From the Division of Molecular Toxicology
Institute of Environmental Medicine
Karolinska Institutet, Stockholm, Sweden

Characterization of ERp29, a novel secretion factor of endoplasmic reticulum

Ernest Sargsyan

Stockholm 2005

Published and printed by Karolinska University Press
Box 200, SE-171 77 Stockholm, Sweden

ISBN 91-7140-337-X

ABSTRACT

ERp29 is a ubiquitously expressed endoplasmic reticulum protein strongly conserved in mammalian species. The N-terminal domain of ERp29 is similar to the thioredoxin domain of protein disulfide isomerase (PDI) lacking however the double-cysteine active site motif. The C-terminal domain is a novel five-helical fold. It was hypothesized that ERp29 may function as a molecular chaperone facilitating folding of the secretory proteins in the ER. The current investigation extended our knowledge of ERp29 providing the most up-to-date account of the genetic, molecular and functional features of this protein.

Gene structure and expression. Characterization and phylogenetic analysis of the rat, murine and human ERp29 genes demonstrated their common origin and close ortholog relationships. Such characteristics of the 5′ flank as CpG island, the absence of TATA-box, multiple transcription start sites in combination with Sp1-dependent basal transcription and ubiquitous gene expression indicate that ERp29 belongs to the group of constitutively expressed housekeeping genes. Exclusive expression of ERp29 in multicellular organisms in concert with its high expression in the specialized secretory tissues suggests a hypothetical secretory role for ERp29.

Functional activity of ERp29. ERp29 is induced upon the proliferation and differentiation of thyroid epithelial cells. Co-immunoprecipitation experiments, sucrose density gradient fractionation of the thyrocytes and affinity chromatography using immobilized Tg demonstrated ATP-independent association of ERp29 with thyroglobulin (Tg), the main secretory protein of thyroid cells, and ER chaperones (BiP and GRp94). Strong association of ERp29 with misfolded Tg and existence of the ER heterocomplexes including ERp29, Tg and other ER chaperones was shown also in the human and rat thyroid glands expressing mutant, transport-incompetent Tg. Surprisingly, despite the presence of the ER-retrieval signal, ERp29 is secreted synchronously with Tg. Overexpression of ERp29 significantly increases secretion of Tg whereas inhibition of ERp29 by RNAi gene silencing had an opposite effect, which strongly suggests the essential secretory function of ERp29 in the ER.

Substrate-binding sites of ERp29. Mutational analysis revealed two potential peptide-binding sites on the ERp29 surface. First, the highly mobile interdomain linker including a unique Cys157 was shown to be important for the functional activity of ERp29. Second functional site is located in the N-terminal domain and as judged by the analysis of the electrostatic surface is an uncharged cleft that might accommodate proteins of sufficiently large size.

ERp29 is a potential target of the unfolded protein response (UPR). We have investigated the potential involvement of ERp29 in the development of UPR. Activation of UPR pathways was demonstrated in the differentiating thyroid epithelial cells and in the endoplasmic reticulum storage diseases (ERSD) caused by the missense mutations in the Tg gene. In both cases we observed induction of ERp29 along with major ER chaperones suggesting that ERp29 is a potential UPR target gene.

In conclusion, our study describes ERp29 as a novel type of an ER auxiliary folding/secretory factor that facilitates transport of thyroglobulin and potentially of other secretory proteins to the cell exterior.

To my family

'That is the story. Do you think there is any way of making them believe it?'

' Not in the first generation', he said, 'but you might succeed with the second and later generations.'

(Plato, 380BC)

LIST OF PUBLICATIONS

I. **Sargsyan, E.,** Baryshev, M., Backlund, M., Sharipo, A., Mkrtchian, S. Genomic organization and promoter characterization of the gene encoding a putative endoplasmic reticulum chaperone, ERp29. *Gene*, 2002, 285, 127-39.

II. **Sargsyan, E.**, Baryshev, M., Szekely, L., Sharipo, A., Mkrtchian, S. Identification of ERp29, an endoplasmic reticulum lumenal protein, as a new member of the thyroglobulin folding complex. *J Biol Chem*, 2002, 277, 17009-15.

III. Baryshev, M., **Sargsyan, E**., Wallin, G., Lejnieks, A., Furudate, S., Hishinuma, A., Mkrtchian, S. Unfolded protein response is involved in the pathology of human congenital hypothyroid goiter and rat non-goitrous congenital hypothyroidism. *J Mol Endocrinol*, 2004, 32, 903-20.

IV. **Sargsyan, E.**, Baryshev, M., Mkrtchian, S. The physiological unfolded protein response in the thyroid epithelial cells. *Biochem Biophys Res Commun*, 2004, 322, 570-6.

V. Baryshev, M., **Sargsyan, E.**, Mkrtchian, S. ERp29 is an essential endoplasmic reticulum factor regulating secretion of thyroglobulin. *manuscript submitted.*

CONTENTS

LIST OF ABBREVIATIONS

Ala	alanine
ATF4	activating transcription factor
ATF6	activating transcription factor 6
ATP	adenosine triphosphate
BiP	immunoglobulin heavy chain binding protein
CFTR	cystic fibrosis transmembrane conductance regulator protein
CHOP	C/EBP homology protein
CNX	calnexin
COP I	coat protein type I
COP II	coat protein type II
CRT	calreticulin
Cys	cysteine
EDEM	ER degradation enhancing α-mannosidase-like protein
eIF2	eukaryotic initiation factor-2
EMSA	electrophoretic mobility shift assay
EOR	ER overload response
ER	endoplasmic reticulum
ERAD	ER associated degradation
ERGIC	ER-Golgi intermediate compartment
Ero1p	endoplasmic reticulum oxidoreductin
ERSD	ER storage disease
ERSE	endoplasmic reticulum stress element
FKBP	FK506 binding protein
Gln	glutamine
Glu	glucose
GPI	glycosylphosphatidyl inositol
GRP94	glucose regulated protein 94
GSH	reduced glutathione
GSSG	oxidized glutathione
HA	influenza hemagglutinin epitope
HSP	heat shock protein
IRE1	inositol-requiring element 1
LRP	low-density lipoprotein receptor–related protein
M6P	mannose-6-phosphate
Man	mannose
MEF	mouse embryonic fibroblast
MFG	milk fat globule
MFGM	milk fat globule membrane
MHC	major histocompatibility complex
MTP	microsomal triglyceride transfer protein
NE	nuclear envelope
NF-κB	nuclear factor-κB
NMR	nuclear magnetic resonance
NSF	N-ethylmaleimide-sensitive fusion

PC2	prohormone convertase 2
PDI	protein disulfide isomerase
PDILT	protein disulfide isomerase-like protein of the testis
PERK	pancreatic ER kinase
PPIase	peptidyl-prolyl isomerases
RACE	rapid amplification of cDNA ends
RAP	receptor-associated protein
RER	rough endoplasmic reticulum
RNAi	RNA interference
Ser	serine
SER	smooth endoplasmic reticulum
siRNA	small interfering RNA
SNAP	NSF attachment protein
SNARE	SNAP receptor
SRP	signal recognition particle
Tg	thyroglobulin
TGN	trans-Golgi network
TNF-α	tumor necrosis factor-α
TRAM	translocation-associated membrane protein
TSH	thyroid-stimulating hormone
Tyr	tyrosine
UPR	unfolded protein response
VTC	vesicular tubular clusters
XBP1	X-box binding protein 1

1 PREFACE

Currently, more than 500 genome sequencing projects are in progress, and about 300 prokaryotic and eukaryotic genomes have been completely sequenced (Genomes OnLine Database; http://ergo.integratedgenomics.com). The primary goal of such extensive genome sequencing efforts is to discover the molecular, biochemical and cellular functions of all the gene products. In the newly sequenced genomes genes are annotated mainly on the basis of sequence homology to already characterized proteins from other genomes. However, about 60% of predicted gene products either have no known homologs or show a homology with known genes and don't provide conclusive functional information (1). Therefore, to infer the function of such genes it is imperative to integrate sequence-based predictions with the traditional biochemical, cell and molecular biology "bench" methods. Such combination of approaches has been applied in current study dedicated to deciphering the regulation and physiological and molecular aspects of the function of a novel endoplasmic reticulum protein, ERp29.

2 INTRODUCTION

All eukaryotic cells secrete proteins. Some tissues, such as muscle, secrete small amounts of proteins whereas others, like endocrine glands are differentiated to release large quantities of specialized secretory proteins.
George Palade and colleagues discovered in 1975 that newly synthesized secretory proteins pass through series of membrane-enclosed organelles where the folding and post-translational modifications occur (fig. 1) (2,3). The "stations" on the way to the extracellular space include the endoplasmic reticulum (ER), the Golgi complex and secretory granules.

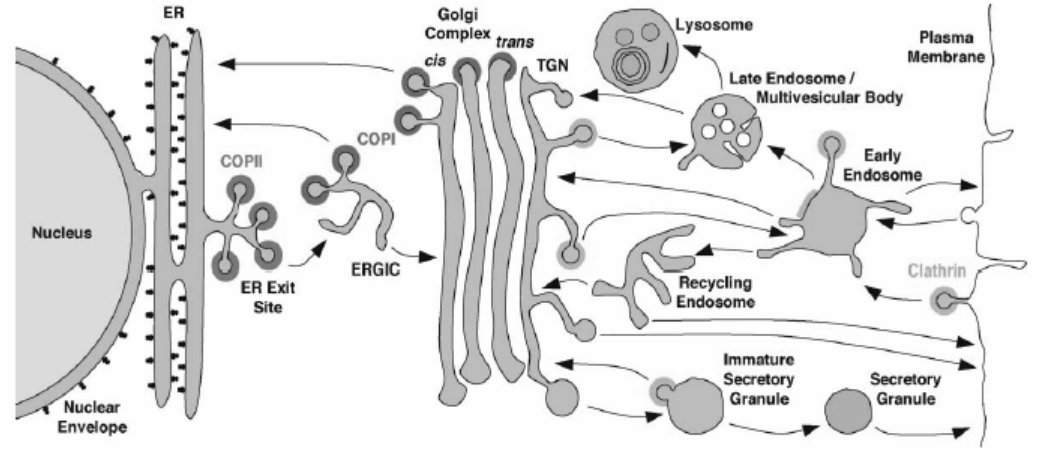

Figure 1. Intracellular Transport Pathways
The scheme depicts the compartments of the secretory, lysosomal/vacuolar, and endocytic pathways. Transport steps are indicated by arrows. Clathrin coats are heterogeneous and contain different adaptor and accessory proteins at different membranes. Only the function of COPII in ER export and of plasma membrane-associated clathrin in endocytosis are known with certainty. Less well understood are the exact functions of COPI at the ERGIC and Golgi complex and of clathrin at the TGN, early endosomes, and immature secretory granules. The pathway of transport through the Golgi stack is still being investigated but is generally believed to involve a combination of COPI-mediated vesicular transport and cisternal maturation. Adapted from Bonifacino (4).

Proteins destined for the plasma membrane, endosomes, or lysosomes share the early stations of this pathway (i.e., the ER and the Golgi complex) with secretory proteins. Trafficking between the compartments of the pathway is mediated by small carrier vesicles that bud from a "donor" compartment with incorporated cargo (secretory or membrane proteins) and travel to an "acceptor" compartment, into which they unload their cargo upon fusion of their membranes.
The processes of budding and fusion are repeated at consecutive transport steps until the cargo reaches its final destination within or outside the cell. To balance the forward movement of cargo organelle homeostasis requires the retrieval of transport machinery components and escaped resident proteins from the acceptor compartments back to the corresponding donor compartments ("retrograde transport"), a process that is also

realized by the vesicular transport. All of these steps are tightly regulated and balanced so that a large amount of cargo can flow through the secretory pathway without compromising the integrity and steady-state composition of the constituent organelles.

2.1 Targeting and translocation of ER proteins

Proteins entering the secretory pathway and ER-resident proteins start the journey from the cytoplasmic ribosomes with the subsequent co-translational translocation across the membrane into the lumen of ER (5-7).
These polypeptides contain hydrophobic ER targeting signals (6-20 amino acids) in their NH_2-terminus. These signals are recognized by the signal recognition particle (SRP) and directed to the ER membrane where components of the translocation machinery effectively couple translation and translocation (8-11).
Translocation takes place at a specific site, the translocon (12), which is an aqueous protein-conducting channel 40-60 Å in diameter (13-15). The core components of the translocation machinery are an integral membrane protein, Sec61, and the translocation-associated membrane protein (TRAM) (16). Sec61 is a trimeric complex that consists of an α-subunit, a multi-spanning integral membrane protein forming the channel walls, and β- and γ-subunits, each spanning the membrane a single time and playing an as yet undetermined role in the translocation (17-19). These four components, presented by two copies per translocon, are believed to be sufficient for the translocation activity. Additionally, a signal peptidase complex is required to cleave the signal peptide after the translocation (16,20).
However, some proteins containing signal sequences do not enter the SRP-dependent pathway, but are instead targeted to and translocated across the ER membrane post-translationally, i.e. after being fully synthesized and released from the ribosome. There are few examples of post-translational translocation in mammalian cells, but this is a major pathway in lower eukaryotes, for example in yeast (21). The unfolded conformation of the substrate proteins in the interval between the release from the ribosomes and passing through the ER membrane is maintained due to interaction with the cytosolic chaperones (22,23).

2.2 The endoplasmic reticulum

The first compartment of the secretory pathway is the endoplasmic reticulum. It is one of the largest cell organelles; its membrane constitutes over one-half of the total membranes in the cell and its lumen, the internal space, occupies over 10% of the cell volume. ER plays a vital role in many cellular processes, including synthesis, folding and post-translational modification of proteins, lipid biosynthesis and Ca^{2+} storage and release (24).

2.2.1 Structure-morphological characteristics of the ER

Despite continuous membrane, the ER can be subdivided into three morphologically distinguishable domains with diverse functions: the nuclear envelope (NE), the peripheral rough (RER) and the smooth ER (SER) (25).
The nuclear ER consists of two sheets of membranes with a lumen. The NE surrounds the nucleus, with the inner and outer membranes connecting only at the nuclear pores.
The peripheral ER is a network of interconnected tubules that extends throughout the cell cytoplasm (26). The RER and SER regions can be distinguished visually due to the morphological differences; for example, the SER is often more convoluted than RER, and the RER tends to be more granular in texture due to the presence of bound ribosomes (24).
Functionally, it is believed that the SER is a site of lipid biosynthesis, detoxification of xenobiotics and calcium regulation, whereas the RER is a site of proteins translocation and translation. Therefore the relative abundance of RER and SER found among different cell types correlates with their functions. Thus, cells producing large amounts of secretory proteins contain mostly RER, whereas steroid-producing, liver (detoxification) and muscle cells (calcium uptake and release for contraction) are abundant with SER (27).
Why are bound ribosomes concentrated in the RER and excluded from SER, rather than being distributed throughout the ER? One proposed explanation for the segregation of ER is that the functions associated with bound ribosomes (translocation and modification of newly synthesized proteins) are more efficient if the proteins performing them are concentrated in one part of the membrane (28).

2.2.2 Protein folding in the ER

Until recently, the principle of protein biogenesis relied entirely on the hypothesis that each peptide can self-assemble into a stable, low free-energy conformation, based on the information encoded by the amino acid sequence itself (29,30).
However, the maturation process in the ER involves co- and post-translational modifications, such as signal-peptide cleavage and disulfide bond formation, *cis-trans* isomerization of prolyl residues and N-linked glycosylation. These circumstances together with the high concentration of proteins in the ER (100-150 mg/ml) (31) lead to an increased possibility for improper intra- and intermolecular associations in the ER. To overcome these obstacles and to facilitate and accelerate the folding of newly synthesized secretory proteins one of the most important tasks of the ER is to provide an appropriate folding environment. Consistent with this role the ER is populated with the "folding assistants". Moreover, to maintain optimal functioning of the folding assistants, the ER provides a controlled milieu characterized by oxidized redox conditions (GSH/GSSG ~3-10) (32) and a highly regulated level of small ions and molecules, including Ca^{2+} (33), ATP (34) and sugars (35).
There are two classes of folding assistants in the ER: folding enzymes and molecular chaperones. The true ER folding enzymes introduce co- and post-translational covalent modifications into the newly synthesized proteins, whereas molecular chaperones facilitate protein folding through series of non-covalent interactions (36). Additionally, molecular chaperones monitor the fidelity of the biosynthetic events in the protein export pathway (ER quality control) (37). Depending on their substrate specificity,

molecular chaperones can be subdivided into 3 groups: general chaperones, lectin chaperones and non-classical molecular chaperones.

2.2.2.1 General chaperones

General molecular chaperones interact transiently with the solvent-exposed hydrophobic surfaces of the nascent or denatured polypeptides and protect them from undesirable aggregation.
BiP is a central molecular chaperone of the ER present in all eukaryotic organisms and identified as an immunoglobulin heavy chain binding protein and as a glucose-regulated protein (therefore it also known as GRP78) (38). It belongs to the family of 70 kDa heat shock proteins (HSP70) and like other members of the family contains two domains: the N-terminal ATPase domain and the C-terminal substrate-binding domain (39). Interaction of BiP with the substrates occurs through the cycles of association/dissociation (40-42), where dissociation is an ATP-dependent process. Binding of BiP to misfolded or unassembled proteins is more permanent (43).
Additionally, BiP binds the Sec61 translocation complex and, hydrolyzing ATP, provides the driving force for anterograde and retrograde protein translocation across the ER membrane (44-46). BiP acts either as a molecular motor, actively pulling the substrate into the lumen (47), or as a Brownian ratchet, preventing the backslide of the nascent chain (48,49).
GRP94 is another abundant ER molecular chaperone that was first described as a protein induced by glucose starvation of Rous Sarcoma Virus-transformed chick embryo fibroblasts, hence the acronym GRP (glucose regulated protein) (50). GRP94 is present in most multicellular organisms (except *Drosophila)* where it accounts for 5-10 % of the luminal content with a concentration of 10 mg/ml (51). However, its role is still enigmatic and the mode of action is poorly understood. There is a hypothesis that GRP94 plays a role in assembly or serving as a scaffold for the multi-protein chaperone complexes that, by virtue of their size and composition, prevent incompletely folded or assembled proteins from continuing through the secretory pathway (52,53). Additionally, it has been discovered that GRP94 possesses antigen-presenting functions and modulates both the innate and adaptive components of the immune system (54).

2.2.2.2 Lectin chaperones

The ER is the only intracellular site where glycoproteins are synthesized and matured. Glycosylation of proteins in the ER is asparagine-linked and accomplished in two steps: co-translational *en bloc* transfer of the oligosaccharide core (Glu_3-Man_9-$GlcNAc_2$) and subsequent removal of glucoses by glucosidase I and II enzymes (35). Such modification might increase solubility and also serve as a marker of a correct folding.
Lectin chaperones (calnexin (CNX) and calreticulin (CRT)), which are unique for the ER, are essential for the folding of glycoproteins. The significance of this system is underscored by the fact that calnexin and calreticulin interact with practically all glycoproteins investigated to date (55).

Calnexin is a transmembrane protein and calreticulin is a soluble protein, very similar to the luminal domain of calnexin. They assist glycoprotein folding through the binding to monoglycosylated form of the substrates that appears either as a trimmed intermediate of the triglycosylated core oligosaccharide or by readdition of a glucose residue by the UDP-glucose: glycoprotein glucosyltransferase to the fully deglucosylated glycan (56-58). The process of disulfide bond formation in the glycoprotein substrates of CNX and CRT is assisted by a glycoprotein-specific thiol-disulfide oxidoreductase ERp57, which is non-covalently associated with both proteins (59,60).

2.2.2.3 Folding catalysts

Proteins synthesized in the ER are often stabilized by intermolecular disulfide bonds that form as the protein folds into its three-dimensional structure (61). Fast and correct folding of such proteins occurs due to the specialized redox environment and the presence of catalysts.
One of the most essential folding catalysts is a protein disulfide isomerase (PDI), a 55 kDa major component of the ER (62-64).
PDI belongs to a large family of dithiol/disulfide oxidoreductases, the thioredoxin superfamily (65). All four domains of PDI are structural homologs of thioredoxin but only two of them contain an independent CXXC active site motif (66). Ero1p and its assistant, ERp44, transfer oxidizing equivalents to these double cystein motifs of PDI, which in turn introduce disulfides into the substrates (oxidase activity) and rearranges incorrect disulfide bonds (isomerase activity) (67,68).
In addition to the oxidoreductase activity PDI may behave also as a general chaperone, inhibiting the aggregation of misfolded proteins (69-71). Chaperone activity does not require the catalytic cysteines and is manifested even with substrates that contain no disulfide bonds (72).
Interestingly, at low concentrations (1/5–1/10 that of substrate) PDI may display an anti-chaperone activity, inducing aggregation and precipitation of normally soluble, misfolded proteins (73).
Other identified members of the oxidoreductase family that contain thioredoxin-like domains are ERp57, ERp72, ERp44, P5, ERdj5, and PDILT (65,74).

Peptidyl-prolyl isomerases (PPIase) (also termed immunophilins) catalyze *cis-trans* isomerization of proline residues, and enhance the rate of protein folding (75,76). PPIases are comprised of two classes: the cyclophilins, which bind drugs of the cyclosporine A group (77), and the FKBPs, which bind drugs of FK506 and rapamycin groups (78).

2.2.2.4 Non-classical molecular chaperones

This group of molecular chaperones may be conditionally divided into two subgroups: dedicated chaperones and molecular escorts (escort chaperones).

The best studied dedicated chaperone is HSP47 (79). This protein performs specific chaperone function throughout all stages of structural maturation of collagen (80,81). Although it may act in association with other ER chaperones, the services of HSP47 are unique to the collagen-secreting cells.
Another member of this subgroup is a microsomal triglyceride transfer protein (MTP), which is essential for the folding of apolipoprotein B. It has been shown that the absence of functional MTP leads to rapid intracellular degradation of apolipoprotein B and development of abetalipoproteinemia (82).

Molecular escorts provide helper function to the protein export pathway and deal with already folded proteins entering the ER exit sites in order to leave to Golgi (83). Unlike the classical ER chaperones, molecular escorts may escape ER and travel through the secretory pathway. Some escort chaperones can also be considered as dedicated chaperones because of their specificity.
The two well-known members of this subgroup are receptor-associated protein (RAP) and 7B2. RAP assists the folding of low-density lipoprotein receptor–related protein (LRP) immediately after its biosynthesis and travels with it to the cell surface in order to protect LRP from the premature interactions with lipoproteins (84-86). Recently it was shown that RAP may function also as a thyroglobulin (Tg) molecular chaperone, since it is involved in Tg secretion (87). 7B2 is expressed in the ER of the neuroendocrine tissues and interacts with the prohormone convertase 2 (PC2) (88,89). This complex leaves the ER and dissociates in the most distal portions of the Golgi complex (90). Absence of 7B2 leads to a slower and less efficient transport of PC2 through the secretory pathway (91) suggesting a molecular escort function for this protein.

2.2.3 Quality control

As already mentioned, in addition to assistance in folding molecular chaperones subject proteins to the stringent quality control to ensure that only correctly folded proteins exit the ER whereas unfolded or misfolded proteins are retained and ultimately degraded (92,93). Such machinery reduces the formation and accumulation of misfolded proteins that may form aggregates and potentially harm the cell. Molecular chaperones control the quality of the newly synthesized chains via the repetitive cycles of association and dissociation. Normal and misfolded forms of proteins are differentiated by the recognition of the structural signals, such as exposed hydrophobic sites or mobile loops in the incompletely folded molecules. Successful export from the ER takes place only when all chaperone-binding sites on the exportable polypeptide are buried. In case of prolonged exposure of structural signals, caused by inability of proteins to reach their native conformation, association with chaperones is more stable and incompletely folded proteins retain in the ER.

2.2.4 The Unfolded Protein Response

Excessive accumulation of the unfolded proteins in the ER invokes the unfolded protein response (UPR), the adaptive cellular reaction that coordinates down-regulation of overall protein synthesis and upregulation of molecular chaperones and other genes implicated in secretion and degradation (94-97).

Impaired folding of proteins occurs under ER stress conditions, that can be caused by different pharmacological agents, such as tunicamycin (inhibition of glycosylation), thapsigargin (Ca^{2+} depletion), brefeldin A (inhibition of intracellular protein transport) (98,99) and by overexpression of structurally defective secretory proteins (ER storage diseases) (see section 2.7) (100,101).

Three distinct pathways of the UPR are identified based on the different UPR sensors that transmit the stress signal across the ER membrane (fig. 2). In quiescent state all of these receptor molecules are silenced by the interaction with BiP, whereas the accumulation of unfolded proteins releases BiP with the ensuing activation of the UPR sensors (102-105).

ATF6α, a 90 kDa membrane-bound protein is translocated to the Golgi and processed by site 1 and site 2 proteases in response to ER stress (102,106-108). The ensuing 50 kDa fragment is an active transcription factor regulating the expression of ER molecular chaperones and other UPR target genes via binding to the endoplasmic reticulum stress element (ERSE) in the corresponding promoters.

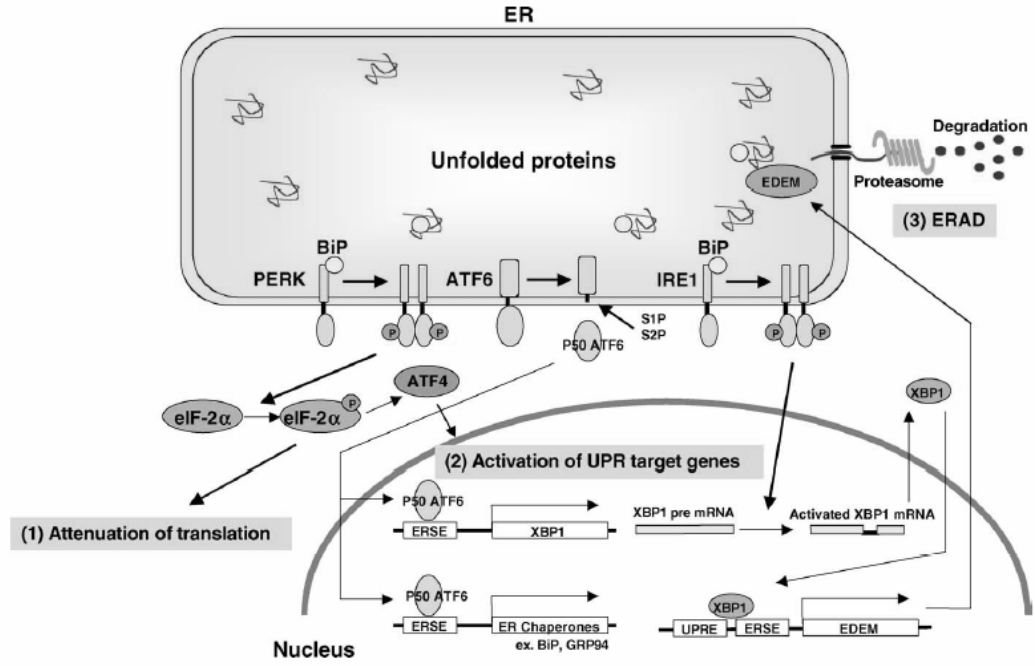

Figure 2. Mechanisms of cell survival in response to ER stress.

Accumulation of misfolded proteins in the ER activates the ER stress sensors, such as IRE1, ATF6, and PERK that mediate ER stress responses. The survival signal is regulated by three main responses; transcriptional activation of the ER target genes that are mediated by IRE1, ATF4, and ATF6; translational attenuation that is mediated by PERK-eIF2α pathway; and the ERAD pathway to restore the folding capacity that is regulated by IRE1–XBP1 pathway. Adapted from Kadowaki (109).

Active IRE1, an ER transmembrane kinase, catalyses the splicing of mRNA of the transcription factor XBP1, removing the 26 bp intron (71,96,110). Synthesis of the

54 kDa protein encoded by the spliced XBP1 mRNA produces a potent transcription factor, involved in the induction of molecular chaperones and also genes implicated in protein degradation and probably secretion (97,111-113). Activation of another sensor of UPR, the transmembrane kinase PERK, is characterized by the general attenuation of the protein translation (114) via the phosphorylation/inhibition of the general translation factor, eIF2α. Additionally, active PERK promotes the expression of the selected set of genes including the transcription factor ATF4 (113,115,116).
The aforementioned signaling mechanisms were extensively studied in the drug-induced ER stress. However, the UPR might be activated also under physiological conditions, when urgent need in the biosynthesis of certain secretory proteins may lead to the overload of ER with the nascent polypeptides. Physiological UPR is studied very poorly. The most known example is the activation of UPR upon differentiation of B-lymphocytes into the antibody-secreting plasma cells and in the β-cells of the pancreas (117,118).

2.2.5 ER-associated degradation

If proteins are not able to pass a quality control and reach their native conformation, even after an extended interaction with the molecular chaperones and despite of rescuing role of UPR, they need to be removed from the folding pathway.
Such extraction is regulated by the ER α-mannosidase I (119) and by EDEM (ER degradation enhancing α-mannosidase-like protein (120), an enzymatically inactive mannosidase-like protein. First, the ER α-mannosidase I cleaves one (121) or more (122) mannose residues from protein-bound *N*-glycans. Then such proteins, tagged with *N*-glycans with a reduced number of mannoses are selectively recognized by EDEM, polyubiquitinated and retrotranslocated via the Sec61 channel back into the cytosol, where they undergo degradation by the proteasomes (fig. 2). This pathway was termed ER-Associated Protein Degradation (ERAD) (93,123).

2.3 ER to Golgi transport

Secretory and membrane proteins that successfully have passed the quality control are transported to specialized regions called ER exit sites and packaged in vesicles (124-127). It has been postulated that the inclusion of cargo in the ER-derived vesicles is a selective process that may involve transmembrane proteins acting as cargo receptors (128-130). These receptors include ERGIC-53/p58 lectin, which selectively and transiently interacts with glycoproteins (131) and p24 protein family members (132), which are implicated in the selection of secreted proteins and exclusion of ER resident proteins.
However, the previously suggested bulk flow hypothesis, i.e. nonselective protein recruitment (133), cannot be completely ruled out, at least for soluble proteins, as some ER-residents may escape the transport vesicles and are recycled back by a special retrieval mechanism (see section 2.4).
Vesicles are formed by type II coats (COPII), supramolecular assemblies of proteins that deform flat membrane patches into round buds, eventually leading to the release of

coated transport vesicles. These vesicles form the vesicular tubular clusters (VTC) or ER–Golgi intermediate compartment (ERGIC) (134), and are subsequently transported to the Golgi complex.
There is more than one type of coated vesicles. Forward movement from ERGIC to Golgi complex as well as retrograde transport to ER requires other types of coats, COPI, whereas clathrin-coated vesicles operate in the post-Golgi locations.
Fusion of the vesicles with target membranes is a specific process and achieved by the interaction of a specific "v-SNARE", carried by all transport vesicles and a cognate "t-SNARE" on the target membranes (135,136).

2.4 Protein retention in the ER

ER residency for the members of the ER folding and maturation machinery is achieved by two complementary systems, one ensuring retention in the ER and one acting to retrieve escapees from the Golgi compartment back into the ER.
Best understood is the retrieval of soluble ER resident proteins bearing the C-terminal H/KDEL-type sequence motif (137-140) and transmembrane proteins bearing either cytosolic di-lysine (KKXX) or di-arginine (RR) motifs or a lumenal HDEL signal (141).
KDEL-bearing proteins that escape from the ER, are recognized by the KDEL receptor, ERD2, ERGIC-resident integral membrane protein (142,143) and cycled back to the ER.
Transmembrane proteins, bearing dibasic motif, either di-lysine or di-arginine, bind the components of COPI and are retrieved by retrograde vesicular transport (144).
Accumulating evidence indicates that the ER localization of resident proteins is attributable not only to their retrieval from post-ER compartments but also to the true retention mechanisms. Thus, it has been shown that the di-lysine retrieval motif can also serve as a retention signal (145). However, there are a number of ER-resident proteins that do not possess any known retrieval sequences. Retention of such proteins might be mediated by weak protein-protein interactions resulting in the formation of large oligomeric structures that are not capable of exiting the ER (146).

2.5 The Golgi complex

On the way to the final destination, secretory and membrane proteins pass through the Golgi apparatus, where modification of oligosaccharide chains of glycoproteins, phosphorylation and other post-translational modifications take place.
Structurally, the Golgi complex consists of cisternal stacks separated by tubulovesicular domains (147,148). Functionally, it is divided into three different subcompartments, the *cis*-, *medial*- and *trans*-Golgi (149), where proteins travel in *cis*-to-*trans* direction.
The exact mechanisms of the protein transport in the Golgi and the role of vesicles in this process remain enigmatic. At least two alternative hypotheses and models have been suggested. In the first model, both anterograde transport of cargo proteins and retrograde transport of Golgi resident proteins are mediated by COPI-coated vesicles.

In the alternative model, cargo proteins progress through the Golgi stacks by a process of cisternal maturation, whereas COPI-coated vesicles balance this process by a return flow of Golgi resident proteins (150,151). In this model, vesicles play only a minor role in the anterograde transport of cargo through the Golgi.

2.6 Post-Golgi biosynthetic trafficking

When proteins exit the Golgi complex they enter the trans-Golgi network (TGN), the major sorting station for newly synthesized proteins and lipids in the biosynthetic pathway. From here a number of traffic pathways emerge such as lysosomal biogenesis via the endosomal system and different constitutive and regulated routes that deliver proteins to the plasma membrane (152,153). Regulated pathway is found in the cells that are specialized for secreting rapidly on demand products such as hormones and neurotransmitters. In this pathway, proteins are initially stored in the secretory granules for later release. The delivery of proteins is believed to be mediated by sorting signals and specific receptors (154-158)

2.7 ERSD

In recent years a number of diseases characterized by inborn errors leading to abnormal protein trafficking have been identified. Functional defects could occur at any of the transport and sorting steps, however the vast majority of mutations affect protein folding in the ER. Such disorders characterized by the accumulation of the proteins in the ER due to the mutations in their primary structure and defects in protein folding are known as endoplasmic reticulum storage diseases (ERSDs) (100,159).
ERSDs can be subdivided into two pathogenetic groups (101).
In the first one, that most ERSDs belong to, the defect lies within the cargo, i.e. the secretory protein is encoded by a mutated gene. In this case disease can develop via two mechanisms. First, the protein activity may lack at its normal site of action and second, more common, the folding mutants may form toxic ER lumenal or cytoplasmic aggregates or toxic degradation products and thus compromise functionality and viability of affected cells.
Typical cases of ERSDs in this group are cystic fibrosis, congenital hypothyroidism, Alzheimer's disease, albinism and many others. Thus, cystic fibrosis is caused by mutation in the cystic fibrosis transmembrane conductance regulator protein (CFTR) that results in the retention of the protein in the ER and eventual degradation by proteasomes (160-162). In congenital hypothyroidism thyroglobulin is accumulated in the ER due to the mutation in the gene and formation of large aggregates (100).
In the second group of ERSDs the defect lies within the gene encoding the protein of the transport machinery. This group, for instance, includes combined coagulation factor V and VIII deficiency, caused by the mutation in ERGIC-53 (163) and abetalipoproteinemia, caused by the mutation in the microsomal triglyceride transfer protein (164,165).

2.8 ERp29

ERp29 is a 24.5 kDa endoplasmic reticulum protein for the first time discovered and characterized in our laboratory (166,167) and almost simultaneously also by others (168,169).

ER localization of ERp29 is suggested by the presence of the N-terminal ER targeting signal and the C-terminal tetrapeptide KEEL, the conserved variant of KDEL, the ER retrieval signal. Experimentally, this was confirmed by the immunofluorescent techniques in different mammalian cells and immunodetection almost exclusively in the microsomal fractions of different tissues. ERp29 is not an abundant ER protein; it accounts for approximately 0.1 % of total microsomal proteins (167). However, the ubiquitous pattern of gene expression suggests an important "house-keeping" function for ERp29 (167,168). Interestingly, in contrast to many ER luminal proteins post-translational modifications such as glycosylation (167), phosphorylation (169) and Ca-binding activity (167,170) are not characteristic for ERp29.

ERp29 is comprised of the N- and C-terminal domains (fig. 3) (171). The amino acid sequence of the N-terminal domain shares 25% of identity with the thioredoxin domain of protein disulfide isomerase and as has been shown by NMR analysis, this similarity extends to their 3D structure. However, this domain of ERp29 does not contain the double-cysteine active site, excluding any PDI-like thiol/disulfide oxidoreductase activity.

Size-exclusion chromatography, cross-linking and dynamic scattering showed that ERp29, like many molecular chaperones, forms homodimers and multimers both *in vivo* and *in vitro* (171,172). Interestingly, ERp29 is the first example of a protein where a thioredoxin-like domain acts as a dimerization module without supporting interactions by covalent bonds or additional contacts by other domains. NMR analysis has predicted that dimerization sites as well as other putative protein binding sites are located in the N-terminal domain.

The C-terminal domain is a novel five - helical fold, which is very similar to the corresponding domain of P5-like PDIs, to the *Drosophila* ortholog of ERp29, Wind (172) and similar proteins from the honey bee and mosquito. It was hypothesized that the most important function of the C-terminal domain is a solubilization of the N-terminal domain (171) and retrieval of the ERp29 in the ER, as it contains the ER-retrieval signal, KEEL.

Two domains of ERp29 are connected by the short mobile loop (amino acid residues 149-159), the type of the structural element often implicated in the protein-protein interactions.

The function of ERp29 at the time of initiation of this thesis work was not yet clarified. The absence of the double-cysteine motif in the thioredoxin domain of ERp29 excludes the oxidoreductase activity. Therefore it was hypothesized that ERp29 probably lost its redox activity in the course of evolution, but may still retain general chaperone features of PDI (172). Indirectly, this notion was supported by such characteristics of ERp29 as co-induction with other molecular chaperones and association with BiP.

However, the substrates or putative partners of ERp29 have not been identified.

Moreover, ERp29 lacks the classical chaperone activity as judged by the inability to protect substrate proteins against thermal aggregation (unpublished data).

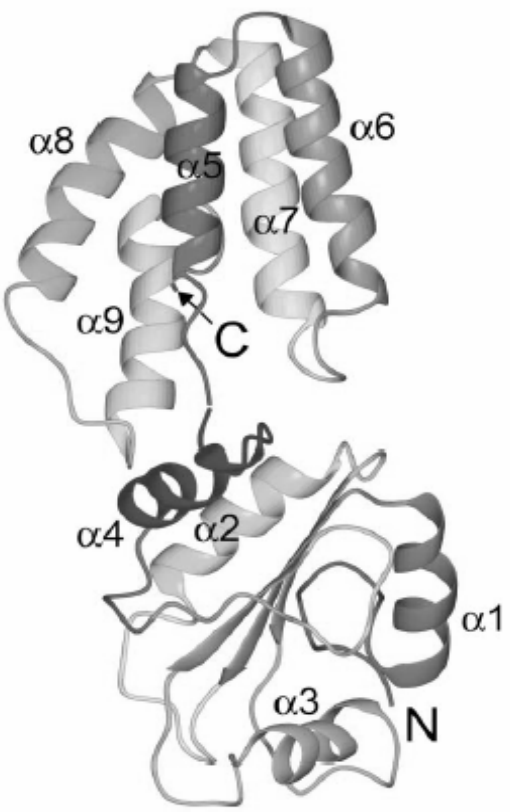

Figure 3. Ribbon Diagram of an ERp29 Monomer as Determined by NMR Spectroscopy
The structures of the N-terminal domain (bottom) and the C-terminal domain (top) were determined individually. The linker segment, interrupted between residues Met154 andPro155 in the drawing, is continuous in the full-length protein.
Adapted from Liepinsh (177).

Almost simultaneously with our group, ERp29 was discovered also in two other laboratories. The Hubbard group isolated ERp29 from the rat enamel cells and demonstrated its upregulation during enamel secretion (168). They demonstrated that ERp29 is highly expressed in the secretory tissues and suggested that it might be involved in the synthesis of secretory protein (173).
The Ferrari group isolated the human analog of ERp29 and termed it ERp28 (169). Although they did not detect induction of ERp29 under ER stress conditions and interaction with BiP, they showed co-immunoprecipitation of human ERp29 with overexpressed hepatitis B small surface antigen, suggesting the unique role in the processing of secretory protiens.
Additionally, in *Drosophila*, the *windbeutel* gene product, Wind, which is highly similar to ERp29, was shown to be important for the function and specific Golgi targeting of the patterning protein Pipe, a putative oligosaccharide-modifying enzyme essential for embryonic development (174).

3 AIMS

The general aim of the project was to characterize and analyze the function and regulation of ERp29.

The more specific objectives were:

- To describe the genomic organization and to characterize the 5´-flanking region of ERp29
- To elucidate the role of ERp29 in the folding and secretion of thyroglobulin, a major secretory product of thyroid cells
- To identify the substrate binfing sites of ERp29
- To clarify the role of UPR in the regulation of ERp29 upon the differentiation of the thyroid epithelial cells and in the ERSDs, caused by missense mutations in thyroglobulin (Tg) gene

4 RESULTS

4.1 Genomic organization of ERp29 (paper I)

The rat ERp29 genomic clone selected by the PCR screening of the P1 phage genomic library with ERp29-specific primers was used for the sequencing of the introns utilizing gene-walking approach. We found that the 6.4 kb-long gene contains three relatively small exons separated by the large (4.7 kb) and small (0.38 kb) introns. The homology search in GenBank and *Celera* databases identified complete sequences of mouse and human ERp29 genes with almost identical exon/intron structures (intron 2 is substantially longer in the mouse and human genes).

The search in all available complete genome databases revealed that ERp29 is expressed exclusively in multicellular organisms, such as vertebrates and some invertebrates (*Drosophila*) and absent from the monocellular life forms.

Phylogenetic analysis of ERp29s from different mammalian species and related PDIs showed clustering of mammalian ERp29 sequences where rodent and murine species form a distinct subgroup. Wind may be tentatively included in the ERp29 family as a distant ortholog member although the comparison of the promoters did not reveal significant similarity suggesting the diverse regulation of these genes.

Cloning of the 5′-flanking region of ERp29 revealed a CpG island in the region spanning from the nucleotide −469 to −2. Such islands are frequently found close to the transcription initiation sites of housekeeping genes (175). Interestingly, the first 600 bp of rat 5′ flanks exhibit more than 60% sequence identity with the human gene and more that 80% similarity with the mouse gene.

Screening of 5′ flanks of all three genes failed to locate canonical TATA-like sequences and CAAT boxes, the elements that commonly specify the transcription start site in most of the genes. However, we mapped two GC boxes, characteristic for constitutively active genes (176). Sp1 binding to GC boxes in such promoters is critical for the transcription initiation (177,178), which is often directed from the multiple sites. Indeed, the modified 5′-RACE procedure, SMART 5′-RACE, identified several transcription start sites at positions −148, −87 and −46.

To delineate the DNA elements responsible for the basal expression of the ERp29 gene and to identify the ccore promoter region, series of 5'-deletion constructs were tested in the rat hepatoma FAO cells using dual-luciferase reporter system. A 337 bp fragment was found to manifest a maximal promoter strength. Conservation of this region between rat, mouse and human sequences indicates its importance for the transcriptional regulation of the ERp29 gene.

Based on the structural and functional analysis of the ERp29 promoter, we sought to identify the trans-acting nuclear factors that might be responsible for the constitutive activation of the ERp29 gene. Electrophoretic gel mobility shift assay (EMSA) demonstrated an interaction of SpI/Sp3 and USF transcription factors with predicted GC and E-box elements respectively, which suggests a key role of these transcriptional factors in the basal expression of ERp29 gene.

4.2 Tissue distribution (paper I)

To examine tissue-specificity of ERp29 gene we hybridized ERp29 cDNA specific probe with the human RNA dot blot array. Expression of ERp29 was detected in all tissues represented on the array with the highest level in the tissues with high secretory activity such as pituitary, adrenal, thyroid and salivary glands, prostate, pancreas, and additionally, liver and kidney. On the other hand, the level of ERp29 expression was substantially lower in testis and ovary, also active secretory tissues, which however, export mostly low molecular weight molecules such as testosterone and estrogen. ERp29 mRNA was detected in all tested fetal tissues, suggesting its expression at the early stages of development.

4.3 ERp29 is a novel folding/secretory factor (paper II, paper III, paper V)

High level of ERp29 gene expression in the thyroid suggested that ERp29 might be implicated in the maturation and/or secretion of Tg, a major secretory product of these cells. This was also supported by the induction of ERp29 mRNA in the thyrocytes treated by the thyroid-stimulating hormone (TSH) (179).
We found that hormonal stimulation of FRTL-5 thyroid epithelial cells up-regulates ERp29 concurrently with other ER resident proteins, such as BiP, GRP94, and PDI. It suggests a putative chaperone-related function for ERp29, in which case one would predict its association with the nascent Tg molecules. Indeed, immunoprecipitation analysis demonstrated co-immunoprecipitation of ERp29, Tg and also ER chaperones, BiP and GRP94, suggesting the existence of large folding heterocomplexes of Tg in the ER. FRTL-5 cell lysate fractionation using sucrose density gradient centrifugation and subsequent analysis of the obtained fractions confirmed the presence of ERp29, BiP and Tg in the same complexes.

To further investigate possible interactions of ERp29 within the Tg folding complexes in the ER, we employed affinity chromatography incubating FRTL-5 lysates with the Sepharose-immobilized Tg and other ligands. It has been shown earlier that molecular chaperones bind immobilized denatured Tg, and other ligand proteins in an ATP-dependent manner (180). Surprisingly, we found that ERp29 associates with both native and denatured Tg although with a stronger affinity to the latter. The binding of ERp29 to the Tg-Sepharose beads was ATP-independent and seemed to be rather specific because no interaction was observed neither with the non-denatured nor the denatured histone.
Association of ERp29 with Tg and existence of the ER heterocomplexes including ERp29, Tg and other ER chaperones was confirmed also in the experiments using specimens from the human and rat thyroid glands expressing mutant, transport-incompetent Tg. Co-immunoprecipitation analysis revealed significant association of ERp29 with mutant Tg and somewhat weaker association with the completely folded Tg. This correlates perfectly with the affinity chromatography data obtained in FRTL-5

cells and lends more support to the putative role of ERp29 as an escort chaperone that may retain association with Tg even after the latter completes its folding.

It has been previously reported that a number of ER-resident proteins including PDI may be secreted (181-184). Despite the presence of ER-retrieval signal substantial amounts of ERp29 were also found in the culture medium of the thyroid cells. Moreover, the synchronous increase of ERp29 and Tg in the medium upon the hormonal stimulation and inhibition of secretion by brefeldin A, an ER-to-Golgi transport inhibitor, indicates true secretory nature of this process. Additionally, immunofluorescence microscopy showed a very high degree of co-localization of ERp29 and Tg, apparently in the same vesicular transport structures.

Collectively, these data clearly suggest a functional significance of ERp29 in the folding/secretion of Tg. To substantiate this assumption we attempted to answer the question whether the secretion of Tg can be regulated by manipulating the ERp29 levels in the thyrocytes. Indeed, overexpression of ERp29 was accompanied by the explicit, nearly 2-fold increase of Tg in medium, which was dependent on the amount of the transfected plasmids. In line with this, the intracellular levels of Tg were reciprocally decreased by the ERp29 overexpression. Choosing the opposite approach we then down-regulated ERp29 using the RNA interference (RNAi) technology. Although the silencing was not particularly strong (~50 % of the control level), it had quite a significant impact on the level of Tg secretion reducing it by ~ 40%. Finally, the involvement of ERp29 in the Tg secretion was demonstrated by the expression of certain ERp29 mutants acting apparently in the dominant-negative fashion.

4.4 Mapping of substrate binding sites in ERp29 (paper V)

Having established the importance of ERp29 for the Tg transport we sought to identify the active site(s) and amino acid residues that are involved in the putative ERp29-Tg interaction. Here we applied mutational analysis using structural information on the ERp29 domains obtained by NMR-spectroscopy and, in addition, recently published crystal structure of Wind.

First, our attention was drawn to the highly conserved linker region that connects two domains of ERp29. Presence of three glycines and secondary structure data predict rather unstructured, surface-exposed and highly mobile loop, the structural element that is frequently involved in the protein function.

Indeed, expression of different deletion mutants in FRTL-5 cells resulted in substantial inhibition of the Tg secretion. Moreover, replacement of the unique solvent-exposed cysteine, located in the linker region, with either alanine or serine had a profound effect reducing the Tg secretion down to nearly 40% of the control level. Collectively, these data highlight the essential role of the linker region for the ERp29 function.

Based on the reported structure of Wind we mutated two amino acids, Tyr and Gln, located in the region corresponding to the functionally important region in Wind. We found that both mutations moderately affect the transport of Tg reducing it down to 60%. These data indicate the importance of this region for the ERp29-Tg association.

NMR study has revealed that two loci in the N-terminal domain may interact with the linker, which was interpreted as either the mechanism for oligomerization of ERp29 or, alternatively, that such association may mimic interaction with unfolded substrates exemplified by the unstructured, mobile linker region. However, expression of the ERp29 mutants where the residues from both sides were replaced couldn't modify the secretion of Tg.
Finally, we sought to investigate the effect of the expression of the individual domains of ERp29. Therefore we constructed N- and C- terminal domains so that they would be targeted and retained in the ER similar to the full-length protein. In both cases expression of domains suppressed Tg secretion, albeit N-terminal construct exhibited much stronger effect.

4.5 Is ERp29 a target of UPR?

4.5.1 Physiological UPR (paper IV)

Induction of ERp29 under ER stress and simultaneous upregulation of ERp29 and molecular chaperones upon stimulation of thyroid cells indicates common mechanisms underlying these phenomena and suggests a potential involvement of UPR.
Since little is known about the development of UPR under physiological conditions we decided to examine the activation of major UPR pathways in the differentiating thyroid cells. We analyzed the kinetics of Tg synthesis in the thyroid cells upon stimulation and compared it with the expression of ERp29 and other ER chaperones. Hormonal treatment had relatively weak effect on the chaperone expression; however the lag phase between the Tg induction and the upregulation of chaperones was quite explicit, suggesting the triggering role of the Tg expression in the physiological UPR in thyroid cells.
Activation of the key UPR sensor molecule, ATF6 was also modest as compared to the ER stress-mediated processing of ATF6α and occurs only for a short period after 4 hours of stimulation. Equally weak activation of ATF6α has been also shown in the differentiating B lymphocytes (117).
Activation of PERK in stimulated thyroid cells was examined by testing the expression of the downstream transcription factors, ATF4 and CHOP. Weak upregulation of both factors was detected by 24-48 h of stimulation, which is consistent with the poor expression of p50ATF6 and similarly weak induction of molecular chaperones.
The IRE1/XBP1 pathway of UPR was studied by monitoring the splicing of 26 bp fragment from the XBP1 mRNA. This reaction is catalyzed by the active IRE1 and results in the reading frame shift and translation of the transcriptionally active XBP1 protein. We failed to detect the presence of the spliced form of XBP1, however the real-time PCR analysis showed strong, time-dependent accumulation of the unspliced form of XBP1 mRNA. Surprisingly, sequencing of the *PstI*-resistant product corresponding to the unspliced form revealed the mixture of unspliced and spliced forms suggesting a certain level of IRE1 endonuclease activity.
In general, we demonstrated modest activation of UPR pathways in the differentiating thyroid cells and proposed the accumulation of Tg nascent chains in the ER as an initial trigger of the UPR.

4.5.2 UPR in the ER-linked pathologies (paper III)

In this work we examined the possible link between the induction of molecular chaperones, including ERp29 and the activation of UPR transducers, ATF6 and XBP1, in the pathologies caused by the expression of mutant Tg.

We analyzed thyroids obtained from patients with congenital goiter and *rdw* dwarf rats with non-goitrous congenital hypothyroidism. Both pathologies are characterized as ERSDs and caused by the missense mutations in the Tg gene (185-187). These mutations make Tg molecules unable to leave the ER, which results in the impairment of the synthesis of thyroid hormones. Such defective molecules are retained in the ER by the quality control machinery with the ensuing induction of the major ER chaperones, BiP and GRP94. We have reproduced these results and additionally demonstrated similar upregulation of ERp29 and a number of cytoplasmic and mitochondrial chaperones, which is an indication of the mobilization of the general cellular and not only the ER resources to cope with such stress situation.

Among the major UPR sensors, the inactive form of ATF6, 90 kDa protein was readily detected in all tissues tested. The active form of ATF6, 50 kDa protein appeared in the mutant patients and also in the homozygote *rdw* rat thyroids unambiguously indicating an active UPR.

Activation of XBP1 was examined as previously, by monitoring the splicing of the 26 bp fragment from the XBP1 mRNA. Most of the human samples tested were found to contain only the unspliced form of XBP1. However, one mutant tissue displayed a faint but distinct band identical to the spliced form of XBP1. Surprisingly, all of the studied rat samples contained both spliced and unspliced forms of XBP1 mRNA, although with the relative enrichment of the spliced fragment in the heterozygous and homozygous animals.

In conclusion, the activation of the transcriptional arm of UPR, as judged by the appearance of the spliced (active) form of XBP1 and processed ATF6 transcription factors is suggested to contribute to the overexpression of ERp29 and other ER chaperones in the examined ERSDs.

5 DISCUSSION

5.1 Genomic organization, expression and tissue distribution of ERp29

Such characteristics of the 5′ flank as CpG island, the absence of TATA-box, multiple transcription start sites in combination with Sp1-dependent basal transcription and also ubiquitous gene expression indicate that ERp29 belongs to the group of constitutively expressed housekeeping genes. However, even though generally considered to have low expression fluctuations, the expression of a number of housekeeping genes varies significantly in different tissues (188,189), which apparently may be caused by quite variable demand for the basic cellular functions in different cell types.

ERp29 has a low expression pattern in tissues with a low secretory activity and higher expression in specialized secretory tissues. Interestingly, in ERp29-enriched tissues ERp29 is approximately equimolar with PDI and BiP suggesting that ERp29 might participate in stoichiometric protein processing events with principal secretory products (190). At the same time, low expression of ERp29 in testis and ovary, also active secretory tissues, which however, export mostly low molecular weight molecules such as testosterone and estrogen, implies that ERp29 might be involved primarily in the processing of protein substrates.

Although ERp29 levels are not particularly high in the brain, recent proteomic and immunochemical studies found that it is enriched in the cerebellum especially in Purkinje neurons (191,192). The cerebellum conventionally is not regarded as a neurosecretory tissue and ERp29 correlates poorly with classical markers of neurosecretion, but strongly with a variety of major membrane proteins. Thus, these authors hypothesize that ERp29 is involved primarily in the production of endomembrane proteins rather than proteins destined for the cell exterior (191).

5.2 Evolutionary expression

ERp29 has an interesting pattern of evolutionary expression. ERp29-like genes are absent in prokaryotic organisms and monocellular eukaryotes and first emerge in flatworms (193). Screening of the nearly complete genomes of metazoans identified ERp29s in arthropods (fruit fly, mosquito, honeybee) and mammals. However, ERp29 is absent from the *C. elegans* (nematode) and fish genomes. This is quite surprising as mammals and fish belong to the same taxonomic unit. Nevertheless, an interesting gene encoding a protein with unknown function with a KEEL ER-retrieval signal and an ER-targeting N-terminal signal peptide has been found in the zebrafish genome (193). It is clustered with human PDI on the dendrogram apparently due to the presence of four thioredoxin-like domains, of which the fold recognition methods predict characteristic PDI-like ***a*** and ***b*** domains. Interestingly, this protein contains no active redox motifs, which implicitly suggest their intermediate position between the redox-active PDIs and ERp29.

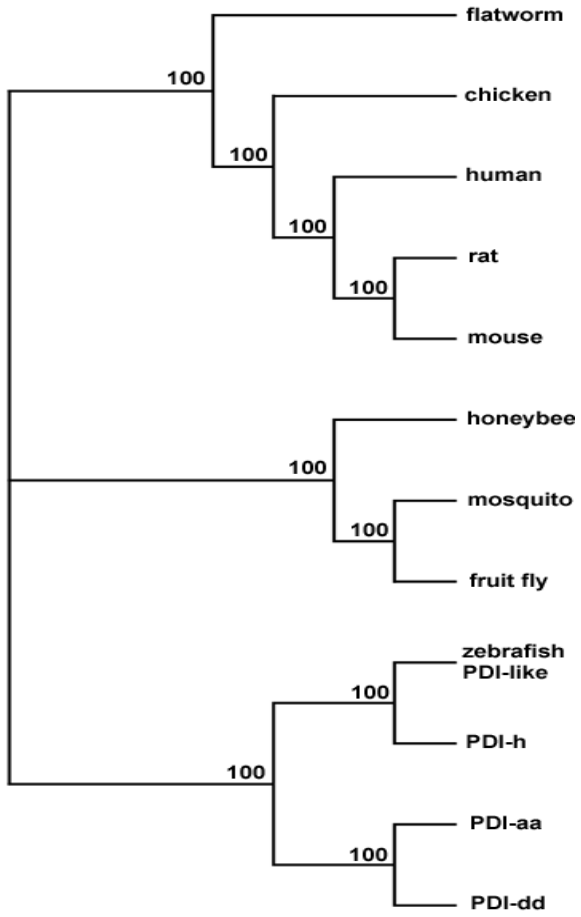

Figure 4. Phylogenetic tree of ERp29s and related PDIs. Phylogenies were inferred by distance matrix analysis using the PHYLIP program. The values supporting each node are derived from 100 resamplings. The image based on the PHYLIP output file was developed by the TreeView software. Adapted from (193)

Phylogenetic tree suggests the common origin of the ERp29 family and a group of PDIs from various organisms (fig. 4). Several P5-like PDIs (65) from plants (represented by the protein from alfalfa) and protists (represented by the amoebal PDI), have similar to ERp29s domain structure with one or two thioredoxin domains and a unique C-terminal domain, which positions them closer to ERp29 as compared to the mammalian PDIs.

In summary, phylogenetic analysis argues in favor of the hypothetical secretory role of ERp29 assuming that the protein export function is most extensively developed in multicellular organisms. It is conceivable that the split of ERp29 from PDI occurred due to the further differentiation of the secretory apparatus and a rising demand for a novel, more specialized functions, such as, for instance, more target- oriented assistance in folding, protection from premature associations and escorting of the proteins to the sites of their final destinations.

5.3 Functional activity of ERp29

The widespread tissue distribution indicates that ERp29 has a functional role of general utility rather than a specialist one used by a minority of cells. However, the physiological function and molecular mechanisms of action of ERp29 until now remained obscure, although most of the studies predicted a role connected with protein maturation and stress defense (167,169,173). Our study carried out in the thyroid epithelial cells provides the most up-to-date functional characterization of ERp29.

A high level of ERp29 gene expression in thyrocytes and its transcriptional regulation by thyroid-stimulating hormone (179) suggested that ERp29 can be implicated in the maturation and/or secretion of Tg. Indeed, co-immunoprecipitation analysis and sucrose density gradient fractionation of the cross-linked thyrocytes indicate the presence of ERp29 in the large heterogeneous Tg folding complexes, containing

additionally a set of folding assistants, such as BiP, GRP94, PDI and calreticulin (194,195).
However, it is not likely that the interaction of ERp29 in these complexes has a true chaperone-like nature as recent work (170) as well as our unpublished data showed that ERp29 lacks the classical chaperone activity. Moreover, the narrow substrate specificity, ATP-independent association with the substrate proteins and association with both denatured and native Tg in *in vitro* binding assay suggest that ERp29 has a unique mode of interaction and may contribute to the final stages of Tg exit from the ER. To some extent such behavior resembles dedicated escort chaperones, such as RAP (196) and the ERp29 analog in *Drosophila,* Wind (197) that remain attached to their substrates even after the completion of the folding and further escort them to the point of destination.

Despite the presence of the ER retrieval signal, substantial amounts of ERp29 have been detected in the medium of the thyroid epithelial cells. Moreover, hormone- and BFA-regulated export of ERp29 and co-localization with Tg in the putative transport structures indicate that ERp29 proceeds via the secretory pathway in a manner typical for secretory proteins, which implicitly suggests a protein escort function for ERp29.
How does ERp29 escape recycling and become secreted? One speculation is that the less conserved variant of the ERp29's retrieval signal (KEEL instead of KDEL) has a weaker retrieval pressure and allows easier exit of ERp29 in association with its substrates. Similarly, another ER escort, RAP has an even less effective signal, RNEL (85). Another explanation is that the C-terminal recognition sequence of ERp29 is shielded by the association of substrate protein (198).

To understand how essential ERp29 is for the secretion of Tg we studied the effect of the manipulated level of ERp29 in the thyroid cells on the rate of Tg export. Overexpression of ERp29 significantly increases the secretion of Tg whereas inhibition of ERp29 by siRNA and dominant-negative mutant forms of ERp29 down-regulate Tg transport. Demonstration of the ERp29-dependent secretion of Tg is the first direct evidence of the essential secretory role of ERp29.

A number of proteomic studies identified ERp29 in a wide range of tissues and organs. Although not providing concrete insights into the physiological role of ERp29, they nevertheless contribute to our knowledge of ERp29 tissue distribution, stress involvement, etc. Thus, induction of ERp29 was demonstrated in the differentiating B cells along with other ER chaperones, foldases and secretion factors, which was described as an evidence of the involvement of UPR in the developmental processes (199). These data in complex with another study where ERp29 was observed in the large ER-localized multiprotein complex that is comprised of the molecular chaperones and unassembled, incompletely folded immunoglobulin heavy chains (52) suggest the importance of ERp29 for the differentiation of the B lymphocytes into antibody-secreting plasma cells.
Another proteomic analysis identified ERp29 in the milk fat globule (MFG) produced during lactation by mammary gland (200). ERp29 was localized in both constituents of MFG, the cytoplasmic lipid droplet (CLD) and MFG membranes (MFGM) that surround CLD. The authors concluded that ERp29 may be instrumental in the lipid secretion from the mammary epithelial cells. However, as MFGM was found to be

enriched also by soluble milk proteins including casein, one can speculate that the presence of ERp29 in MFG might be explained by its putative secretory protein escort function. This suggestion is supported also by the low expression level in the lipid-oriented cells and preferential localization in the rough ER (192).
Finally, several proteomic studies have reported altered ERp29 expression upon various stress conditions. ERp29 has been shown to be induced in the mouse macrophages, treated with anthrax lethal toxin (201), in the liver of rats treated with hepatotoxic doses of bromobenzene (202) and in rat intestinal epithelial IEC-6 cells upon gamma-ray irradiation (203). Besides that, ERp29 has been shown to be induced in several tumor cell lines (204).
Taken together these studies implicitly endorse the role as secretory assistant for ERp29 and its involvement in the cell defense against various stress situations thus supporting the hypothetical biological role of ERp29 suggested in our study.

5.4 Substrate binding sites of ERp29

What is the molecular basis of the putative ERp29/Tg interactions? NMR spectroscopy has previously determined the structure of ERp29 domains and predicted several potential peptide-binding sites (171).
High degree of mobility of the interdomain linker suggests an important role for this region. Indeed, all deletion mutants from this fragment were negatively affecting the secretion of Tg. Such an effect may have different explanations.
It can be argued that the deletions (including glycines) hinder the relative dynamics of two domains that might be important for the ERp29-substrate interactions. Another explanation is that deletions in the linker might indirectly affect the dimer assembly. Finally, it cannot be ruled out that the protein binding site is located either within the linker or in its immediate vicinity.
The interdomain linker contains also a solvent-exposed unique cysteine (Cys157), conserved between mammals and arthropods (fig. 5). Although it is probably not involved in the catalytic reactions employing redox chemistry it was interesting to investigate if this residue is of any functional relevance for ERp29. Indeed, replacement of Cys157 with alanine or serine resulted in the drastic reduction of the Tg secretion. Examination of the local structure around Cys157 in the C-terminal domain revealed that in the majority of 20 NMR structures the thiol group of Cys157 is involved in the intricate net of van der Waals/hydrogen bond interactions between the residues from helix 5. Removal of the thiol (Cys157Ala) or even replacement for the smaller and more hydrophilic hydroxyl group (Cys157Ser) might disrupt this interaction and allow helix 5 to move away. Hence, it seems that Cys157 plays a major role in the structural integrity of ERp29 and more specifically in the stabilization of its C-terminal domain.
However, the question remains as to whether this residue and/or the whole linker are directly involved in the interaction with the ERp29 protein substrates. Recently, a study was published where the authors investigating the biophysical properties of the Cys157Ser ERp29 mutant revealed the reduced surface hydrophobicity and increased proteolytic lability of the mutant protein thus suggesting its important structural role (205). Noting the high hydrophobicity of the linker, the authors suggested that it could represent a potential site of interaction with the unfolded substrates.

Mutations of two residues from the two different loci that were predicted by NMR study to be important for oligomerization of ERp29 showed no effect on the ERp29 secretory function. This might be interpreted as either the oligomerization is not

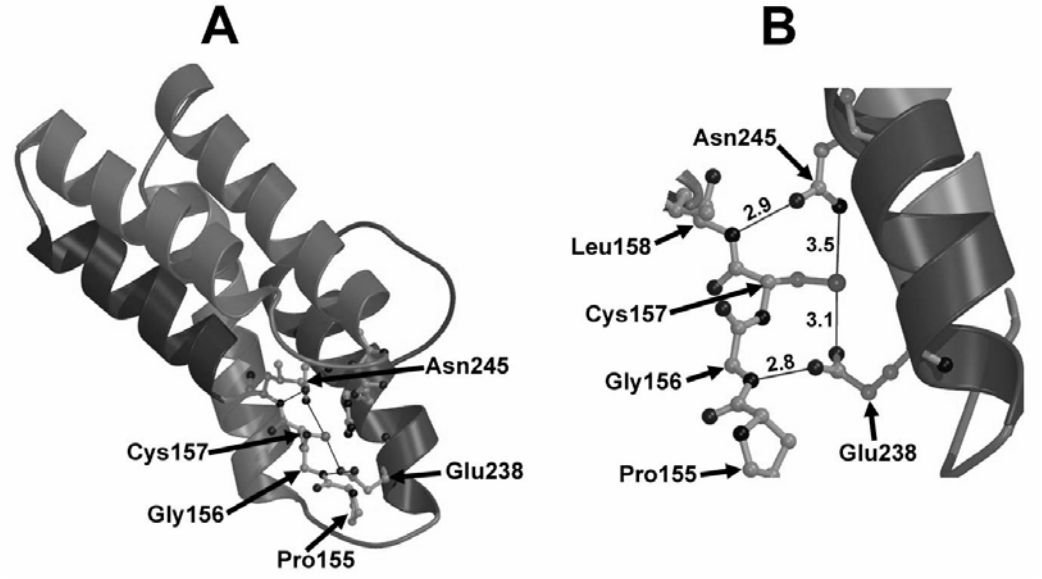

Figure 5. Cys157 is important for the structural integrity of ERp29.

A. Ribbon diagram of the C-terminal domain of ERp29 (PDB accession G7DE). **B.** Enlarged view of the N-terminus of the domain including Cys157 and surrounding amino acids shown as ball and stick model. Labels point to Cα± atoms of the corresponding amino acids. Indicated are distances (in Å…) between Cys157 and Glu238 and Asn245. Molecular graphics were produced by MOLSCRIPT (35).

important for the function of ERp29 or the oligomerization of recombinant His-tagged ERp29 is an artefactual event.

Further analysis of potential peptide binding sites on the ERp29 surface is based on the reported structure of Wind, the ERp29 ortholog from *Drosophila* (206,207). We found that the residues of the putative substrate binding site, identified in the thioredoxin domain of Wind are conserved between Wind and ERp29 and are sufficiently well superimposed structurally. Mutation of amino acids from the corresponding region of ERp29 (Tyr64 and Gln70) led to the pronounced drop in the export of Tg. Analysis of the electrostatic surface of the N-terminal domain revealed an uncharged cleft between these residues (approximately 13 Å wide) that might accommodate proteins of sufficiently large size. Moreover, further extrapolations of Wind data to ERp29 allows the speculation that the aromatic nature of Tyr and not the polar interactions are important for the putative peptide binding in this region.

Another putative functional locus of Wind, a cluster of conserved residues, is located in the C-terminal domain (206). It indicates a more functional role of this domain rather than being merely responsible for the retention and/or solubility of the full-length protein. As the chimera composed of the N-terminal domain of Wind and the C-terminal domain of ERp29 is able to functionally replace Wind it might be speculated that these data could be of functional relevance also for ERp29. Although further mutational analysis is needed to confirm this, one exciting hypothesis is that the conformational changes induced by the mutations/deletions in the interdomain linker

might affect the functional cluster of the C-terminal domain, thus impairing the function of ERp29.

5.5 ERp29 is a potential target of the Unfolded Protein Response

Early experiments demonstrated stress-induced upregulation of ERp29 simultaneously with a number of ER chaperones that have been shown to be regulated by UPR (167). Although the promoter region of ERp29 lacks the ERSE motif characteristic for UPR target genes and the induction of ERp29 was not as strong as compared to the major ER chaperones it was hypothesized that ERp29 is a potential target of UPR. Similar weak stress-induced upregulation has been shown for PDI, which also lacks ERSE consensus sequence (208).
In this study we demonstrated induction of ERp29 in the hormone-stimulated thyroid cells and in the typical ERSDs, human and rat variant of thyroid diseases. The elevation of ER chaperones under these conditions was known before (209,210), however little is known about the mechanisms and kinetics of UPR development. To clarify the potential role of UPR in the upregulation of ERp29 and other ER chaperones we examined the activation of the UPR sensors.

Despite the observed activation of all studied UPR pathways, it seems unlikely that ATF6 or PERK pathways are involved in the regulation of ERp29. As mentioned above, the 5´-flanking region of ERp29 lacks ERSE motif for binding to ATF6 and downstream transcription factor of PERK, ATF4. Additionally, we have shown that overexpression of ATF6 does not affect the ERp29 mRNA level and the expression of luciferase reporter carrying ERp29 promoter fragments (our unpublished data).
The induction of the ATF6 pathway upon ER stress is usually accompanied by activation of IRE1 and subsequent splicing of XBP1 mRNA.
Detection of the spliced form of XBP1 in mutant human tissue and in all of the studied rat samples with the relative enrichment in the heterozygous and homozygous animals indicates the involvement of this pathway in the development of UPR and induction of molecular chaperones.
Although PCR analysis failed to detect the presence of spliced form of XBP1 in the stimulated thyroid cells induction of the unspliced form of XBP1 implicitly points to the involvement of this pathway in the development of UPR. Interestingly, sequencing of the *PstI*-resistant product migrating as an unspliced fragment revealed a mixture of spliced and unspliced forms of XBP1 indicating certain level of IRE1 activation. The existence of such hybrid forms of XBP1 mRNA caused by a moderate ER stress was also reported by others (211).
The unlikely involvement of ATF6 and PERK pathways in the regulation of ERp29 and the presence of potential XBP1-binding sites in the 5´-flanking region of ERp29 prompted us to speculate that ERp29 expression is controlled primarily via the XBP1/IRE1 pathway. Indeed, overexpression of the spliced form of XBP1 enhances the expression of luciferase reporter carrying ERp29 promoter fragments and notably induces ERp29 mRNA level (our unpublished data). Additionally, we have shown that while in the wild type mouse embryonic cells (MEF) ERp29 mRNA is induced by

tunicamycin or thapsigargin, in the XBP1-/- and IRE1-/- MEFs such upregulation is abolished (unpublished observations).

Interestingly, another ER protein, EDEM, was also found to be regulated in a similar manner (97). The kinetics of ERp29 induction coincides with that of the EDEM and is substantially delayed as compared with the ATF6 target protein, BiP. This would indicate that the XBP1-dependent subset of ER proteins including ERp29 is needed on the later stages of UPR in order to facilitate the traffic of increased amounts of correctly folded proteins and/or to direct misfolded species to the proteasomal machinery.

Our data are also in agreement with another study where XBP1 has been shown to be important for the regulation of ER chaperones and secretion factors including ERp29 during the differentiation of B lymphocytes into the antibody-secreting plasma cells (199).

6 CONCLUSIONS

- Ubiquitous expression in combination with structural and functional characteristics of the ERp29 gene indicates that ERp29 belongs to the group of constitutively expressed housekeeping genes. A relatively high expression in specialized secretory tissues suggests the function connected with the processing of secretory proteins from the ER.

- Expression in multicellular organisms and absence in monocellular life forms also speaks in favor of a hypothetical secretory role of ERp29 assuming that the protein export function is most extensively developed in the multicellular organisms.

- Involvement of ERp29 in the heterogeneous thyroglobulin folding complexes in the thyroid epithelial cells concomitantly with other ER chaperones and preferential interaction with misfolded thyroglobulin suggests chaperone-like properties for ERp29.

- An increased secretion of ERp29 in the cells stimulated to produce large amounts of Tg and ERp29-dependent regulation of Tg traffic suggest a more precise definition of ERp29 function as an ER secretion factor/escort chaperone.

- Mutational analysis suggests an important role of the interdomain linker of ERp29 for biological function. The unique cysteine, located in this region seems to be essential for the structural integrity of the C-terminal domain, and hence, for the whole protein. An additional putative substrate-binding site is located in the N-terminal thioredoxin domain representing an uncharged cleft that may accommodate proteins of sufficiently large size.

- Induction of ERp29 concomitantly with the major ER chaperones in the cells undergoing physiological UPR or ER stress caused by the accumulation of transport-incompetent mutant secretory proteins suggest an involvement of UPR in the regulation of ERp29.

7 ACKNOWLEDGEMENTS

This thesis is a result of research carried out at the Division of Molecular Toxicology, Institute of Environmental Medicine, Karolinska Institute.

I would like to express my sincere gratitude and thanks to all those who contributed to this work:

Dr. Souren Mkrtchian, my supervisor, for giving me an opportunity to work in a small "ERp29 group" and believing in me. It was an honour for me to work with a real and devoted scientist. I appreciate very much your patience, constant support, discussions and lots of advice concerning science and other aspects of life.

Prof. Magnus Ingelman-Sundberg, my co-supervisor, for letting me join the group, for financial support and for providing a stimulating and at the same time friendly working environment.

Mikhail Barishev, for teaching me the molecular biology methods, for having patience to answer all my questions and for the interesting conversations during the lunches.

Anatoly Sharipo, for fruitful collaboration and interesting ideas.

Daniel, for thorough reading of my thesis and for useful comments.

My roommates, for the great time I had with you:
Cristina, for your optimism and kindness, for the scientific discussions, interesting chats and funny jokes. Sussi and Tove, for bringing some Swedish atmosphere into our office room. Eleni, for being not only a good office mate but also a good neighbour in Lappis.

All the other colleagues and friends in the lab:

present
Alvin, Amm, Angelica, Anita, Bengt, Diana, Erika, Inger, Jessica, Jue, Margareta, Maria K, Sarah, Sun, Susanne, Ylva, Yvonne, Åsa

past
Agneta, Anna W, Etienne, Fang, Katarina, Magdalena, Maria B, Marià, MariCarmen, Matteo, Mia, Micke, Natassa, Niclas, Oliver, Pelin, Roman, Sandra, Shinichi, Veronica, Vita, Yoon

our students
Anna Ö, Jesper, Karina, Pia

for the fantastic working environment and for the friendly atmosphere in the lab. I will never forget the wonderful time I had with you inside and outside the lab. And such

activities as bowling, curling, “innebandy” (especially, if your team is a winner)… they are just unforgettable.

Prof. Armen Trchounian and prof. Anna Boyajian, for being my first teachers and for helping me make my first steps in science.

My friends and colleagues in Armenia, for remembering me. I miss all of you very much and I am always happy to see you during my short visits. Special thanks to Vitja and Tatul, for being true friends whom I can always rely on.

My new friends in Sweden, for helping me to survive in a new country, “far-far away”, especially during the first months.

My parents and my sister, for the endless love and for the constant support in realization of my plans. My grandfather, who will always be unique for me and whom I will always admire.

Finally, my wife Katja, for the true love and for bringing more meaning to my life.

8 REFERENCES

1. Zarembinski, T. I., Hung, L. W., Mueller-Dieckmann, H. J., Kim, K. K., Yokota, H., Kim, R., and Kim, S. H. (1998) *Proc Natl Acad Sci U S A* **95**, 15189-15193
2. Palade, G. (1975) *Science* **189**, 347-358
3. Ellgaard, L., and Helenius, A. (2003) *Nat Rev Mol Cell Biol* **4**, 181-191
4. Bonifacino, J. S., and Glick, B. S. (2004) *Cell* **116**, 153-166
5. Kalies, K. U., and Hartmann, E. (1998) *Eur J Biochem* **254**, 1-5
6. Rapoport, T. A., Jungnickel, B., and Kutay, U. (1996) *Annu Rev Biochem* **65**, 271-303
7. Wilkinson, B. M., Regnacq, M., and Stirling, C. J. (1997) *J Membr Biol* **155**, 189-197
8. von Heijne, G. (1990) *J Membr Biol* **115**, 195-201
9. Blobel, G., and Dobberstein, B. (1975) *J Cell Biol* **67**, 852-862
10. Blobel, G., and Dobberstein, B. (1975) *J Cell Biol* **67**, 835-851
11. Walter, P., Ibrahimi, I., and Blobel, G. (1981) *J Cell Biol* **91**, 545-550
12. Walter, P., and Lingappa, V. R. (1986) *Annu Rev Cell Biol* **2**, 499-516
13. Hamman, B. D., Chen, J. C., Johnson, E. E., and Johnson, A. E. (1997) *Cell* **89**, 535-544
14. High, S. (1995) *Prog Biophys Mol Biol* **63**, 233-250
15. Martoglio, B., and Dobberstein, B. (1996) *Trends Cell Biol* **6**, 142-147
16. Gorlich, D., Hartmann, E., Prehn, S., and Rapoport, T. A. (1992) *Nature* **357**, 47-52
17. Gorlich, D., and Rapoport, T. A. (1993) *Cell* **75**, 615-630
18. Johnson, A. E., and van Waes, M. A. (1999) *Annu Rev Cell Dev Biol* **15**, 799-842
19. Schnell, D. J., and Hebert, D. N. (2003) *Cell* **112**, 491-505
20. Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G., and Frank, J. (1997) *Science* **278**, 2123-2126
21. Stirling, C. J. (1999) *Microbiology* **145 (Pt 5)**, 991-998
22. Chirico, W. J., Waters, M. G., and Blobel, G. (1988) *Nature* **332**, 805-810
23. Deshaies, R. J., Koch, B. D., Werner-Washburne, M., Craig, E. A., and Schekman, R. (1988) *Nature* **332**, 800-805
24. Baumann, O., and Walz, B. (2001) *Int Rev Cytol* **205**, 149-214
25. Vertel, B. M., Walters, L. M., and Mills, D. (1992) *Semin Cell Biol* **3**, 325-341
26. Terasaki, M., and Jaffe, L. A. (1991) *J Cell Biol* **114**, 929-940
27. Sitia, R., and Meldolesi, J. (1992) *Mol Biol Cell* **3**, 1067-1072
28. Bergmann, J. E., and Fusco, P. J. (1990) *J Cell Biol* **110**, 625-635
29. Anfinsen, C. B. (1967) *Harvey Lect* **61**, 95-116
30. Anfinsen, C. B. (1973) *Science* **181**, 223-230
31. Koch, G. L. (1987) *J Cell Sci* **87 (Pt 4)**, 491-492
32. Hwang, C., Sinskey, A. J., and Lodish, H. F. (1992) *Science* **257**, 1496-1502
33. Sambrook, J. F. (1990) *Cell* **61**, 197-199
34. Clairmont, C. A., De Maio, A., and Hirschberg, C. B. (1992) *J Biol Chem* **267**, 3983-3990
35. Kornfeld, R., and Kornfeld, S. (1985) *Annu Rev Biochem* **54**, 631-664
36. Ellis, R. J., van der Vies, S. M., and Hemmingsen, S. M. (1989) *Biochem Soc Symp* **55**, 145-153
37. Ellgaard, L., Molinari, M., and Helenius, A. (1999) *Science* **286**, 1882-1888
38. Haas, I. G., and Wabl, M. (1983) *Nature* **306**, 387-389
39. Munro, S., and Pelham, H. R. (1986) *Cell* **46**, 291-300
40. Flynn, G. C., Chappell, T. G., and Rothman, J. E. (1989) *Science* **245**, 385-390
41. Blond-Elguindi, S., Fourie, A. M., Sambrook, J. F., and Gething, M. J. (1993) *J Biol Chem* **268**, 12730-12735
42. Blond-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F., and Gething, M. J. (1993) *Cell* **75**, 717-728

43. Gething, M. J., and Sambrook, J. (1992) *Nature* **355**, 33-45
44. Vogel, J. P., Misra, L. M., and Rose, M. D. (1990) *J Cell Biol* **110**, 1885-1895
45. Sanders, S. L., Whitfield, K. M., Vogel, J. P., Rose, M. D., and Schekman, R. W. (1992) *Cell* **69**, 353-365
46. Nicchitta, C. V., and Blobel, G. (1993) *Cell* **73**, 989-998
47. Glick, B. S. (1995) *Cell* **80**, 11-14
48. Matlack, K. E., Misselwitz, B., Plath, K., and Rapoport, T. A. (1999) *Cell* **97**, 553-564
49. Simon, S. M., Peskin, C. S., and Oster, G. F. (1992) *Proc Natl Acad Sci U S A* **89**, 3770-3774
50. Shiu, R. P., Pouyssegur, J., and Pastan, I. (1977) *Proc Natl Acad Sci U S A* **74**, 3840-3844
51. Koch, G., Smith, M., Macer, D., Webster, P., and Mortara, R. (1986) *J Cell Sci* **86**, 217-232
52. Meunier, L., Usherwood, Y. K., Chung, K. T., and Hendershot, L. M. (2002) *Mol Biol Cell* **13**, 4456-4469
53. Ma, Y., and Hendershot, L. M. (2004) *J Chem Neuroanat* **28**, 51-65
54. Li, Z., Dai, J., Zheng, H., Liu, B., and Caudill, M. (2002) *Front Biosci* **7**, d731-751
55. Ellgaard, L., and Frickel, E. M. (2003) *Cell Biochem Biophys* **39**, 223-247
56. Hammond, C., and Helenius, A. (1994) *J Cell Biol* **126**, 41-52
57. Ware, F. E., Vassilakos, A., Peterson, P. A., Jackson, M. R., Lehrman, M. A., and Williams, D. B. (1995) *J Biol Chem* **270**, 4697-4704
58. Cannon, K. S., and Helenius, A. (1999) *J Biol Chem* **274**, 7537-7544
59. High, S., Lecomte, F. J., Russell, S. J., Abell, B. M., and Oliver, J. D. (2000) *FEBS Lett* **476**, 38-41
60. Oliver, J. D., van der Wal, F. J., Bulleid, N. J., and High, S. (1997) *Science* **275**, 86-88
61. Benham, A. (2003) *Antioxid Redox Signal* **5**, 355-357
62. LaMantia, M. L., and Lennarz, W. J. (1993) *Cell* **74**, 899-908
63. Laboissiere, M. C., Sturley, S. L., and Raines, R. T. (1995) *J Biol Chem* **270**, 28006-28009
64. Gilbert, H. F. (1998) *Methods Enzymol* **290**, 26-50
65. Ferrari, D. M., and Soling, H. D. (1999) *Biochem J* **339 (Pt 1)**, 1-10
66. Noiva, R. (1999) *Semin Cell Dev Biol* **10**, 481-493
67. Anelli, T., Alessio, M., Mezghrani, A., Simmen, T., Talamo, F., Bachi, A., and Sitia, R. (2002) *Embo J* **21**, 835-844
68. Cabibbo, A., Pagani, M., Fabbri, M., Rocchi, M., Farmery, M. R., Bulleid, N. J., and Sitia, R. (2000) *J Biol Chem* **275**, 4827-4833
69. Freedman, R. B., Hirst, T. R., and Tuite, M. F. (1994) *Trends Biochem Sci* **19**, 331-336
70. Wang, C. C., and Tsou, C. L. (1993) *Faseb J* **7**, 1515-1517
71. Wang, C. C. (1998) *Ann N Y Acad Sci* **864**, 9-13
72. Quan, H., Fan, G., and Wang, C. C. (1995) *J Biol Chem* **270**, 17078-17080
73. Sideraki, V., and Gilbert, H. F. (2000) *Biochemistry* **39**, 1180-1188
74. van Lith, M., Hartigan, N., Hatch, J., and Benham, A. M. (2005) *J Biol Chem* **280**, 1376-1383
75. Freskgard, P. O., Bergenhem, N., Jonsson, B. H., Svensson, M., and Carlsson, U. (1992) *Science* **258**, 466-468
76. Fransson, C., Freskgard, P. O., Herbertsson, H., Johansson, A., Jonasson, P., Martensson, L. G., Svensson, M., Jonsson, B. H., and Carlsson, U. (1992) *FEBS Lett* **296**, 90-94
77. Lodish, H. F., and Kong, N. (1991) *J Biol Chem* **266**, 14835-14838
78. Fruman, D. A., Burakoff, S. J., and Bierer, B. E. (1994) *Faseb J* **8**, 391-400
79. Nakai, A., Satoh, M., Hirayoshi, K., and Nagata, K. (1992) *J Cell Biol* **117**, 903-914
80. Smith, T., Ferreira, L. R., Hebert, C., Norris, K., and Sauk, J. J. (1995) *J Biol Chem* **270**, 18323-18328
81. Satoh, M., Hirayoshi, K., Yokota, S., Hosokawa, N., and Nagata, K. (1996) *J Cell Biol* **133**, 469-483

82. Wetterau, J. R., Aggerbeck, L. P., Bouma, M. E., Eisenberg, C., Munck, A., Hermier, M., Schmitz, J., Gay, G., Rader, D. J., and Gregg, R. E. (1992) *Science* **258**, 999-1001
83. Willnow, T. E., Armstrong, S. A., Hammer, R. E., and Herz, J. (1995) *Proc Natl Acad Sci U S A* **92**, 4537-4541
84. Williams, S. E., Ashcom, J. D., Argraves, W. S., and Strickland, D. K. (1992) *J Biol Chem* **267**, 9035-9040
85. Bu, G., and Marzolo, M. P. (2000) *Trends Cardiovasc Med* **10**, 148-155
86. Bu, G., Rennke, S., and Geuze, H. J. (1997) *J Cell Sci* **110 (Pt 1)**, 65-73
87. Lisi, S., Chiovato, L., Pinchera, A., Marcocci, C., Menconi, F., Morabito, E., Altea, M. A., McCluskey, R. T., and Marino, M. (2003) *J Endocrinol Invest* **26**, 1105-1110
88. Braks, J. A., and Martens, G. J. (1994) *Cell* **78**, 263-273
89. Mbikay, M., Seidah, N. G., and Chretien, M. (2001) *Biochem J* **357**, 329-342
90. Braks, J. A., Van Horssen, A. M., and Martens, G. J. (1996) *Eur J Biochem* **238**, 505-510
91. Zhu, X., and Lindberg, I. (1995) *J Cell Biol* **129**, 1641-1650
92. Hammond, C., and Helenius, A. (1995) *Curr Opin Cell Biol* **7**, 523-529
93. Kopito, R. R. (1997) *Cell* **88**, 427-430
94. Rutkowski, D. T., and Kaufman, R. J. (2004) *Trends Cell Biol* **14**, 20-28
95. Kaufman, R. J. (1999) *Genes Dev* **13**, 1211-1233
96. Patil, C., and Walter, P. (2001) *Curr Opin Cell Biol* **13**, 349-355
97. Yoshida, H., Matsui, T., Hosokawa, N., Kaufman, R. J., Nagata, K., and Mori, K. (2003) *Dev Cell* **4**, 265-271
98. Resendez, E., Jr., Attenello, J. W., Grafsky, A., Chang, C. S., and Lee, A. S. (1985) *Mol Cell Biol* **5**, 1212-1219
99. Liu, E. S., Ou, J. H., and Lee, A. S. (1992) *J Biol Chem* **267**, 7128-7133
100. Kim, P. S., and Arvan, P. (1998) *Endocr Rev* **19**, 173-202
101. Rutishauser, J., and Spiess, M. (2002) *Swiss Med Wkly* **132**, 211-222
102. Shen, J., Chen, X., Hendershot, L., and Prywes, R. (2002) *Dev Cell* **3**, 99-111
103. Shen, J., Snapp, E. L., Lippincott-Schwartz, J., and Prywes, R. (2005) *Mol Cell Biol* **25**, 921-932
104. Liu, C. Y., Schroder, M., and Kaufman, R. J. (2000) *J Biol Chem* **275**, 24881-24885
105. Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P., and Ron, D. (2000) *Nat Cell Biol* **2**, 326-332
106. Ye, J., Rawson, R. B., Komuro, R., Chen, X., Dave, U. P., Prywes, R., Brown, M. S., and Goldstein, J. L. (2000) *Mol Cell* **6**, 1355-1364
107. Yoshida, H., Haze, K., Yanagi, H., Yura, T., and Mori, K. (1998) *J Biol Chem* **273**, 33741-33749
108. Haze, K., Yoshida, H., Yanagi, H., Yura, T., and Mori, K. (1999) *Mol Biol Cell* **10**, 3787-3799
109. Kadowaki, H., Nishitoh, H., and Ichijo, H. (2004) *J Chem Neuroanat* **28**, 93-100
110. Tirasophon, W., Welihinda, A. A., and Kaufman, R. J. (1998) *Genes Dev* **12**, 1812-1824
111. Shen, X., Ellis, R. E., Lee, K., Liu, C. Y., Yang, K., Solomon, A., Yoshida, H., Morimoto, R., Kurnit, D. M., Mori, K., and Kaufman, R. J. (2001) *Cell* **107**, 893-903
112. Yoshida, H., Matsui, T., Yamamoto, A., Okada, T., and Mori, K. (2001) *Cell* **107**, 881-891
113. Calfon, M., Zeng, H., Urano, F., Till, J. H., Hubbard, S. R., Harding, H. P., Clark, S. G., and Ron, D. (2002) *Nature* **415**, 92-96
114. Harding, H. P., Zhang, Y., and Ron, D. (1999) *Nature* **397**, 271-274
115. Harding, H. P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., and Ron, D. (2000) *Mol Cell* **6**, 1099-1108
116. Scheuner, D., Song, B., McEwen, E., Liu, C., Laybutt, R., Gillespie, P., Saunders, T., Bonner-Weir, S., and Kaufman, R. J. (2001) *Mol Cell* **7**, 1165-1176

117. Gass, J. N., Gifford, N. M., and Brewer, J. W. (2002) *J Biol Chem* **277**, 49047-49054
118. Harding, H. P., Zeng, H., Zhang, Y., Jungries, R., Chung, P., Plesken, H., Sabatini, D. D., and Ron, D. (2001) *Mol Cell* **7**, 1153-1163
119. Liu, Y., Choudhury, P., Cabral, C. M., and Sifers, R. N. (1999) *J Biol Chem* **274**, 5861-5867
120. Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L. O., Herscovics, A., and Nagata, K. (2001) *EMBO Rep* **2**, 415-422
121. Tremblay, L. O., and Herscovics, A. (1999) *Glycobiology* **9**, 1073-1078
122. Hosokawa, N., Tremblay, L. O., You, Z., Herscovics, A., Wada, I., and Nagata, K. (2003) *J Biol Chem* **278**, 26287-26294
123. Plemper, R. K., Bohmler, S., Bordallo, J., Sommer, T., and Wolf, D. H. (1997) *Nature* **388**, 891-895
124. Bannykh, S. I., Rowe, T., and Balch, W. E. (1996) *J Cell Biol* **135**, 19-35
125. Bannykh, S. I., and Balch, W. E. (1997) *J Cell Biol* **138**, 1-4
126. Schekman, R., and Orci, L. (1996) *Science* **271**, 1526-1533
127. Barlowe, C. (1998) *Biochim Biophys Acta* **1404**, 67-76
128. Mizuno, M., and Singer, S. J. (1993) *Proc Natl Acad Sci U S A* **90**, 5732-5736
129. Balch, W. E., McCaffery, J. M., Plutner, H., and Farquhar, M. G. (1994) *Cell* **76**, 841-852
130. Kuehn, M. J., Herrmann, J. M., and Schekman, R. (1998) *Nature* **391**, 187-190
131. Hauri, H. P., Kappeler, F., Andersson, H., and Appenzeller, C. (2000) *J Cell Sci* **113 (Pt 4)**, 587-596
132. Dominguez, M., Dejgaard, K., Fullekrug, J., Dahan, S., Fazel, A., Paccaud, J. P., Thomas, D. Y., Bergeron, J. J., and Nilsson, T. (1998) *J Cell Biol* **140**, 751-765
133. Wieland, F. T., Gleason, M. L., Serafini, T. A., and Rothman, J. E. (1987) *Cell* **50**, 289-300
134. Schweizer, A., Fransen, J. A., Matter, K., Kreis, T. E., Ginsel, L., and Hauri, H. P. (1990) *Eur J Cell Biol* **53**, 185-196
135. Sollner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J. E. (1993) *Nature* **362**, 318-324
136. Rothman, J. E., and Warren, G. (1994) *Curr Biol* **4**, 220-233
137. Munro, S., and Pelham, H. R. (1987) *Cell* **48**, 899-907
138. Pelham, H. R. (1989) *Annu Rev Cell Biol* **5**, 1-23
139. Pelham, H. R. (1990) *Trends Biochem Sci* **15**, 483-486
140. Pelham, H. R., Hardwick, K. G., and Lewis, M. J. (1988) *Embo J* **7**, 1757-1762
141. Teasdale, R. D., and Jackson, M. R. (1996) *Annu Rev Cell Dev Biol* **12**, 27-54
142. Lewis, M. J., Sweet, D. J., and Pelham, H. R. (1990) *Cell* **61**, 1359-1363
143. Semenza, J. C., Hardwick, K. G., Dean, N., and Pelham, H. R. (1990) *Cell* **61**, 1349-1357
144. Cosson, P., and Letourneur, F. (1994) *Science* **263**, 1629-1631
145. Andersson, H., Kappeler, F., and Hauri, H. P. (1999) *J Biol Chem* **274**, 15080-15084
146. Hortsch, M., Griffiths, G., and Meyer, D. I. (1985) *Eur J Cell Biol* **38**, 271-279
147. Thyberg, J., and Moskalewski, S. (1985) *Exp Cell Res* **159**, 1-16
148. Ladinsky, M. S., Mastronarde, D. N., McIntosh, J. R., Howell, K. E., and Staehelin, L. A. (1999) *J Cell Biol* **144**, 1135-1149
149. Mellman, I., and Simons, K. (1992) *Cell* **68**, 829-840
150. Martinez-Menarguez, J. A., Prekeris, R., Oorschot, V. M., Scheller, R., Slot, J. W., Geuze, H. J., and Klumperman, J. (2001) *J Cell Biol* **155**, 1213-1224
151. Mironov, A. A., Beznoussenko, G. V., Nicoziani, P., Martella, O., Trucco, A., Kweon, H. S., Di Giandomenico, D., Polishchuk, R. S., Fusella, A., Lupetti, P., Berger, E. G., Geerts, W. J., Koster, A. J., Burger, K. N., and Luini, A. (2001) *J Cell Biol* **155**, 1225-1238
152. Robinson, M. S., and Bonifacino, J. S. (2001) *Curr Opin Cell Biol* **13**, 444-453
153. Traub, L. M., and Kornfeld, S. (1997) *Curr Opin Cell Biol* **9**, 527-533
154. Brown, D. A., Crise, B., and Rose, J. K. (1989) *Science* **245**, 1499-1501
155. Kornfeld, S., and Mellman, I. (1989) *Annu Rev Cell Biol* **5**, 483-525
156. Kornfeld, S. (1992) *Annu Rev Biochem* **61**, 307-330

157. Lisanti, M. P., Caras, I. W., Davitz, M. A., and Rodriguez-Boulan, E. (1989) *J Cell Biol* **109**, 2145-2156
158. Scheiffele, P., Peranen, J., and Simons, K. (1995) *Nature* **378**, 96-98
159. Callea, F., Brisigotti, M., Fabbretti, G., Bonino, F., and Desmet, V. J. (1992) *Liver* **12**, 357-362
160. Riordan, J. R., Rommens, J. M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J. L., and et al. (1989) *Science* **245**, 1066-1073
161. Cheng, S. H., Gregory, R. J., Marshall, J., Paul, S., Souza, D. W., White, G. A., O'Riordan, C. R., and Smith, A. E. (1990) *Cell* **63**, 827-834
162. Ward, C. L., Omura, S., and Kopito, R. R. (1995) *Cell* **83**, 121-127
163. Nichols, W. C., Seligsohn, U., Zivelin, A., Terry, V. H., Hertel, C. E., Wheatley, M. A., Moussalli, M. J., Hauri, H. P., Ciavarella, N., Kaufman, R. J., and Ginsburg, D. (1998) *Cell* **93**, 61-70
164. Sharp, D., Blinderman, L., Combs, K. A., Kienzle, B., Ricci, B., Wager-Smith, K., Gil, C. M., Turck, C. W., Bouma, M. E., Rader, D. J., and et al. (1993) *Nature* **365**, 65-69
165. Gregg, R. E., and Wetterau, J. R. (1994) *Curr Opin Lipidol* **5**, 81-86
166. Fang, C., Mkrtchian, S., and Ingelman-Sundberg, M. (1997) *Biotechniques* **23**, 52, 54, 56, 58
167. Mkrtchian, S., Fang, C., Hellman, U., and Ingelman-Sundberg, M. (1998) *Eur J Biochem* **251**, 304-313
168. Demmer, J., Zhou, C., and Hubbard, M. J. (1997) *FEBS Lett* **402**, 145-150
169. Ferrari, D. M., Nguyen Van, P., Kratzin, H. D., and Soling, H. D. (1998) *Eur J Biochem* **255**, 570-579
170. Hubbard, M. J., Mangum, J. E., and McHugh, N. J. (2004) *Biochem J* **383**, 589-597
171. Liepinsh, E., Baryshev, M., Sharipo, A., Ingelman-Sundberg, M., Otting, G., and Mkrtchian, S. (2001) *Structure (Camb)* **9**, 457-471
172. Mkrtchian, S., Baryshev, M., Matvijenko, O., Sharipo, A., Sandalova, T., Schneider, G., Ingelman-Sundberg, M., and Mkrtchiana, S. (1998) *FEBS Lett* **431**, 322-326
173. Hubbard, M. J., McHugh, N. J., and Carne, D. L. (2000) *Eur J Biochem* **267**, 1945-1957
174. Konsolaki, M., and Schupbach, T. (1998) *Genes Dev* **12**, 120-131
175. Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992) *Genomics* **13**, 1095-1107
176. Dynan, W. S., Sazer, S., Tjian, R., and Schimke, R. T. (1986) *Nature* **319**, 246-248
177. Pugh, B. F., and Tjian, R. (1991) *Genes Dev* **5**, 1935-1945
178. Purnell, B. A., and Gilmour, D. S. (1993) *Mol Cell Biol* **13**, 2593-2603
179. Kwon, O. Y., Park, S., Lee, W., You, K. H., Kim, H., and Shong, M. (2000) *FEBS Lett* **475**, 27-30
180. Gething, M. J. (1999) *Semin Cell Dev Biol* **10**, 465-472
181. Bruneau, N., Lombardo, D., Levy, E., and Bendayan, M. (2000) *Microsc Res Tech* **49**, 329-345
182. Xiao, G., Chung, T. F., Pyun, H. Y., Fine, R. E., and Johnson, R. J. (1999) *Brain Res Mol Brain Res* **72**, 121-128
183. Wiest, D. L., Bhandoola, A., Punt, J., Kreibich, G., McKean, D., and Singer, A. (1997) *Proc Natl Acad Sci U S A* **94**, 1884-1889
184. Mezghrani, A., Courageot, J., Mani, J. C., Pugniere, M., Bastiani, P., and Miquelis, R. (2000) *J Biol Chem* **275**, 1920-1929
185. Hishinuma, A., Furudate, S., Oh-Ishi, M., Nagakubo, N., Namatame, T., and Ieiri, T. (2000) *Endocrinology* **141**, 4050-4055
186. Hishinuma, A., Takamatsu, J., Ohyama, Y., Yokozawa, T., Kanno, Y., Kuma, K., Yoshida, S., Matsuura, N., and Ieiri, T. (1999) *J Clin Endocrinol Metab* **84**, 1438-1444
187. Kim, P. S., Ding, M., Menon, S., Jung, C. G., Cheng, J. M., Miyamoto, T., Li, B., Furudate, S., and Agui, T. (2000) *Mol Endocrinol* **14**, 1944-1953

188. Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A., and Heinen, E. (1999) *J Biotechnol* **75**, 291-295
189. Warrington, J. A., Nair, A., Mahadevappa, M., and Tsyganskaya, M. (2000) *Physiol Genomics* **2**, 143-147
190. Hubbard, M. J. (2002) *Proteomics* **2**, 1069-1078
191. MacLeod, J. C., Sayer, R. J., Lucocq, J. M., and Hubbard, M. J. (2004) *J Comp Neurol* **477**, 29-42
192. Shnyder, S. D., and Hubbard, M. J. (2002) *J Histochem Cytochem* **50**, 557-566
193. Mkrtchian, S., and Sandalova, T. (2005) *Antioxidants & redox signaling*
194. Kuznetsov, G., Chen, L. B., and Nigam, S. K. (1994) *J Biol Chem* **269**, 22990-22995
195. Kuznetsov, G., Chen, L. B., and Nigam, S. K. (1997) *J Biol Chem* **272**, 3057-3063
196. Willnow, T. E., Nykjaer, A., and Herz, J. (1999) *Nat Cell Biol* **1**, E157-162
197. Sen, J., Goltz, J. S., Konsolaki, M., Schupbach, T., and Stein, D. (2000) *Development* **127**, 5541-5550
198. Herrmann, J. M., Malkus, P., and Schekman, R. (1999) *Trends Cell Biol* **9**, 5-7
199. van Anken, E., Romijn, E. P., Maggioni, C., Mezghrani, A., Sitia, R., Braakman, I., and Heck, A. J. (2003) *Immunity* **18**, 243-253
200. Wu, C. C., Howell, K. E., Neville, M. C., Yates, J. R., 3rd, and McManaman, J. L. (2000) *Electrophoresis* **21**, 3470-3482
201. Chandra, H., Gupta, P. K., Sharma, K., Mattoo, A. R., Garg, S. K., Gade, W. N., Sirdeshmukh, R., Maithal, K., and Singh, Y. (2005) *Biochim Biophys Acta* **1747**, 151-159
202. Koen, Y. M., and Hanzlik, R. P. (2002) *Chem Res Toxicol* **15**, 699-706
203. Bo, Z., Yongping, S., Fengchao, W., Guoping, A., and Yongjiang, W. (2005) *Proteomics* **5**, 426-432
204. Myung, J. K., Afjehi-Sadat, L., Felizardo-Cabatic, M., Slavc, I., and Lubec, G. (2004) *Proteome Sci* **2**, 8
205. Hermann, V. M., Cutfield, J. F., and Hubbard, M. J. (2004) *J Biol Chem*
206. Ma, Q., Guo, C., Barnewitz, K., Sheldrick, G. M., Soling, H. D., Uson, I., and Ferrari, D. M. (2003) *J Biol Chem* **278**, 44600-44607
207. Barnewitz, K., Guo, C., Sevvana, M., Ma, Q., Sheldrick, G. M., Soling, H. D., and Ferrari, D. M. (2004) *J Biol Chem* **279**, 39829-39837
208. Tasanen, K., Parkkonen, T., Chow, L. T., Kivirikko, K. I., and Pihlajaniemi, T. (1988) *J Biol Chem* **263**, 16218-16224
209. Santisteban, P., Kohn, L. D., and Di Lauro, R. (1987) *J Biol Chem* **262**, 4048-4052
210. Kim, P. S., and Arvan, P. (1993) *J Biol Chem* **268**, 4873-4879
211. Shang, J., and Lehrman, M. A. (2004) *Biochem Biophys Res Commun* **317**, 390-396