma cell lines frequently express a basic leucine zipper transcription factor, XBP-1, to extremely high levels (26, 204). XBP-1 was also shown to be essential for normal plasma cell differentiation (156, 157, 204). Myeloma cells and mature plasma cells share the ability to secrete huge quantities of immunoglobulin, thanks to their highly developed secretory pathway. Histologically these cells have characteristically basophilic cytoplasm due to the large numbers of ribosomes associated with their extensive ER (Fig. 3). It seems likely that XBP-1’s role in plasma cell differentiation is related to regulation of ER expansion in response to increased demand (96). Furthermore, new therapeutic agents being developed to treat myeloma may exploit the dependence of this neoplasm on UPR activation (2, 66). Proteosome inhibitors have proven to be highly effective in the killing of myeloma cells and in early trials have achieved remarkable clinical remissions in otherwise drug-resistant disease (81). Their selectivity toward myelomas might reflect the reliance of these cells on efficient ERAD, which itself is dependent on XBP-1 signaling (96, 221). The first drug of this class, bortezomib (Velcade), has now been approved by the FDA for use in otherwise refractory disease (17).

   XBP-1 regulates many genes of the unfolded protein response in metazoans, including genes involved in generation of the ER membrane (177). It is upregulated at the protein level during ER stress by a remarkable mechanism. The XBP-1 primary transcript is translated, albeit inefficiently, to produce a protein without significant transactivation activity (98). ER stress-dependent splicing of XBP-1 mRNA involves excision of a short intron (26nt in mammals) causing a frame shift in its open reading
IRE1β appears limited to the gut epithelium (198). The significance of this tissue-dependent isoform expression is unclear but may point toward cell type specific differences in UPR signaling. IRE1α−/− animals are not viable, whereas IRE1β mutants appear well, although are more prone to chemical-induced colitis (11, 185, 192).

The IRE1-dependent arm of the UPR is the most ancient in evolutionary terms, since the entire UPR of lower eukaryotes, including that of yeast, is regulated by this molecule (30, 31, 40, 184, 185). Although in yeast the target transcription factor HAC1 bares little homology to XBPI, the activation of Ire1p is identical. It was, in fact, historically the first UPR pathway to be described in molecular terms (14, 22, 50, 85, 123, 167, 173, 202). In yeast, after Ire1p-mediated intron excision from the HAC1 mRNA, religation by the tRNA ligase Rgl1p generates an efficiently translated message (123, 161, 172). The analogous ligase in higher organisms has not yet been identified and remains an important unanswered question.

2. Insulin resistance and ER stress

In addition to linking secretory capacity to demand through regulation of ER expansion, IRE1 also plays an important role in integrating ER stress signaling with other stress and growth factor sensitive pathways.

An important feature of type 2 diabetes is peripheral insulin resistance. Normally, activated insulin receptors phosphorylate proximal signaling molecules, such as insulin receptor substrate 1 (IRS-1), on tyrosine residues, which transduce the effects of insulin through interaction with cytosolic targets (180, 182). In obesity (genetic or dietary), this tyrosine phosphorylation is inhibited by JNK-dependent serine phosphorylation of IRS-1 (4, 5). Surprisingly, the mechanism of obesity-related JNK activation appears to involve ER stress.

For unclear reasons, liver and adipose cells from obese mice show biochemical evidence of ER stress with PERK activation, eIF2α phosphorylation, and BiP induction (139). It has previously been shown that ER stress, through IRE1 activation, can directly trigger the JNK cascade (192). Activated IRE1 recruits the scaffolding protein TRAF2 to the ER membrane (192), which triggers a mitogen-activated protein (MAP) kinase cascade leading to JNK activation (129, 130). Thus IRE1 offers a plausible mechanistic link between obesity and peripheral insulin resistance. In this model, obesity-associated ER stress would contribute to insulin resistance by causing JNK activation through IRE1/TRAF2 with subsequent IRS-1 serine phosphorylation (Fig. 4). Consistent with this hypothesis, IRE1−/− cells fail to active JNK or phosphorylate IRS-1 during ER stress unlike wild-type controls. Supporting a causal link between peripheral ER stress and insulin resistance, a recent study has demonstrated protection against obesity-induced type 2 diabetes in mice...
by overexpression of an ER chaperone, while knockdown of the same chaperone was diabetogenic (128). Furthermore, in the Akita mouse, systemic rather than \( \beta \)-cell overexpression of an ER chaperone appeared to improve peripheral insulin sensitivity (138).

In light of this, it is possible that the peripheral insulin resistance seen in the \( \text{eIF2}^\alpha_{\text{S51A}} \) heterozygous mice (above) might have resulted from increased signaling through the IRE1-JNK arm of the UPR attempting to compensate for impaired PERK-dependent signaling. Consistent with this speculation, \( \text{PERK}^{-/-} \) cells, which are impaired in \( \text{eIF2}^\alpha \) phosphorylation, experience basically higher IRE1 activation (55).

It can be seen that the role of UPR signaling in diabetes is far from straightforward. At the level of the pancreatic \( \beta \)-cell, the UPR protects against developing type I diabetes by ensuring that this highly secretory tissue maintains efficient function despite wide swings in ER protein flux. Conversely, chronic ER stress signaling (perhaps related to obesity) appears involved in the etiology of peripheral insulin resistance. We can at present only speculate as to how the UPR progresses from regulation to dysregulation over the passing years. It is known that chronic exposure to elevated saturated free fatty acids induces ER stress (83). Consequently, obesity-related ER stress may reflect the evolutionarily recent phenomenon of chronic nutrient excess. The dysregulation of ER stress signaling seen in this disease may reflect the lack of selection pressure to adapt this homeostatic mechanism to chronic inappropriate activation.

3. ER stress and lipid metabolism

A) Homocysteine and atherosclerosis. Among the known risk factors for atherosclerosis, ER stress would currently be ranked low, if at all. This may begin to change as evidence accumulates implicating the UPR in disordered lipid metabolism. The relationship between the UPR and lipid accumulation is a complex one. On the one hand, many target genes of the UPR involve lipid synthesis, in part to allow expansion of the ER itself; on the other hand, there is increasing evidence that perturbation of the lipid environment within the ER can activate UPR signaling.

Homocysteinemia is associated with the development of atherosclerosis, certainly in patients with rare inherited disorders of amino acid metabolism, but also in the wider population, as indicated by large epidemiological studies (42, 70, 201). Although homocysteine is a sulfur-containing amino acid, some of its toxic effects appear to involve dysregulation of cholesterol and triglyceride biosynthesis (205). Patients with inherited hyperhomocysteinemia or laboratory animals fed a homocysteine-rich diet develop hepatic steatosis in the absence of marked rises in plasma lipids. The likely explanation is local lipid synthesis, and consistent with this, homocysteine has been shown to activate lipogenic signaling via the sterol regulated element-binding proteins (SREBPs) (205). Surprisingly, ER stress contributes to the activation of SREBP by homocysteine. Livers of homocysteine-fed mice contain raised levels of ER chaperones, as do cultured cells exposed to high levels of homocysteine in vitro. The importance of ER stress in this lipogenic signaling was demonstrated by BiP overexpression, which ameliorated SREBP induction in response to homocysteine.

The relationship between the UPR and SREBP, while clearly evident, is still not entirely explained in mechanis-
tic terms. The SREBPVs are transmembrane proteins that, under sterol-sufficient conditions, are retained within the ER through interaction with other ER membrane proteins, Insig-1 and Insig-2 (213, 216). When cholesterol levels are low, this interaction is weakened, whereupon SREBP progresses to the Golgi to be cleaved sequentially by two serine proteases, S1P and S2P (33). This cleavage liberates the cytosolic portion of SREBP as a soluble protein, which enters the nucleus to bind specific DNA elements, thus transactivating target genes involved in lipid synthesis. One possible mechanism of ER stress-induced SREBP activation, for which there is evidence, involves stress-dependent loss of Insig-1 from the ER (98).

B) SREBP AND ATF6. Interestingly, the third known class of ER stress sensor, exemplified by ATF6, is unrelated to PERK or IRE1. ATF6 is another ER transmembrane protein, that undergoes processing similar to SREBP (24, 60, 102). It is normally retained within the ER through interaction, not with Insig proteins, but with BiP (168). During ER stress, ATF6 is released from the ER and moves on to the Golgi. There it is processed by the very same proteases that act on SREBP, to liberate a soluble translocation factor (218). Unlike SREBP, the targets of ATF6 are classical UPR genes (133, 199, 220). Indeed, ATF6 was isolated in a search for ligands of an ER stress-responsive DNA element (ESRE). Despite their separate transcriptional programs, there may still be some cross-talk between ATF6 and SREBP. A recent report suggests that within the nucleus, cleaved ATF6 can bind SREBP2 directly to inhibit its transactivation potential and oppose lipogenesis (224).

Following from the identification of ATF6, there have been several related transcription factors identified, which share similar domain structures and localization to the ER (Luman (35, 113, 154), OASIS (88), CREBH (225), ATF6β (59), CREB4 (178)). The existence of such a variety of UPR signaling molecules at first seems puzzling, but a clue to their roles may come from tissue distribution. In the cases of OASIS and CREBH, at least, there is clear evidence for tissue specificity, OASIS being found in astrocytes of the central nervous system, while CREBH is restricted to the liver. This raises the possibility that ER stress may elicit tissue-specific transcriptional responses. In the case of CREBH for example, it has already been shown that its activation by ER stress leads to induction of genes of the acute phase response, providing a direct link between hepatic ER stress and systemic inflammation (225).

C) CHOLESTEROL’S IMPACT ON THE ER. The dependence of cholesterol metabolism on ER stress signaling may reflect the impact of sterols on ER function. This is well illustrated by the macrophage in development of atherosclerosis. During the initial phase of atherosclerotic lesion formation, foam cells have a characteristic histological finding within the vessel intima. These are macrophages that have taken up oxidized lipoprotein particles and become laden with cholesterol. This cholesterol is stored as esters within large lipid vesicles, which give foam cells their foamy appearance. Over time, death of these cells, likely due to the toxic effects of unesterified cholesterol, results in deposition of extracellular cholesterol within the plaque. Unlike its esters, free cholesterol efficiently inserts into lipid bilayers and can alter membrane physical properties. The ER membrane, being poor in free cholesterol (16), may be especially sensitive to cholesterol loading. Indeed, recent findings indicate that for free cholesterol to have its toxic effect it must be trafficked to the ER (45). This trafficking results in activation of UPR signaling and caspase activation, and eventually in macrophage apoptosis. In addition, ER stress caused by free cholesterol loading of macrophages promotes chemokine secretion, and this may contribute to the formation of “vulnerable” atherosclerotic lesion, which are prone to rupture as a function of their high inflammatory cell infiltrate (105).

III. CONCLUDING REMARKS

The complexity of the mammalian UPR, mediated as it is by three distinct classes of ER sensors and numerous transcription factors, may provide a degree of redundancy; however, perhaps a more interesting interpretation of this complexity may be to allow higher organisms a more subtle response to ER stress. For example, the division of UPR genes between multiple transcription factors likely enables portions of the gene expression program to be induced as is appropriate either to the duration or intensity of ER stress. While ATF6 and XBP-1 transactivate many of the same genes, their expression programs show important differences (97). Binding of ATF6 to the BiP promoter is sufficient for maximal transactivation, while XBP-1 binding is sufficient for maximal induction of EDEM (97, 221). In contrast, many genes require both arms to be active in order for best induction (214). It has been suggested that an early, ATF6-dominated, response may attempt to compensate for ER stress through chaperone induction alone. While XBP-1 activation, which is more sustained, may, in addition, promote protein degradation through the induction of ERAD components. Accordingly, IRE1α−/− cells have been shown defective in glycoprotein ERAD (221).

Repeatedly, models of ER dysfunction have generated models of diabetes, and investigations into rare genetic diabetic syndromes have revealed unexpected links to ER stress. The field remains fertile for further study, since mutations in ER stress-inducible genes such as WFS1 (Wolfram syndrome) (74, 179, 181) and P58IPK (93), which lead to diabetic syndromes, have yet to be fully understood. Furthermore, the study of apparently unrelated disease states including atherosclerosis, viral
infection, and multiple myeloma may also benefit from
our increasing understanding of ER stress in their patho-
physiology.

ACKNOWLEDGMENTS

Address for reprint requests and other correspondence:
S. J. Marciniak, Cambridge Institute for Medical Research, Univ.
of Cambridge, Wellcome Trust/MRC Building, Hills Road, Cam-
bridge CB2 2XY, UK (e-mail: sjm20@cam.ac.uk); and D. Ron,
Skirball Institute of Biomolecular Medicine, New York University
School of Medicine, New York, NY 10016 (e-mail: ron@saturn.med.nyu.edu).

REFERENCES

1. Abastado JP, Miller PF, Jackson BM, and Hinnenbusch AG.
Suppression of ribosomal reinitiation at upstream open reading
frames in amino acid-starved cells forms the basis for GCN4 trans-
3. Adesnik M, Lande M, Martin T, and Sabatini DD. Retention of
mRNA on the endoplasmic reticulum membranes after in vivo
dissassembly of polysomes by an inhibitor of initiation. J Cell Biol
4. Aguirre V, Uchida T, Yenush L, Davis R, and White MF.
Aguirre V, Werner ED, Giraud J, Lee YH, Shoelson SE, and
White MF. Phosphorylation of Ser537 in insulin receptor sub-
strate-1 blocks interactions with the insulin receptor and inhibits
5. Aifantis I, Gounari F, Scorrono L, Borowski C, and von Boeh-
mer H. Constitutive pre-TCR signaling promotes differentiation
through Ca2+ mobilization and activation of NF-kappaB and NFAT.
6. Alexopoulou L, Holt AC, Medzhitov R, and Flavell RA.
Recognition of double-stranded RNA and activation of NF-kappaB
7. Allen JR, Nguyen LX, Sargent KE, Lipson KL, Hackett A, and
Urano F. High ER stress in beta-cells stimulates intracellular
degradation of misfolded insulin. Biochem Biophys Res Commun
8. Berlanga JJ, Santoyo J, and De Haro C. Characterization of a
mammalian homolog of the GCN2 eukaryotic initiation factor 2b
CE. Mechanism of interferon action. Purification and substrate
specificities of the double-stranded RNA-dependent protein kinase
from untreated and interferon-treated mouse fibroblasts. J Biol
Cho JH, West AB, and Ron D. Increased sensitivity to dextran
sulfate sodium colitis in Ire1b deficient mice. J Clin Invest 107:
Dynamic interaction of BiP and the ER stress transducers in the
12. Bole D, Hendershot L, and Kearney J. Post translational associ-
aton of immunoglobulin heavy chain binding protein with the
nascent heavy chains in non secreting and secreting hybridomas.
13. Bork P and Sander C. A hybrid protein kinase-RNase in an
D, Kaufman RJ, Ma D, Coen D, Ron D, and Yuan J. A selective
inhibitor of eIF2 alpha dephosphorylation protects cells from ER stress.
ENDOPLASMIC RETICULUM AND DISEASE


Amyloid diseases:


Molinari M, Calanca V, Galli C, Lucca P, and Paganetti P.
Mikulits W, Pradet-Balade B, Habermann B, Beug H, Garcia-
Mannik Y, Kaneto H, Kawamori D, Yoshiuchi K, Hatazaki M,
Munro S and Pelham HR.
Mueller PP and Hinnebusch AG.
Mueller PP, Harashima S, and Hinnebusch AG.
Meyer M, Caselmann WH, Schluter V, Schreck R, Hofschneider
Meyrath C, Hara K, Aso S, Narita M, and Ishihara K.
Marciniak SJ, Yun CY, Oyadomari S, Novoa I, Zhang Y, Jun-
Lu R and Misra V.
Lomas DA and Mahadeva R.


164. Stirling J and O’Hare P. CREB4, a transmembrane bZIP transcription factor and co-factor for 


206. Williams NP, Hinnenbusch AG, and Donahue TF. Mutations in the structural genes for eukaryotic initiation factors 2 alpha and 2 beta of *Saccharomyces cerevisiae* disrupt translational control of GCN4 mRNA. *Proc Natl Acad Sci USA* 86: 7515–7519, 1989.


