DISCOVERY OF THE EDEM SUBFAMILY IN THE FAMILY 47

GLYCOSYLHYDROLASES

by

STEVEN W. MAST

(Under the Direction of Kelley W. Moremen)

ABSTRACT

In the Endoplasmic Reticulum (ER), misfolded proteins are retro-translocated to the cytosol and degraded by the proteasome in a process known as ER-associated degradation (ERAD). Early in this pathway, a proposed lumenal ER lectin, termed EDEM (ER degradation enhancing α-mannosidase like protein) in mammalian cells, and a homologous protein in *S. cerevisiae,* termed Htm1p (homologous to mannosidase protein), appear to recognize misfolded glycoproteins in the ER, disengage the nascent molecules from the folding pathway, and facilitate their targeting for degradation. In humans there are a total of three EDEM homologs, while in *S. cerevisiae* two homologs may be present. The amino acid sequences of these proteins are different from other lectins, but are closely related to the Class I mannosidases (Family 47 glycosylhydrolase). In this study, gene disruption of *HTM1* resulted in accumulation of the misfolded glycoprotein, mutant carboxypeptide Y (CPY*), while disruption of the second *S. cerevisiae* homolog had no effect on the degradation of CPY*. A study of Htm1p point mutations suggested that the catalytic site is not essential for function, while the oligosaccharide binding site of Htm1p is necessary for function. We have also characterized one of the EDEM homologs from *H. sapiens*, which we have termed EDEM2 (C20orf31). Using recombinantly generated EDEM2, no α-1,2 mannosidase activity was observed. In HEK293 cells, recombinant EDEM2 is localized to the ER where it can associate with misfolded α1-antitrypsin. Overexpression of EDEM2 accelerates the degradation of misfolded α1-antitrypsin indicating that the protein is involved in ER-associated degradation.

INDEX WORDS: EDEM, EDEM2, ERAD, ER-associated degradation, ER quality control, α-mannosidase

DISCOVERY OF THE EDEM SUBFAMILY IN THE FAMILY 47 GLYCOSYLHYDROLASES

by

STEVEN W. MAST

B.A., University of Virginia, 1994

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2004

DISCOVERY OF THE EDEM SUBFAMILY IN THE FAMILY 47

GLYCOSYLHYDROLASES

by

STEVEN W. MAST

Major Professor: Kelley Moremen

Committee: Jeff Dean
Claiborne Glover
Kojo Mensa-Wilmot
Michael Pierce

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2004

ACKNOWLEDGEMENTS

I would like to express appreciation to my major professor, Dr. Kelley Moremen, for his guidance and patience in my work and in the process of writing. The members of my advisory committee, Dr. Michael Pierce, Dr. Claiborne Glover, Dr. Kojo Mensa-Wilmot, and Dr. Jeff Dean, have also been of great assistance in contributing to my professional development. I am grateful to Dr. Rick Sifers, Departments of Pathology, and Molecular and Cellular Biology, Baylor College of Medicine, who contributed his time and efforts.

In addition, I would like to thank the members of Dr. Moremen's and Dr. Pierce's labs. They have provided training, fellowship, encouragement, and guidance.

I also thank my friends and family, especially my parents, who have provided much support during this time.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

LITERATURE REVIEW

The Endoplasmic Reticulum (ER) serves as a quality control checkpoint in the synthesis of polypeptides. Properly folded proteins are allowed to proceed through the secretory pathway, while misfolded proteins are retained within the secretory pathway before they are degraded. The mechanisms by which chronically misfolded proteins in the ER are recognized as substrates for degradation instead of as substrates for chaperones are poorly understood. For example, different proteins are needed for the degradation of soluble versus membrane proteins, and different proteins are also needed for the degradation of glycoproteins versus non-glycoproteins.

In the ER, degradation of glycoproteins requires trimming by an α1,2-mannosidase. This trimming step generates a glycan signal. Previous studies suggested that an unidentified lectin recognizes the glycan signal, and that lectin targets the misfolded glycoprotein for degradation (1). In *S. cerevisiae*, the lectin has been suggested to be Htm1p (homologous to mannosidase), a protein with sequence similarity to the Family 47 glycosylhydrolases (2). Htm1p is also known as Mnl1p (mannosidase like protein)(3). Homologs of Htm1p have been identified throughout the eukaryotic kingdom. The mammalian homologs, EDEM1 (ER degradation enhancing α-mannosidase like protein) and EDEM2 also appear to play a role in the degradation of misfolded glycoproteins. In order to understand the function of the HTM1/EDEM homologs, a summary of ER quality control and the Family 47 glycosylhydrolases will first be discussed.

I. The asparagine-linked oligosaccharide biosynthetic pathway

Proteins of the secretory pathway are translocated into the ER lumen. Post-translational modifications occur, including signal sequence cleavage, membrane insertion, N-linked glycosylation, disulfide bond formation, addition of glycosylphophatidylinositol anchors, and formation of oligomeric complexes. This process is slow and inefficient, yet the ER can reach high levels of efficiency in professional secretory cells (4).

Synthesis of N-linked oligosaccharides starts with the addition of monosaccharides to dolichol, a polyisoprenoid lipid carrier that anchors the oligosaccharide to the membrane. The end product of lipid-linked oligosaccharide synthesis for N-linked oligosaccharides is $Glc_3Man_9GlcNAc_2$ covalently attached by pyrophosphoryl linkage to dolichol; this oligosaccharide structure is conserved in all eukaryotes, except for trypanosomes (5). As the polypeptide is synthesized and extruded into the lumen of the ER, the oligosaccharide is added co-translationally to nascent proteins through the action of the oligosaccharyl transferase (OST) (Figure 1). OST is a multimeric complex. In *S. cerevisiae*, the OST complex includes nine different membrane-bound subunits: Nlt1p/Ost1p, Ost2p, Ost3p, Ost4p, Ost5p, Ost6p, Stt3p, Swp1p, and Wbp1p (6). For efficient transfer the OST prefers the oligosaccharide, $Glc_3Man_9GlcNAc_2$. The OST adds the oligosaccharide to the asparagine residues with the Asn-Xxx-Ser/Thr consensus sequence, where Xxx denotes any amino acid except proline. The glycan may increase protein stability, solubility, act as a tag for intracellular trafficking, and function as markers in quality control.

As soon as the N-linked oligosaccharide is transferred to the nascent polypeptide trimming of the oligosaccharide invariably involves removal of all glucose residues and removal of at least one mannose. The importance of the presence of the fully glucosylated oligosaccharide will be discussed in the next section. α-Glucosidase I catalyzes the removal of the outer α1,2 linked glucose residue. α-Glucosidase II catalyzes the removal of the inner α1,3 linked glucose residues. Removal of the innermost glucose is much slower than removal of the terminal glucose.

An alternative mechanism for removal of glucose is trimming by endo α-mannosidase. This is the only processing glycosidase that cleaves an internal glycosidic linkage, generating $Man_8GlcNAc_2$ (A isomer) (Figure 1) from $Glc_{1-3}Man_9GlcNAc_2$ (7). It is also capable of acting on oligosaccharides with truncated mannose branches, $Glc_1Man_{8-4}GlcNAc_2$. Endo α-mannosidase activity has been shown to be cell-specific (8). α-Glucosidase and α-mannosidase inhibitors have no effect on endo α-mannosidase (9), which corresponds with endo α-mannosidase not sharing homology with the α-glucosidases or α-mannosidases, and no other related proteins have been found (10).

With the exception of endo α-mannosidase, mannose trimming occurs independently of the presence of the glucose residues; however, on the branch that the glucose residue is attached the glucose residues need to be removed before mannose trimming can occur on that branch. In *Saccharomyces*, removal of a single mannose residue by ER mannosidase I (*MNS1*) is the only mannose cleavage step, and disruption of *MNS1* resulted in a mutant that was deficient in cleavage of $Man_9GlcNAc_2$ to $Man_8GlcNAc_2$. Δ*mns1* cells exhibited normal growth and mannan extension (11). In mammalian cells, inhibition of ER mannosidase I with kifunensine treatment also does not result in changes in cell morphology or growth. The trimming by ER mannosidase

I is the last fully conserved step in glycoprotein processing between yeast and mammals, and its conservation may be a result of its role in quality control. Following the ER mannosidase I trimming step, in *S. cerevisiae* further processing occurs in the Golgi where mannan structures are formed by mannosyltransferases. In higher eukaryotes, other α-mannosidases can act in the ER. ER mannosidase I can produce the $Man_8GlcNAc_2$ B isomer from $Man_9GlcNAc_2$, while endo αα-mannosidase can act on glucosylated oligosaccharides to produce the $Man_8GlcNAc_2$ A isomer (7). ER mannosidase II (glycosylhydrolase Family 38) can trim $Man_9GlcNAc_2$ to a $Man_8GlcNAc_2$ C isomer (12,13). Expression of ER mannosidase II is variable, and in HEK293 cells its action is not detectable (14). If the protein is assembled correctly, it is exported to the Golgi, where further mannose trimming by Golgi α1,2-mannosidases (glycosylhydrolase Family 47) cleave the oligosaccharide to $Man_5GlcNAc_2$; these enzymes are localized to the Golgi, but they may also be found in the ER in some cell types (15). Trimming by Golgi mannosidase II (glycosylhydrolase Family 38) is the last α-mannosidase involved in the biosynthesis of N-linked oligosaccharides. Golgi mannosidase II and Golgi glycosyltransferases catalyze the last steps of oligosaccharide synthesis, branching and extending glycans into a diverse group of high mannose, hybrid, or complex glycans.

II. ER quality control

The biosynthesis and folding of secretory and membrane proteins in the ER is subject to quality control (Figure 2). The ER quality control process can be viewed as having two opposing functions: the correct folding of nascent polypeptides and the degradation of terminally unfolded intermediates. For the former process, chaperones and folding enzymes facilitate the

proper folding and assembly of proteins and oligomeric complexes in the lumen of the ER. Transcription of many of these folding chaperones can be induced by the unfolded protein response (UPR) through an ER to nucleus signal transduction pathway. On the latter side of ER quality control, general protein translation can be suppressed by UPR, while transcription of the ER-associated degradation (ERAD) machinery is induced, misfolded proteins are prevented from anteriograde transport by retrieval to the ER, and misfolded proteins are eliminated by ERAD or lysosomal/vacuolar disposal.

As polypeptides are imported into the lumen of the ER and undergo the process of folding, some proteins are able to fold on their own, but many require assistance for folding and assembly. The quality control mechanism of the ER needs to be able to distinguish between native and unfolded proteins to ensure that properly folded proteins can be exported, while non-native proteins are paired with chaperones to facilitate correct folding. To discriminate between folded and unfolded proteins, chaperones and folding enzymes, such as BiP, GRP94, protein disulfide isomerase, calnexin, and calreticulin each detect different features, such as exposed hydrophobic sequences, unpaired cysteines, or glycan structures. This system is highly redundant in mammalian cells, and if one chaperone fails to bind to an unfolded protein, it is probable that another one will (16). Apart from actively folding proteins, certain chaperones may also provide quality control by retention of unfolded proteins in the ER. The sequestering of these proteins allows the protein to fold on its own without interference from the milieu of proteins in the ER. The formation of aggregates may also exclude proteins from vesicles that transport material from the ER (4).

There is evidence for 2 distinct chaperone pathways available for facilitating nascent glycoprotein folding. In the calnexin/calreticulin cycle (Figure 3), monoglucosylated

glycoproteins are bound by calnexin or calreticulin (17-19). Calnexin and calreticulin are homologous lectins, which act as molecular chaperones in conjunction with the thiol oxidoreductase, ERp57. Release from calnexin and calreticulin is mediated by the action of □-glucosidase II. Upon release, three options are available to the glycoprotein. First, a protein released from the chaperone in a properly folded state is no longer retained in the ER. Transport to the Golgi occurs by either receptor-mediated export or by bulk flow of fluids and membrane. A few examples of cargo receptors include ERGIC-53 which recognizes high mannose structures on cathepsin C and blood clotting factors V and VIII, LST1 which recognizes plasma membrane H^+-ATPase, and p24 which recognizes invertase (20). Second, if the protein has not properly folded after its release from calnexin or calreticulin, it can be recognized by (UDP)-glucose:glycoprotein glucosyltransferase (GT), which reglucosylates non-native glycoproteins. GT is the folding sensor in the calnexin/calreticulin folding cycle, and it recognizes both the carbohydrate and nascent polypeptide (21,22). With the re-addition of the glucose, the glycoprotein is able to bind to calnexin/calreticulin for another attempt at folding. Third, upon release from calnexin/calreticulin, the protein is degraded. Prolonged ER-retention increases the probability of trimming by ER mannosidase I. This enzyme generates structures likely to be recognized by EDEM1, which directs the misfolded glycoprotein for degradation.

While calnexin/calreticulin acts on glycoproteins, another chaperone "pathway" is necessary for non-glycoproteins. This pathway utilizes the presence of unfolded regions of proteins containing hydrophobic residues, and this chaperone system appears to exist as a large complex. In mammalian cells, a multimeric complex composed of hsp70 proteins (BiP, GRP170), GRP94, several oxidoreductases (CaBP1, Erp72, and Protein disulfide isomerase (PDI)), Erdj3 (hsp40 protein with PDI, thioredoxin, and J domains), cyclophilin B (a peptdiyl-

prolyl isomerase), UDP-glucosyltransferase, and SDF2-L1 (member of the O-mannosyltranferase family and inducible by ER stress) was detected (23). With the exception of Erdj3 and to a lesser extent, PDI, this complex is still maintained in the absence of unfolded proteins. Proteins are able to interact with both chaperone systems, but the interactions are separated both temporally and spatially. In one example, vesicular stomatitis virus glycoprotein G binds first to BiP and then to calnexin (24). Immunofluorescence studies also show distinct localizations of calreticulin/calnexin and BiP/PDI/GT within the ER of cells where proteasomal degradation was inhibited (25).

Compared to mammalian cells, *S. cerevisiae* has a stripped down quality control system, lacking the homologs for GRP94, calreticulin, and GT. Without GT, *S. cerevisiae* lacks a major component of the calnexin cycle; however, other yeasts do have GT. Cne1p, the yeast homolog of calnexin, still plays an undetermined role in the folding and degradation of proteins. *Saccharomyces* lacking Cne1p have a slower degradation rate of misfolded pro-□□factor compared to wild-type yeast (26), while in *S. pombe,* a knockout of calnexin is lethal (27). An important backup system for quality control is present in *Saccharomyces;* misfolded proteins that escape the ER are diverted from the Golgi to the vacuole for degradation. In mammalian cells, rerouting of misfolded proteins from the Golgi to the lysosomes is possible but does not appear to be a major quality control pathway (20).

II-1. Calnexin and calreticulin

Calnexin and calreticulin are homologous lectins that act as chaperones for glycoproteins in the ER. The lumenal domains of calnexin and calreticulin are highly similar. Calnexin is a type-I integral membrane protein. Calnexin associates with the translocon via Sec61□, and the

interaction is regulated by the phosphorylation of the cytosolic domain of calnexin (28-30). Calreticulin is a soluble homolog with a carboxyl-terminal KDEL ER retrieval sequence. The protein sequence of calnexin and calreticulin is composed of three regions: the N-, P-, and C-domains. The N-domain comprises a sheet-rich globular structure. In the crystal structure of calnexin, this globular domain contains the monovalent glycan binding site and a putative Ca^{2+} binding site (31). The lectin domain of calnexin is structurally similar to ERGIC-53, the galectins, and the legume lectin family, but little sequence similarity exists between them (32). The P-domain is rich in prolines and contains short sequence repeats. This forms a hairpin motif, which forms a curved arm that extends away from the lectin domain. Erp57, a thiol oxidoreductase of the PDI family, acts as a co-chaperone with calnexin and calreticulin. Erp57 is found non-covalently associated with calnexin and calreticulin, and Erp57 promotes the formation of correct disulfides in the folding protein (33). Experiments using deletion constructs and experiments using NMR have shown that Erp57 binds to the tip of the P-domain of calreticulin (34,35). The C-domain of calreticulin is enriched in acidic amino acids, and is known to play a role in calcium binding. In contrast, the C-domain of calnexin does not share this sequence. Instead, this region acts as a linker between the membrane anchor and the globular lectin domain.

The role of cofactors has not been fully determined. Ca^{2+} appears to play a structural role. Calnexin also has a Zn^{2+}-dependent site of interaction with Erp57, and the addition of 10-100μM Zn^{2+} results in a conformational change in calreticulin, increasing its hydrophobicity and its chaperone activity, although the physiological relevance of this is unclear (34). Calnexin and calreticulin both bind ATP *in vitro*, but the physiological significance of ATP binding also

remains unclear (36,37). It is unknown whether the ATPase activity is from calnexin or an associated protein (32,38).

Unlike many other chaperones that recognize only the protein, calnexin and calreticulin are lectins that recognize glucosylated proteins, and subsequent protein-protein interactions may facilitate the folding process. Calnexin and calreticulin recognize the terminal glucose moiety on high mannose N-linked glycans ($Glc_1Man_{7-9}GlcNAc_2$) in the ER. Molecular chaperones appear to act by keeping the substrate sequestered, rather than by actively promoting folding, although an active chaperone folding role has been suggested (39). The P-domain provides a partially solvent-shielded environment, lowering the chance of aggregation that could result with other folding intermediates.

Calnexin and calreticulin interact with practically all glycoproteins that have been studied (40). For the most part, calnexin and calreticulin have the same glycan specificity and recognize many of the same substrates, but these lectins can have different roles in glycoprotein maturation (references within (41)). However, different glycoprotein substrates have been observed for the two lectins, most likely as a result of their localization. The membrane localization of calnexin versus the soluble localization of calreticulin accounts for some of the difference in specificity. For example, when calreticulin is anchored to the membrane, it has the same substrate specificity as calnexin. When calnexin was expressed as a soluble ER protein, the proteins it bound were similar to calreticulin (42). There are examples where the soluble lumenal domain of calnexin cannot replace calreticulin (43). Additionally, when calnexin and calreticulin bind to the same protein, they interact with different oligosaccharides or interact at different stages of the maturation process (40). They both can bind to MHC class I heavy chain and influenza hemagglutinin but do so at distinct stages of the folding process. Calreticulin associates with the

glycans on the hemagglutinin top domain, while calnexin interacts with the glycans found in the membrane-proximal stem domain of hemagglutinin.

There are two models for interaction of calnexin and calreticulin with the substrate. In the first model, the chaperone function is solely attributed to the lectin activity. *In vitro* experiments with RNAse B show that complexes can be dissociated solely by oligosaccharide modification, and studies of glycoprotein folding using *T. cruzi* support a lectin-only chaperone activity (44-46). There is also evidence suggesting that calnexin and calreticulin associate with non-native proteins by polypeptide interactions (references within (39,47)). First, complexes with glycoproteins cannot be disrupted by enzymatic removal of oligosaccharides. Second, calnexin associates with proteins that do not contain oligosaccharides, including proteins that do not naturally contain Asn-linked glycans. Third, treatment of cells with castanospermine or the use of cells deficient in either glucosidase I or glucosidase II to block the formation of $Glc_1Man_9GlcNAc_2$ does not stop the association of calnexin with various glycoproteins. In this study, different glycoproteins exhibited differences in the strength of the polypeptide interaction (47). Fourth, *in vitro* assays showing that calnexin and calreticulin can suppress aggregation, independent of the glycosylation state (36,37,39). This study indicates the globular domain is responsible for suppressing most of the aggregation of nonglycosylated proteins. Polypeptide contacts with the substrate also occur in the P domain, and although this domain alone cannot suppress aggregation, it does enhance the chaperone activity (34). However, the temperature at which the *in vitro* studies were carried out was above the melting temperature of calnexin (personal communication, Ari Helenius).

Release of the glycoprotein from these lectin chaperones occurs by the action of glucosidase II. Glucosidase II trimming is not dependent on the folding state of the glycoprotein,

but its activity is affected by the oligosaccharide structure. Deglucosylation of $Glc_1Man_8GlcNAc_2$ is 80% less efficient than it is for $Glc_1Man_9GlcNAc_2$ (48). Many studies on calnexin/calreticulin use the glucosidase inhibitors, castanospermine or deoxynojirimycin, to prevent the formation of monoglucosylated intermediates. With these inhibitors, access to calnexin and calreticulin is prevented. When the calnexin/calreticulin cycle is bypassed by glucosidase inhibitors, the stringency of quality control is reduced (49,50). Using influenza HA as a model substrate, the presence of either calnexin or calreticulin was sufficient to prevent forward transport of misfolded HA; however, when cells were treated with glucosidase inhibitors, HA was secreted. BiP and other chaperones were unable to correctly fold the protein or to effectively retain HA (40). In other studies using glucosidase inhibitors, it was found that high molecular weight complexes are formed and glycoproteins were retained in the ER (41).

II-2. UDP-glucose:glycoprotein glucosyltransferase

Removal of glucose from the N-linked oligosaccharide by glucosidase II frees the glycoprotein from the calnexin/calreticulin cycle. Monoglucosylated glycoproteins can be recreated by the UDP-glucose:glycoprotein glucosyltransferase (GT), allowing re-entry into the cycle. GT activity has been detected in mammalian, insect, protozoan, plant, and fungal cells, but no activity has been detected in *S. cerevisiae* (51).

GT is an ~160 kDa soluble protein found in the lumen of the ER that adds a single glucose in an α1,3 linkage to the same mannose from which a glucose was removed by glucosidase II (52,53). The efficiency with which GT acts on the oligosaccharide varies, ranging from an efficiency of 1 for $Man_9GlcNAc_2$, 0.5 for $Man_8GlcNAc_2$, and 0.15 for $Man_7GlcNAc_2$ (19).

GT acts as a sensor of folding. It glucosylates misfolded glycoproteins that have partially structured conformations but does not act on native glycoproteins. GT binds to hydrophobic peptides and has been shown to act on a domain level so that only glycans in a structurally destabilized area are reglucosylated while other glycans in a folded domain of the same protein are not glucosylated (54). GT does not recognize glycopeptides with only a few amino acids. The proper folding of glycoproteins likely hinders its recognition by GT, as the innermost GlcNAc is recognized by GT and can be inaccessible in folded proteins (53).

II-3. Glucosidase I and II

α-Glucosidase I is an α1,2 glucosidase, which removes the outermost, or terminal, glucose residue. α-Glucosidase I is a type-II transmembrane protein, which belongs to the Swiss Prot glycosylhydrolase Family 63. It is 92 kDa in *S. cerevisiae* and 85 kDa in mammalian cells, has no apparent requirement for divalent cations (12), and has a pH optimum of 6.4-6.8 (9). α-Glucosidase II (glycosylhydrolase Family 31) purified from rat liver was found to be a heterodimer of α and β subunits of 110 kDa and 80 kDa, respectively. α-Glucosidase II is a soluble protein, and it appears to be retained in the ER by the KDEL ER retention sequence located in the β subunit. Although glucosidase I and glucosidase II are ER resident proteins, they have distinct localizations within the ER (12). Cell lines deficient in glucosidase I (CHO derived Lec23) or glucosidase II (BW5147PHA^{R}2.7 mouse melanoma) are able to produce complex oligosaccharides, suggesting the presence of an alternative pathway for glucose trimming (9). In the glucosidase II-deficient cell line, Golgi endo α-mannosidase was able to bypass the glucosidase II step, allowing for processing to complex structures (55). This enzyme

is likely to act on glucosylated oligosaccharides that have escaped the ER. In contrast, *S. cerevisiae* cells with an inactive glucosidase I mutant show no alternative pathway(56).

II-4. HSP70/BiP

The stress inducible 70 kDa heat shock proteins (Hsp70) and constitutively expressed heat shock cognate proteins (Hsc70) are part of a family of molecular chaperones that bind and release polypeptides in an ATP-dependent cycle (57). The Hsp70 analog, BiP/GRP78 (Kar2p in *S. cerevisiae*) is the most abundant ER chaperone, and it performs many roles. BiP recognizes unfolded polypeptides and unassembled subunits. BiP retains these proteins in the ER and inhibits their aggregation. BiP interacts with a wide variety of polypeptides, which usually exhibit high hydrophobicity (58). BiP plays an essential part in polypeptide translocation. In protein import, BiP interacts with the J-domain of Sec63p (59). BiP maintains the permeability barrier of the ER by sealing the lumenal end of the translocon before and after translocation (60). BiP also has an important role in modulating UPR (see UPR section). BiP is required for the ERAD of many misfolded proteins. For example, BiP is required for ERAD of CPY*, mammalian Pi Z in yeast, and pro □ factor, which are all soluble proteins (57). BiP is not required for 4 known transmembrane ERAD substrates.

The structure of BiP has not been determined, but the structure of the ATPase domain of Hsc70 (61) and the substrate binding domain of DnaK (62) have been determined. Interaction with a peptide stimulates ATP hydrolysis, and the resulting conformational change increases the stability of the complex by closing a lid to trap the substrate. Exchange of ADP for ATP is needed for release of the substrate. When ATP is bound, the helical lid opens (57) and the polypeptide is released (63). BiP has a weak intrinsic ATPase, but there are co-chaperones,

which stimulate ATPase activity. The ATPase activity of the HSP40 protein promotes stable substrate binding. In *S. cerevisiae*, Kar2p (BiP) interacts with three HSP40 proteins, Sec63p, Scj1p, and Jem1p. Sec63p is part of the translocon. Scj1p is an ER lumenal protein and acts with BiP to facilitate folding. Jem1p is a single spanning ER membrane protein and is required for membrane fusion during karyogamy along with Kar2p and Sec63p (58). In mammalian cells, there are several Hsp40 proteins, with which BiP could potentially interact.

II-5. HSP40

The hsp40 family members are co-chaperones that regulate BiP. The hsp40 proteins catalyze the ATPase reaction, thus facilitating the binding and release of substrate. The ATPase cycle of hsp70 proteins is regulated positively and negatively by several chaperones and cofactors (DnaJ, GrpE, Hip, Hop, and Bag-1), although the mammalian ER homologs are not as well characterized as the cytosolic Hsp40s (23). In the ER, the main hsp40 members are Erdj1-5 (4,64,65) and Sec63. Like the other hsp40 members, these proteins interact with BiP via the J-domain in an ATP-dependent fashion.

II-6. HSP90

The hsp90 family of chaperones participates in the maturation of proteins. This family of proteins is expressed in response to cell stress, including heat shock, starvation, and oxidation. GRP94 is one of the most abundant ER resident proteins and is an ER paralog of cytosolic Hsp90 (23), and it facilitates folding of a subset of membrane-bound proteins, such as Toll-like receptors and integrins, as well as secreted proteins, such as IgG. GRP94 differs from Hsp90 in that it exhibits weak ATP-binding and hydrolysis activity.

The structure of GRP94 has been solved, and GRP94 contains a N-terminal domain that binds the nucleotide, a charged region that acts as a flexible tether, and a C-terminal dimerization domain (66). A second dimerization site has not been found, but multimeric hsp90 aggregates have been observed.

II-7. Peptidyl-prolyl isomerases

Peptidyl-prolyl isomerases (PPIs) catalyze the cis-trans isomerization of peptide bonds N-terminal to proline residues and are involved in the folding of newly synthesized proteins. PPIs are divided into three main classes, which are structurally distinct: cyclosporine A-binding cyclophilins, the FK506-binding proteins, and the Parvulin-like PPIs.

II-8. Ero1

In yeast and mammalian cells, members of the Ero1 family oxidize PDI. Oxidizing equivalents flow from Ero1 to cargo protein via PDI through disulfide exchange. Ero1 is a FAD-containing, membrane-associated protein that is oxidized directly by molecular oxygen although it can probably use alternate electron acceptors under anaerobic conditions (67). Control of FAD levels may be a posttranslational way for the cell to regulate Ero1 (68). In *S. cerevisiae,* there is one essential ERO1 gene, but Erv2p performs a similar function (69). There are two mammalian ERO1 genes, but no Erv2p orthologs. hEro1-Lα is constitutively expressed, while hEro1-Lβ as well as the yeast Ero1p are induced by UPR (70,71). In mammalian cells, both forms of Ero1 can alter the redox state of PDI, but do not affect the oxidoreductase, Erp57. Ero1p has also been proposed to play a role in the degradation of proteins by catalyzing the redox-dependent dissociation of PDI from unfolded proteins (72).

II-9. Oxidoreductases

In the highly oxidizing environment of the ER, disulfide bond formation is favored. The redox potential of the ER is –230mV compared to –150mV for the cytoplasm (57). Protein disulfide isomerase (PDI) is the prototypical oxidoreductase in the ER. It is an essential protein that constitutes ~2% of the protein in the ER, and it is one of the major chaperones that catalyzes the rearrangement of disulfide bonds (68). There are four PDI homologs in yeast (EUG1, MPD1, MPD2, and EPS1) and dozens more homologs in higher eukaryotes, including ERp72, P5, PDIR, TMX, ERp44, Erdj5, ERp19/ERp19, and ERp46 (68). These proteins contain one or more thioredoxin-like domains, but it is not clear what role these homologs have, and whether they have roles in reduction, oxidation, and/or isomerization. PDI belongs to the thioredoxin superfamily and consists of two homologous catalytic domains (a and a') separated by two homologous non-catalytic domains (b and b') and ends with an acidic c-domain that contains the KDEL retrieval sequence. The primary peptide substrate binding domain is found in b'. The b domain probably has a structural role (73). All four of the domains appear to have the same thioredoxin fold. Thioredoxin is a 12 kDa protein that is a disulfide reductase in the cytosol. The chaperone activity of PDI prevents non-productive interactions from occurring and is not dependent on the redox state of PDI, but the enzymatic activity of PDI is dependent on the redox state. Therefore, it can act as an oxidase, reductase, or isomerase (74). Mammalian PDI catalyzes protein folding when in an oxidized state, but catalyzes protein unfolding when reduced (72).

ERp57 shares extensive similarity to PDI, including similar domain organization, abb'a', although ERp57 lacks the charged region at the C-terminus that PDI contains. The b and b'

domains appear to be the principle binding site with calreticulin (75). Through its interaction with calnexin and calreticulin, ERp57 acts as glycoprotein-specific oxidoreductase and is able to catalyze the formation, reduction, and isomerization of disulfide bonds. Erp57 forms a 1:1 complex with calnexin or calreticulin (69). The affinity of Erp57 and the calreticulin P-domain was ~9 μM; however, the interaction is likely stabilized in a mixed disulfide ternary complex with the glycoprotein substrate (41). It has been suggested that ERp57 plays a role in reducing the intrachain disulfides in MHC class I molecules (76), similar to the activity proposed for PDI in retro-translocation of the cholera toxin A chain (72). It also contains a cysteine protease activity, which may act in peptide trimming of MHC class I ligands (77).

III. ER-associated degradation

In the ER, misfolded proteins are retained and targeted for degradation by the proteasome after retro-translocation to the cytosol and ubiquitination. This process is referred to as ER-associated degradation (ERAD). There are different pathways in ERAD for nonglycosylated and glycosylated proteins, although both can undergo ubiquitin-mediated proteasomal disposal (Figures 4 and 5). Ubiquitin-independent proteasomal degradation pathways (78,79) and alternative non-proteasomal degradation pathways also exist, but little is known about their mechanism of action (80).

How non-glycoproteins are targeted for ERAD is unknown, but recognition likely involves sensing the maturation state of newly synthesized proteins. For glycoproteins, several studies in mammalian cell lines have shown that the removal of mannose is important for targeting of mutant proteins for degradation in the ubiquitin-proteasomal degradation pathway.

Proteasomal-dependent degradation of T cell receptor CD3-δ subunit (81), pro/pre α factor (26), α_2-PI_{Nara}, α_2-$PI_{Okinawa}$ (82), α1-PI Z (83), and α1-$AT_{Hong\ Kong}$ (84) are inhibited by 1-deoxymannojirimycin which acts on Class 1 α-mannosidases. This work indicates that trimming by a Class 1 α-mannosidase triggers targeting for degradation by proteasomes, and that the sensing mechanism is dependent on the oligosaccharide. Genetic manipulation of yeast has shown that the combination of glucosidase and ER mannosidase I action is necessary for disposal of misfolded glycoproteins. Processing of the oligosaccharides on CPY* to $Man_8GlcNAc_2$ B isomer increased the rate of degradation of this glycoprotein, whereas blocking the processing of glycans on CPY* at $Glc_{1-3}Man_9GlcNAc_2$ or $Man_9GlcNAc_2$ on CPY* caused an accumulation of the glycoprotein in the ER (Figure 6). Similarly, transfer of the truncated oligosaccharides ($Man_{7-5}GlcNAc_2$) to CPY* caused an accumulation of CPY*(1,85). Since anteriograde transport and maturation of properly folded CPY does not appear to be dependent on a specific oligosaccharide structure, the investigators concluded the $Man_8GlcNAc_2$ isomer B structure was recognized for targeting the unfolded glycoprotein for degradation. A list of proteins which are known to be required for degradation of CPY* as well as other soluble and membrane-bound proteins from *S. cerevisiae* is shown in Table 2 (57).

The model for degradation of glycoproteins involves a glycan-based signal and an additional polypeptide-based component that involves recognition of a non-native structure (86). As described, ER mannosidase I apparently forms the glycan signal determinant, and the timing of the glycan modification plays an important role in the process. The low basal concentration of ER mannosidase I accompanied by its slow trimming allows the protein time to fold. Those proteins, which have not acquired a native state by the time their glycans are trimmed, become substrates for degradation. This was observed during overexpression of ER mannosidase I, which

diverted newly synthesized proteins to the proteasome instead of allowing time for correct folding and secretion of a functional protein (86). However, conformational maturation prevents recognition by the ERAD system.

EDEM1 appears to be a receptor that recognizes the degradation signal on the misfolded glycoprotein, removes the misfolded glycoprotein from the calnexin cycle, and sorts the aberrant protein into the degradation pathway. When terminally misfolded or unassembled proteins are recognized as substrates for degradation, they are retrotranslocated through the Sec61p into the cytosol in an as yet undetermined manner. Evidence for Sec61 as the channel for retro-translocation includes interaction of misfolded proteins (87-89)and toxins (90)with Sec61. Toxins proteins use retrograde transport to enter the cytosol. Additionally, in *sec61* mutants, degradation of several ER substrates was prevented (91). Retro-translocation uses a different mechanism than is used in import and is described here for *S. cerevisiae*. A retro-translocon complex appears to be built of at least Sec61p, Hrd1p/Der3p, Hrd3p, Cue1p, and Ubc7p (Figure 5). The ER membrane proteins Hrd1p/Der3p and Hrd3p interact with Sec61p; it is thought they may act to reprogram the translocon for retrograde transport or deliver misfolded proteins to the translocon from the lumenal side (92). Kar2p is also essential for ERAD of several proteins; its roles may include unfolding polypeptides for retrotranslocation and targeting misfolded proteins for degradation. Prior to retro-translocation, the intrachain disulfide bonds are reduced in the substrate (93). Transport across the membrane into the cytosol requires the AAA-ATPase Cdc48p (p97)/Npl4p/Ufd1p complex. On the cytosolic face, Cue1p recruits Ubc7p, an ubiquitin-conjugating enzyme, to the ER membrane. The ubiquitin ligase adds ubiquitin to the □□amino group of lysines on the target protein, or it is added to another ubiquitin, forming a polyubiquitin chain on the protein. Once across the membrane, N-linked oligosaccharides are removed by

cytosolic N-glycanase. Polyubiquitination of the protein serves as a signal, targeting it for degradation by the trypsin, chymotrypsin, and post-glutamyl hydrolase-like activities of the 26S-proteasome (94).

The α1-antitrypsin variants, N_{HK} and PI Z, are both retained in the ER in hepatoma cells, but they are degraded by different means. Treatment with the proteasomal inhibitor, lactacystin, arrested N_{HK} turnover but did not alter the turnover rate of PI Z. Instead, PI Z is eliminated by a nonproteasomal system. The means of its disposal are unknown, but this pathway is sensitive to tyrosine phosphatase inhibition. In proteasomal degradation, there is an association between calnexin and the misfolded protein, which precedes the degradation phase. Sifer's data indicates that the combined trimming by ER mannosidases I and II partitions PI Z into the non-proteasomal pathway by preventing an association with calnexin. However, overexpression of ER mannosidase I in the same cell line diverted PI Z to mostly proteasomal degradation (86).

When PI Z was expressed in HEK293 cells a distinct mechanism for PI Z disposal was observed. In HEK293 cells, PI Z is degraded via the ubiquitin-proteasomal pathway, in contrast to the non-proteasomal pathway in the Hepa1 hepatoma cell line. The extent of PI Z secretion is 4-fold higher in HEK293 cells than Hep1a, suggesting an apparent lack of polymerization of PI Z in HEK293 cells. Another difference between the cell lines is that in HEK293 cells there is a post-ER localization of the UDP-glucose:glycoprotein glycosyltransferase, while it has a reticular localization in Hep1a cells (95). Although, it is unclear what the mechanistic differences are, the evidence supports distinct pathways in different cell types.

III-1. EDEM1

Hosokawa et al. (96) cloned a cDNA corresponding to mouse EDEM1 (in the literature also referred to as EDEM). They also identified the human EDEM1 homolog (KIAA0212), which had been previously cloned (97). Mouse and human EDEM1 proteins share 92% identity. EDEM1 is an ER stress inducible gene (96), that is induced by the IRE1-XBP1 pathway (described in section IV) (98,99). The studies were done using mouse EDEM1: a type-II ER transmembrane protein, which is localized to the ER, lacks α1,2-mannosidase activity, interacts with misfolded α1-antitrypsin, and accelerates its degradation via the proteasome (96), while downregulation of EDEM1 prolonged degradation (100). EDEM1 was not able to accelerate degradation of N_{HK}, when cells were treated with kifunensine (α1,2-mannosidase inhibitor), castanopermine (α-glucosidase inhibitor), or treated with mannosamine (prevents oligosaccharide precursor elongation resulting in transfer of $Man_{5-7}GlcNAc_2$ glycans), indicating that EDEM1 prefers a Man8 glycan structure. Also, EDEM1 did not accelerate the degradation of wild-type α1-antitrypsin (101).

Oda et al. (101) found that EDEM1 (C-terminal HA tagged) appears to accelerate degradation by accepting substrates from calnexin. Molinari et al. (100) arrived at a similar conclusion using misfolded β-secretase as a substrate. EDEM1 interacts with calnexin through its transmembrane domain, and the complex exists in the absence of bound substrate. EDEM1 does not interact with calreticulin. If the ERAD substrate was not allowed to bind to calnexin by treatment of cells with glucosidase inhibitors or by using glucosidase I deficient cell lines, then EDEM1 overexpression did not accelerate degradation. This suggested that EDEM1 extracted

misfolded glycoproteins from calnexin, but not glycoproteins undergoing productive folding (100).

Hosokawa et al. (14) observed that ERAD of the α1-antitrypsin variant, N_{HK}, was enhanced by overexpression of either ER mannosidase I or EDEM1 in HEK293. Co-expression of ER mannosidase I and EDEM1 had a pronounced effect on the degradation of N_{HK} compared to either ER mannosidase I or EDEM1 alone. Additionally, they were not able to detect an interaction between ER mannosidase I and EDEM1. In HEK293 cells, recognition of a glycoprotein as a substrate for degradation appears to be mediated by EDEM1 acting as the rate-limiting step (86). In contrast, in a murine hepatoma cell line, Hepa1a, formation of the degradation signal by ER mannosidase I appears to be the rate limiting step as EDEM1 overexpression did not accelerate the degradation of N_{HK} (86).

III-2. *HTM1*

In *Saccharomyces*, a non-essential gene was identified that is required for the degradation of mutant glycoproteins. In the study by Jakob et al. (2), disruption of the *HTM1* (homologous to mannosidase) reduced the rate of degradation of mutant glycoproteins, such as CPY*, ABC-transporter Pdr5-26p, and the oligosaccharyltransferase subunit Stt3-7p, but not of the non-glycoprotein mutant Sec61-2p. Additionally, protein-bound oligosaccharides were isolated and analyzed by HPLC. This experiment indicated that disruption of *HTM1* does not alter the processing of oligosaccharides, as its oligosaccharide profile was the same as the profile of wild-type cells ($Man_8GlcNAc_2$), while Δ*htm1* cells contained $Man_9GlcNAc_2$. In the study by Endo (3), disruption of the *HTM1* gene also reduced the rate of degradation of CPY*, but not of the unglycosylated misfolded pro-α-factor, which had its 3 consensus glycosylation sites removed.

In Δ*htm1* cells, a fraction of CPY* received α1,6-mannose addition, but no α1,3-mannose was detected, indicating some CPY* reached the early Golgi whereas CPY* is strictly retained in the ER in wild-type cells. They also found that a C-terminal HA-tagged Htm1p is localized to the ER.

III-3. P-type ATPases

In *S. cerevisiae*, Cod1p/Spf1p and Pmr1p are P-type ATPases that catalyze the ATP-dependent transport of ions across membranes. Pmr1p is a Ca^{2+}/Mn^{2+} ion pump found in the medial Golgi, while Cod1p is ER-localized. Consistent with their localizations, degradation of the ER protein Hmg2p is disrupted in Δ*cod1* but not in Δ*pmr1*, and Golgi-localized carbohydrate modifications more compromised in Δ*pmr1* cells. Together they function to maintain cation homeostasis in the secretory pathway, and the extensive amount of vesicular exchange between the ER and Golgi can explain why both are needed for ion homeostasis (102). Cod1p is regulated by UPR, while Pmr1p is not. Deletion or either one slows the degradation rate of CPY*. The effects are additive in *cod1 pmr1* double mutant, with slower turnover of CPY* (102). There are several possible explanations why Pmr1 and Cod1p are needed for ERAD. *Pmr1*-null yeast display mislocalization of several proteins in the secretory pathway, so the defect may be a result of a general mislocalization of proteins. Another explanation could be the role of Pmr1p in maintaining Ca^{2+} levels in the ER (103). Cod1p and Pmr1p are only required for ERAD of a subset of proteins. Inactivation of Cod1p reduced degradation of soluble CPY* and of membrane spanning glycoprotein, Stt3-7p, but the *cod1 pmr1* double mutant was able to degrade the misfolded non-glycosylated integral membrane protein, Ste6-166p, efficiently (102). Furthermore, translocation and core glycosylation are normal in the double mutant (102).

Examination of N-glycans showed wild-type cells contained primarily $Man_8GlcNAc_2$, while Δ*cod1* or Δ*pmr1* cells contained $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ in equivalent amounts. The double mutant accumulated mostly $Man_9GlcNAc_2$. From this study, they concluded that ER mannosidase I was impaired by an imbalance of Ca^{2+}.

III-4. The translocon

Protein translocation and integration of membrane proteins occurs at the translocon, an aqueous pore in the ER membrane where polypeptides can enter the ER lumen (reviewed in (104)). The core components of the mammalian translocon are the translocon-associated membrane protein (TRAM) and Sec61, which is composed of α, β, and γ subunits.

In *S. cerevisiae*, translocation can occur co- or post-translationally. Sec61 complex, Sec63 complex, and BiP (Kar2p) are involved in co- and post-translational translocation (105). The Sec61 complex consists of Sec61p, Sbh1p, and Sss1p (the homologs of Sec61α, Sec61β, and Sec61γ, respectively). Sec61p, the core component of the translocon, is a 53 KDa integral membrane protein that spans the membrane 10 times. Sbh1p and Sss1p are small C-terminal anchor proteins. The Sec63 complex includes Sec62p, Sec63p, Sec71p, and Sec72p (106). Sec63p is a 68 KDa integral membrane protein and contains a lumenal J-domain that interacts with Kar2p (107). The role of Kar2p may be to act as a ratchet that prevents translocating proteins from slipping back into the cytosol (108).

The crystal structure of the SecY complex from *Methanococcus jannaschii* has been determined (109). SecY is the bacteria/archaea homolog of the Sec61 complex. The SecY complex acts as the channel for polypeptide transport. Although the complex is oligomeric, the structure suggests that a single copy of the SecY complex acts as the channel. The channel has

an hourglass shape, lined by a ring of hydrophobic residues at the narrowest point, which may act to form a seal around the translocating polypeptide. At the top of the hourglass, the diameter is 20-25 Å. The diameter at the center of the hourglass (the ring) is 5-8 Å. The proposed model for an increase in the ring size involves a shift in the position of the helices. In the open position, the pore would be 15-20 Å by 10-15 Å. The channel is plugged by a transmembrane helix of SecY, effectively separating the cytosol and lumen. Displacement of the plug is necessary for translocation.

Compared to its role with lumenal proteins, the role of the translocon appears to be much a more active in the integration of membrane proteins. Initially, a transmembrane sequence enters the translocon pore and can be cross-linked to Sec61□ and TRAM. After lengthening of the polypeptide by ~20 residues, it associates only with TRAM and remains with TRAM until translation is complete; then the protein diffuses laterally in the membrane.

The translocon also appears to be involved in retrograde protein transport from the ER lumen to the cytoplasm (88,90,91,110). Sec61p, Kar2p, and Sec63p are involved in retro-translocation. All of the components have probably not been identified, and the mechanism has not been fully elucidated (105).

III-5. p97/Cdc48

The p97-Ufd1-Npl7 complex is required for protein retro-translocation across the ER membrane to the cytosol. p97 (called Cdc48 yeast) is a member of the AAA ATPase family (reviewed in (111,112)). Several adaptor proteins appear to direct p97 to different roles within the cell. The adaptors, Ufd1 and Npl4, function in retrotranslocation. Retro-translocation from the ER lumen to the cytosol occurs in two distinct steps (113). The first step is passage across

the membrane, and the second step is release of the substrate. Polyubiquitination alone does not provide sufficient force for retro-translocation. Although it may keep the polypeptide from slipping back into the ER lumen, the polyubiquitin serves as a recognition signal for the p97-Ufd1-Npl7 complex, which provides the driving force for retro-translocation (114). p97 appears to have a dual role in retro-translocation; initially the nonubiquitinated protein is recognized by p97, and subsequently, the polyubiquitinated chain is recognized in a synergistic manner by the p97-Ufd1-Npl4 complex (115). Additionally, p97 can recruit the proteasome to the cytosolic face of the ER (113).

III-6. Ubiquitination in ERAD

Polyubiquitin serves as the marker for degradation. This modification is mediated by a complex of enzymes and occurs in the cytosol. With a chain of at least 4 ubiquitins, the protein can be recognized and degraded by the proteasome. In the first step, the ubiquitin-activating (E1) enzyme uses ATP hydrolysis to form a thioester linkage between a cysteine of the E1 active site and ubiquitin. The activated ubiquitin is transferred from the E1 enzyme to a cysteine in the active site of an ubiquitin-conjugating enzyme (E2 or Ubc). The E2 can then transfer ubiquitin to the ϵ-amino group of a lysine on the substrate and subsequently can form polyubiquitin chains on the substrate. The E2 can accomplish this in 3 ways: 1) the E2 directly ubiquitinates the substrate, 2) the E2 transfers the ubiquitin to an ubiquitin-protein ligase (E3 enzyme) which then ubiquitinates the substrate, or 3) an E2-E3 complex ubiquitinates the substrate. The E3 ubiquitin ligase acts in the specific recognition of substrates for degradation, and the list of proteins included as E3 enzymes is growing.

Hrd1p is an ubiquitin ligase identified in *S. cerevisiae,* and it is needed for the disposal of both ER membrane and soluble misfolded proteins and for the regulation of HMG-CoA reductase (HMGR) (116). A human homolog was identified, and its expression is induced during ER stress by the action of IRE1 and ATF6. Expression of human HRD1 in HEK293 cells protects against ER stress-induced apoptosis apparently by degrading unfolded proteins that accumulate in the ER (117). In *S. cerevisiae,* the HRD pathway is comprised of Hrd1p, Hrd3p, Ubc7p, and sometimes Der1p. Der1p is an integral ER membrane protein of unknown function; however, for some proteins, all of which are soluble, Der1p is needed for degradation (118).

Hrd1p is an ER multi-spanning transmembrane protein. Its cytosolic domain contains the RING-H2 motif that interacts with the E2 enzymes (Ubc7p or Ubc1p) and brings them into close proximity with a substrate (119). The E2 enzyme recognized by Hrd1p is Ubc7p, which is anchored to the ER membrane by Cue1p. Ubc1p is an alternative E2, which can be used in this complex. The transmembrane domain of Hrd1p forms a complex with Hrd3p, which regulates the RING domain of Hrd1p. Hrd3p has a single membrane-spanning domain, but most of the protein resides in the ER lumen, and it has been proposed that Hrd3p has a primary role in initial substrate recognition (120). Genetic data indicates that Hrd1p and Hrd3p form a complex with Sec61p, acting as a retro-translocon (92).

As the ubiquitin ligases act on the level of substrate specificity, it is not surprising that alternatives to the Hrd1p pathway exist. Hrd1p independent–proteolysis (HIP) is another mechanism for ERAD in *S. cerevisiae*. Ubiquitination of HIP substrate uses the ubiquitin ligase, Rsp5p. Rsp5p is a HECT domain ubiquitin ligase. This domain mediates E2 binding, and the structure of this domain appears similar to the RING domain. Like the HRD pathway, HIP is regulated by UPR, and both misfolded soluble (CPY* and PrA*) and integral membrane proteins

(Sec61-2p) are substrates for HIP. One major difference is that HIP substrates require transit between the ER and Golgi before degradation by the proteasome. In the Golgi, the substrate may receive a modification where upon returning to the ER the protein is targeted for degradation. The most likely Golgi modification that could act as a degradation signal is the addition of an α-1,6-mannose. Disruption of the OCH1, the α-1,6-mannosyltransferase in *S. cerevisiae*, has no effects on the degradation kinetics of CPY* (121). Hrd1p appears to be the primary degradation pathway for CPY*, although CPY* degradation is completely stopped only when both pathways are blocked (122).

A ubiquitin ligase which recognizes high mannose oligosaccharide exists. SCF^{Fbx2} (Skp1-Cullin1-Fbx2-Roc1) is a multi-subunit ubiquitin ligase complex found in the cytosol. In the complex, the F-box protein, Fbx2, selects target proteins for ubiquitination. The F-box protein is recruited to the complex by Skp1. Roc1 is a RING finger protein. It interacts strongly with Cul1 and also recruits the E2 enzyme (123). In this complex, Fbx2 specifically binds to proteins with high-mannose oligosaccharides and contributes to their ubiquitination. Fbx2 is expressed mainly in neuronal cells in the adult brain. Expression of mutant Fbx2, which is unable to form the SCF complex, blocked the degradation of ERAD substrates, CFTRΔF508 and TCR-α (124).

There are 3 other ubiquitin ligases that are known to have a role in ERAD, Doa10p, Gp78, and Parkin. They all possess a RING domain, use the E2 enzymes Ubc6 and Ubc7, and recognize non-native proteins. Doa10p is found in *S. cerevisiae,* and like Hrd1p it is an ER transmembrane ubiquitin ligase. However, its model substrates are different from Hrd1p. Doa10p recognizes solvent-exposed residues in unassembled subunits of protein complexes (119). A role for Doa10p in ERAD is inferred by increased UPR signaling in *DOA10* null yeast (125). Parkin and gp78 are mammalian proteins. gp78 is a transmembrane ubiquitin ligase that

recruits mammalian UBC7 and can mediate the degradation of the T cell antigen receptor CD3-□ (126). Parkin is mutated in patients with autosomal recessive juvenile Parkinsonism, and it likely binds unfolded proteins, including □□synuclein and the Pael receptor (119).

III-8. N-glycanase

Yeast Png1 and its mammalian homolog are cytosolic peptide:N-glycanases that most likely act prior proteasomal degradation. Png1 is a member of the transglutaminase-like superfamily. Recognition of substrate likely involves both the oligosaccharide and a portion of the polypeptide backbone, as Png1 deglycosylates misfolded proteins but does not act on folded proteins (127). It also prefers high mannose glycans over complex type oligosaccharides (128).

III-9. Proteasome

Proteasomes are responsible for most (80-90%) of the cellular protein turnover (reviewed in (129) and (130)). They are found in the cytosol and nucleus where their action includes the degradation of misfolded proteins, cell-cycle regulators, oncogenes, the activation or degradation of transcription factors, and processing of MHC-class I antigens. Substrates of the proteasome are marked for degradation by the attachment of multiple ubiquitins. The proteasome is able to recognize the ubiquitin tag, unfold the protein, and digest the protein, but before proteolysis the ubiquitins are removed. The complete 26S proteasome is 2.5 MDa and is composed of one or two19S regulatory caps and the 20S proteolytic core complex. The 19S cap is composed of 17 subunits and contains the ubiquitin receptor and also contains the ATPase subunits, which most

likely unfold the substrate for the protease activities of the 20S proteasome. The 20S proteasome has 14 different subunits, 7 distinct α subunits and 7 distinct β subunits, arranged in an $\alpha_7\beta_7\beta_7\alpha_7$ structure with a cylindrical shape. The proteolytic sites are compartmentalized inside the barrel, so only unfolded proteins can be acted on. Substrates with hydrophobic or aromatic, basic, or glutamate residues at the P1 site (N-terminal of the scissile bond) can be cleaved, so the proteasome can be described as having chymotrypsin-like, trypsin-like, and peptidylglutamyl peptide hydrolytic activities. After digestion, peptides of 4 to 25 residues are released.

IV. The unfolded protein response

The Unfolded Protein Response (UPR) is a signal transduction pathway between the ER and the nucleus, which allows the cell to adapt to changing conditions in the early secretory pathway and protect the cell from protein misfolding. The UPR is also important in the development of cells that carry a high load of secretory protein, such as plasma cells. The UPR is not induced by heat shock but is activated by ER stress. BiP appears to act as the master regulator of the pathway, controlling the expression of most, if not all UPR target genes (131). The effects of UPR are wide ranging, regulating translocation, folding, covalent modification, glycosylation, cargo selection, transport from the ER and Golgi, retrograde transport from the Golgi to the ER, vacuolar targeting, membrane biosynthesis, secretion, and degradation. While ERAD is capable of disposing of misfolded proteins under normal conditions, under conditions of ER stress, the UPR facilitates degradation by upregulating ERAD machinery and decreasing the general protein load. Loss of ERAD leads to a constitutive UPR induction, and loss of both UPR and ERAD greatly decreases cell viability (132).

In *S. cerevisiae,* Ire1p and Hac1p act in a simple, linear pathway to activate UPR (Figure 7). Ire1p is a type-I transmembrane ER-resident protein. Kar2p binds to the lumenal domain of Ire1p in unstressed cells. When bound by Kar2p, Ire1p is monomeric and unphosphorylated. During ER stress Kar2p dissociates from Ire1p, presumably to act as a chaperone for unfolded proteins, and Ire1p oligomerizes and autophosphorylates (133). Trans-autophosphorylation of Ire1p activates the cytoplasmic domain of Ire1p, and its ribonuclease activity excises a 252 nucleotide intron from HAC1 mRNA. The excision removes a stem loop, which prematurely stops translation. Rgl1p, a tRNA ligase, religates HAC1 mRNA, allowing for the production of Hac1p. Hac1p is a bZIP transcription factor, and it binds to UPR elements present in the promoter, leading to increased transcription of these genes (134). Regulation of the pathway uses Kar2p in a feedback loop. Upon rebinding of Kar2p, Ire1p is inactivated.

In higher eukaryotes, the ER stress response is extended in scope (Figure 8). UPR can be divided into three phases: an early phase of translational repression, a middle phase of transcriptional activation, and an end phase of protein synthesis in which the ER and secretory system are remodeled (133). In mammalian cells, three ER-resident transmembrane proteins have been identified as ER stress sensors, IRE1, PERK, and ATF6.

PERK is a stress-induced kinase, which inhibits protein translation during ER stress. PERK is a type-I transmembrane ER protein. The lumenal domain resembles IRE1 and acts as a stress sensor in the same manner. Dissociation of BiP from PERK results in PERK oligomerization and trans-autophosphorylation of the cytosolic domain. PERK activation leads to phosphorylation of the □ subunit of eIF2□ kinase (translation initiation factor 2□) (135), which inhibits assembly of the 80S ribosome and leads to a general inhibition of protein synthesis (136). In addition to a general inhibition of protein synthesis, PERK-mediated eIF2□

phosphorylation activates stress-induced gene expression by another pathway. Under normal conditions ATF4 mRNA is poorly translated; however, during ER stress, amino acid starvation, or arsenite treatment rapid protein synthesis of ATF4 occurs. Genes downstream of ATF4 are generally important for recovery from various stresses (136). The PERK pathway is modulated by feedback from GADD34. GADD34 requires PERK activation for its up-regulation and associates with a protein phosphatase, which dephosphorylates eIF2 (Figure 8). PERK is also regulated by $p58^{IPK}$, which is downstream of IRE1. As this occurs late in the UPR time course, $p58^{IPK}$ may act to terminate the UPR response (136).

ATF6 is a leucine zipper transcription factor that appears to be essential for ER chaperone induction, including BiP, PDI, and GrP94. Contradictory work has proposed that ATF6 is not responsible for the majority of UPR target genes, including BiP (99) . ATF6 is a type-II transmembrane protein with a lumenal sensing domain and cytosolic domain, which is the transcription factor. ATF6 is found in the ER membrane bound to BiP. Binding of BiP inhibits the Golgi localization signals of ATF6, apparently masking the signal by binding nearby. During ER stress BiP dissociates from ATF6, allowing transport of ATF6 to the Golgi where it is cleaved by the site-1 protease and the site-2 protease. The cystosolic domain of ATF6 is released from the membrane and transported to the nucleus where it binds to ERSE (ER stress response elements) and acts as a transcription factor (131).

IRE1 is a kinase and an endoribonuclease. Two Ire1p homologs (and) exist. IRE1 is broadly expressed and is essential in mice; however, Ire1 is found only in the gastrointestinal tract and is not essential. IRE-/- and XBP1-/- mouse embryos are lethal at an early stage (133). The mechanism of IRE1 activation is conserved between yeast and mammals, and XBP-1 is the mammalian equivalent of *S. cerevisiae* HAC1. XBP-1 transcription is induced by ATF6, but the

splicing of XBP-1 mRNA by IRE1 leads to production of an active transcription factor (131). The IRE1 pathway may be responsible for the activation of ERAD, and the following proteins have been identified as being XBP1 dependent: human HRD1 (ubiquitin ligase) (117), EDEM1 (98,99), Hsp40-like genes ($p58^{IPK}$, Erdj4, HEDJ), RAMP-4 (implicated in glycosylation and stabilization of membrane proteins in response to stress), and PDI-5 (homolog of PDI) (99). XBP-1 also has an essential role in hepatocyte and plasma cell development, both cell types that secrete large amounts of proteins.

Of the three ER stress sensors and their pathways, PERK is activated most rapidly, complete as soon as 30 minafter exposure to ER stress. ATF6 cleavage occurs soon after ER stress, but it requires transport of the cytosolic domain to the nucleus as well as induction of transcription and protein translation. The response of IRE1 is delayed relative to the others. XBP1 mRNA is expressed at low levels in nonstressed cells, and XBP1 mRNA requires upregulation by ATF6 (136). The time delay of IRE1 presumably allows the chaperone folding arm of the pathway to act before the degradation arm of the pathway is activated. UPR can initiate proapoptotic pathways, but the means by which this is accomplished remains poorly understood, and involves the CHOP proapoptotic transcription factor (which is under the control of ATF4 and ATF6) (133), capsase-12 , and the c-jun N terminal kinase (JNK) (136).

V. Diseases and ER quality control

After failure in one of the steps in quality control, disease may manifest itself by deficiency of a functional protein, by accumulation of misfolded proteins within the cell, or by

chronic ER stress resulting in apoptosis of certain types of cells. Several examples of disease are listed in Table 3, and a few are described below.

The ΔF508 form of the cystic fibrosis transmembrane conductance regulator (CFTR) is the most common mutation, which causes cystic fibrosis. CFTRΔF508 protein is functional, but it is completely retained in the ER and is ubiquitinated and rapidly degraded by the proteasome (137). Even wild-type CFTR chloride channel maturation is inefficient in folding. Only 25% of the wild-type CFTR proteins reach their intended cellular location of the plasma membrane (138).

α_1-antitrypsin (α_1-AT) is a glycoprotein secreted by the liver into the extracellular fluid, and its action in the lungs protects elastin fibers by inhibiting neutrophil elastase. α_1-AT deficiency causes chronic obstructive pulmonary disease, while the accumulation in the liver results in cirrhosis. There are several allelic variations of α_1-AT, which are retained in the ER and degraded by ubiquitin-dependent and -independent proteasomal mechanisms. The most common variation is caused by a single amino acid substitution; this functional mutant protein (PI Z) favors spontaneous polymerization of late folding intermediates and accumulates in the ER of the hepatocyte (78,83). In another allelic variation of α_1-AT deficiency, a truncated mutant, α_1-antitrypsin$_{\text{HONG KONG}}$, remains associated with calnexin until shortly before its degradation (27).

α_2-Plasmin Inhibitor (α_2-PI) is a plasma glycoprotein of the serine protease inhibitor superfamily, and its main function is to inhibit plasma-mediated fibrinolysis (82). Its deficiency causes hemostatic plugs to be prematurely dissolved before the restoration of injured vessels. α_2-PI$_{\text{NARA}}$ and α_2-PI$_{\text{Okinawa}}$ are genetic mutants, which are retained in the ER and then degraded by the proteasome without ubiquitination.

The formation of protein aggregates is also the main feature of several neuronal diseases. It has been suggested that in neurodegenerative diseases, failure of the UPR leads to accumulation of misfolded proteins in the ER and results in apoptosis (117). For example, mutations in presenilin-1(PS1) gene cause the early onset of familial Alzheimer's disease (FAD) by disturbing UPR activation. The FAD-linked PS1 mutation inhibits the activation of all the ER stress transducers (IRE1, PERK, and ATF6). It remains unclear how the PS1 mutation acts, but PS1 has been shown to interact with IRE1. Perhaps a BiP/mutant PS1/IRE1 complex cannot dissociate under conditions of ER stress, thus inhibiting UPR activation (139).

In Parkinson's disease (PD), neuronal degeneration may involve excessive ER stress. The drugs used to model Parkinson's disease activate the UPR in cultured neuronal cells. Neurons with an impaired UPR response were more sensitive to the PD-mimicking drugs, indicating that the ER stress caused by these drugs likely promotes cell death. In a familial form of PD, the disease is caused by a mutation in the *Parkin* gene. Parkin is an ubiquitin ligase, and loss of its activity results in the accumulation of its substrate protein in neurons, ER stress, and eventually cell death. A second form of familial PD stems from a mutation in □□synuclein. In neuronal cell cultures, overexpression of the mutant □□synuclein leads to formation of cytoplasmic aggregates, disruption of the ubiquitin-proteasomal degradation, and cell degeneration (140).

Another way that quality control is related to disease is the viral down-regulation of host proteins. Viruses are able to avoid host defenses by using the quality control pathway to their advantage. Viral antigenic fragments are presented by major histocompatability complex (MHC) class I molecules on the surface of infected cells. CD8 T cells recognize the MHC I-antigen complex and are able to destroy the infected cell. Human cytomegalovirus encodes ER-resident

transmembrane proteins, US2 and US11, which target MHC class 1 molecules for ubiquitin-mediated proteasomal degradation. US2 and US11 bind to MHC class I heavy chains and mediate an accelerated retrotranslocation to the cytosol apparently using Sec61p. However, US2 and US11 are unable to extract the heavy chain from the ER by themselves (141), and the mechanism for retrotranslocation is not understood.

HIV down-regulates host cell CD4 receptors in infected cells. CD4 is expressed on a subset of T cells, and it plays an important role in the development and maintenance of the immune system. It is also the coreceptor for HIV entry into the cell. The Vpu protein of HIV-1 is an integral membrane protein whose function is to regulate virus release from a post-ER compartment and to downreglulate CD4 and MHC class I molecules (142). Vpu induces the formation of a complex of proteins with CD4 that recruits Skp1. Skp1 is part of the Skp1-Cdc53/cullin-F-box protein (SCF) complex, which is one of the major ubiquitin ligases of the cell cycle (143). Degradation of CD4 has been shown to be an ubiquitin-mediated and proteasomal-dependent event.

VI. Glycosylhydrolase Family 47

VI-1. Class I α-mannosidases

Class I α-mannosidases, which form glycosylhydrolase Family 47, are found in eukaryotic organisms from yeast to plants and humans. The Class I α-mannosidases are distinct from the Class 2 α-mannosidases (glycosylhydrolase Family 38) in that they have different structures, mechanisms of action, and biochemical properties. The Class I α-mannosidases can be divided into 4 main subgroups (Figure 9). The first subgroup is composed of ER

mannosidase I, which is a type-II membrane protein consisting of a N-terminal cytoplasmic domain, a single transmembrane spanning domain, and a large ER lumenal C-terminal catalytic domain (144). In yeast, the cytoplasmic tail is only two amino acids, while sequence analysis of other organisms indicates that the cytoplasmic tail is variable in length and sequence. ER mannosidase I cleaves α1,2-mannose linkages, but its *in vivo* activity trims oligosaccharides from a $Man_9GlcNAc_2$ structure to $Man_8GlcNAc_2$ (isomer B), although it is capable of inefficiently trimming to $Man_5GlcNAc_2$. The enzymes require calcium, have a pH optima of 6-6.5, and are inhibited by 1-deoxymannojirimycin (dMNJ), kifunensine, and EDTA (145-148).

The second subgroup is Golgi mannosidase I. The Golgi α1,2-mannosidases are also type-II membrane proteins like ER mannosidase I. These enzymes trim $Man_{9-8}GlcNAc_2$ to $Man_5GlcNAc_2$. In the process, $Man_8GlcNAc_2$ isomers A and C are generated, and the mannose that is preferentially trimmed by ER mannosidase I is cleaved last and at a much slower rate. This indicates that there is a complementary and nearly non-overlapping specificity in substrate specificity between ER mannosidase I and Golgi α1,2-mannosidases (149). There are several isoforms of the Golgi α1,2-mannosidases (148,150,151). Golgi mannosidase IA, IB, and IC are derived from distinct genes (152,153) that are most likely not redundant, but have diverged into proteins with tissue-specific expression patterns (148). Northern blot analysis of Golgi mannosidases IA, IB, and IC shows that they have very distinct patterns of tissue specific expression, and suggests that mannosidase IB is important in embryonic development, while mannosidase IA and IC may act as housekeeping genes (12).

Within the catalytic domain, Golgi mannosidase IA shares 71% and 65% identity with Golgi mannosidase IB and IC, respectively. Golgi mannosidase IA, IB, and IC share many biochemical and enzymatic characteristics, including inhibition by dMNJ and kifunensine,

requirement for a divalent cation, and high mannose oligosaccharide substrate specificity (149,153). Recombinant Golgi mannosidases exhibit a difference in the order of removal of mannose residues from $Man_9GlcNAc_2$. Mannosidase IA and IC generate the Man8A isomer from $Man_9GlcNAc_2$, while mannosidase IB generates a mixture of Man8A and Man8C isomers (153). Golgi mannosidase IA and IB produce the same $Man_{7-5}GlcNAc_2$ isomers; however, Golgi mannosidase IC produces an even mix of two Man7 isomers. One of the Man7 isomers is equivalent to the isomer formed by Golgi mannosidases IA and IB, while the other isomer formed is also produced by insect and fungal α1,2-mannosidases (149). This suggested that the enzymes bind differently to $Man_9GlcNAc_2$. The orthologous proteins have a species- and cell-specific localization. These enzymes act predominantly in the Golgi as seen in immunofluorescence studies of mouse and human Golgi mannosidase IA, but ER localization has also been found for the pig mannosidase IA (148).

The third subgroup is the fungal secreted mannosidase subfamily. Two fungal mannosidases have been cloned and purified, and includes the *Penicillium citrinium* and *Asperillus saitoi* α1,2-mannosidases. These enzymes have an activity which is similar to the Golgi α1,2-mannosidases (154,155). They differ from the ER mannosidase I and the Golgi α1,2-mannosidases in that they contain a cleavable signal sequence and are secreted. They also have a lower pH optimum.

The EDEM subgroup is the fourth subgroup and will be covered in the Discussion.

VI-2. Structure

Crystal structures have been solved for the catalytic domain of several Class I mannosidases, including *Saccharomyces cerevisiae* ER mannosidase I (156), *Homo sapiens* ER

mannosidase I (157), Golgi mannosidase IA, *Trichoderma reesei* □1,2-mannosidase (158), and *Penicillium citrinium* □1,2-mannosidase (154). The structures are essentially the same, an $(\square\square)_7\square$barrel, which is approximately 50 Å x 50 Å x 50 Å in size (Figure 10). The barrel is constructed of fourteen consecutive □ helices, which alternate from outside to inside, forming seven parallel outer helices, concentric to the inner helices and anti-parallel to them. The two ends of the barrel are distinct. On one side, short loops connect the helices; this end is plugged by the □-hairpin, preventing the protein from being an open channel. On the opposite side of the barrel, longer loops and □-strands connect the helices. The □ strands pack together to form a series of antiparallel □ sheets. On this end, the barrel is open, allowing access to the active site, which is at the bottom of the cavity located just above the □ hairpin (156).

Comparing the conservation of protein sequence for the solved structures, the sequence similarity is not very high. For example, there is only a 38% identity between yeast and human ER mannosidase I in the catalytic domain. Also in comparing the catalytic domain, *T. reesei* □1,2-mannosidase shares 27% identity with *S. cerevisiae* ER mannosidase I, and *Penicillium citrinium* shares 23% identity with *S. cerevisiae* ER mannosidase I and 28% identity with human ER mannosidase I . Yet the carbon backbones of these proteins overlay very well (Figure 11). The helices represent an area of high conservation. The □-sheets are the region expected to be involved in binding and are not conserved. Additionally, the solvent-exposed residues are not conserved, with the exception of the residues inside the barrel. The residues of the active site are highly conserved and also overlap in the various crystal structures (Figure 12). In human and *S. cerevisiae* ER mannosidase I and *P. citrinium* □1,2-mannosidase, the conserved catalytic site residues and the calcium cation overlay with RMS deviation of 0.52 & 0.62 Å, respectively (154).

VI-3. Mechanism of Action

The Class I mannosidases are inverting glycosylhydrolases (149,159). The classic mechanism involves two carboxylic acids separated by a distance of ~9.5 Å. One acts as a general base by removing a proton from water; the other acts as a general acid by donating a proton to the leaving group (Figure 13). The end result is an inversion of the anomeric configuration at C1 and formation of a hydroxyl group at O-2. The mechanism for the Class I mannosidases appears to deviate from the classic mechanism (Figure 13), in that an activated water molecule acts as the catalytic acid (Figure 14).

In the model for the catalytic mechanism of human ER mannosidase I, Asp463 or Glu599 act the catalytic base, abstracting a proton from water. The activated water then attacks the C-1 atom. Glu330 acts as the catalytic acid. A water molecule is hydrogen-bonded to Glu330, and that water molecule is also hydrogen-bonded to the calcium. This water acts as an intermediate between Glu330 and the oligosaccharide. The end result is addition of a hydroxyl group to C-1 and inversion of configuration at C-1.

The structures of ER manosidase I with the inhibitors, 1-deoxy-mannojirimycin and kifunensine, support the idea that the protein binds optimally to the $^{1}C_{4}$ (all axial) conformation of the sugar in the active site (157). With an α1,2-linked mannose substrate bound in the active site in a $^{1}C_{4}$ conformation, a subsequent flattening of the ring at C-1 into a skewed boat conformation was proposed to be the transition state conformation.

VI-4. Calcium

Calcium is essential for activity but likely does not have a role in catalysis. Instead, its role is probably to stabilize the saccharide in the active site (Figure 15) and the enzyme. The calcium is located at the top of the β hairpin, near the active site. The ion is coordinated by 8-fold pentagonal bipyradimal coordination either directly, or indirectly via water molecules to residues that are all conserved. In the crystal structure, the inhibitors, Kif and DMJ, bind at the active site, with calcium coordinating hydroxyls of the inhibitors and stabilizing the rings of both inhibitors in the axial $^1C_4$ conformation. The structure of the active site and the $^1C_4$ conformation suggest that these enzymes optimally bind the transition state of the oligosaccharide substrate.

VI-5. Oligosaccharide binding site

As the structure of the catalytic site is conserved, the differences in substrate specificity need to be explained by interactions between the rest of the oligosaccharide and its interaction with the enzymes. The known structures of the Class I mannosidases have the same global structure and conserved residues at the active site. *T. reesei* and *P. citrinium* β1,2-mannosidase aglycone-binding sites are more spacious than ER mannosidase I, providing a wider range for binding oligosaccharides (160). *T. reesei* β1,2-mannosidase also has shorter loops at the surface, which allows glycoproteins a closer approach to the catalytic site. These factors are proposed to allow a greater degree of freedom for the β1,6 and β1,3 branches of the oligosaccharide. In modeling of the oligosaccharide (154,158) a greater variability in the conformations of the *T. reesei* and *P. citrinium* enzymes was observed, while the allowable conformations in ER mannosidase I were highly restricted. The difference in specificity is probably determined by specific interactions between the oligosaccharide and the mannosidases

throughout the binding site. One important residue is Arg^{273} of *S. cerevisiae* ER mannosidase I. This residue interacts with NAG2, M3, and M4 (Figure 15). In mammalian Golgi α1,2-mannosidases and in *T. reesei*, this residue is generally a leucine. In *P. citrinium* α1,2-mannosidase the change is from Arg to Gly. Replacement of Arg^{273} with leucine in yeast ER mannosidase I alters its specificity, allowing ER mannosidase I to trim to $Man_5GlcNAc_2$, although the order of trimming differs from the Golgi/fungal mannosidase mechanism.

CHAPTER 2

INTRODUCTION

Carbohydrate-active enzymes are classified into families by the CAZy database (http://afmb.cnrs-mrs.fr/CAZY/). The database includes organisms from eukaryotes, bacteria, and archaea. In CAZy, the Family 47 glycosylhydrolases are composed of multiple α1,2 mannosidases involved in the maturation of N-linked oligosaccharides into complex structures. Genome sequencing projects revealed novel coding regions with sufficient similarity to fall into the Family 47 glycosylhydrolases. These sequences did not fall into the established subgroups of the family, but instead formed a new subgroup, which we refer to as EDEM. However, the processing of glycoproteins in the biosynthetic pathway was already explained by the activities of the known proteins, thus opening the question of what the function of the EDEM group was.

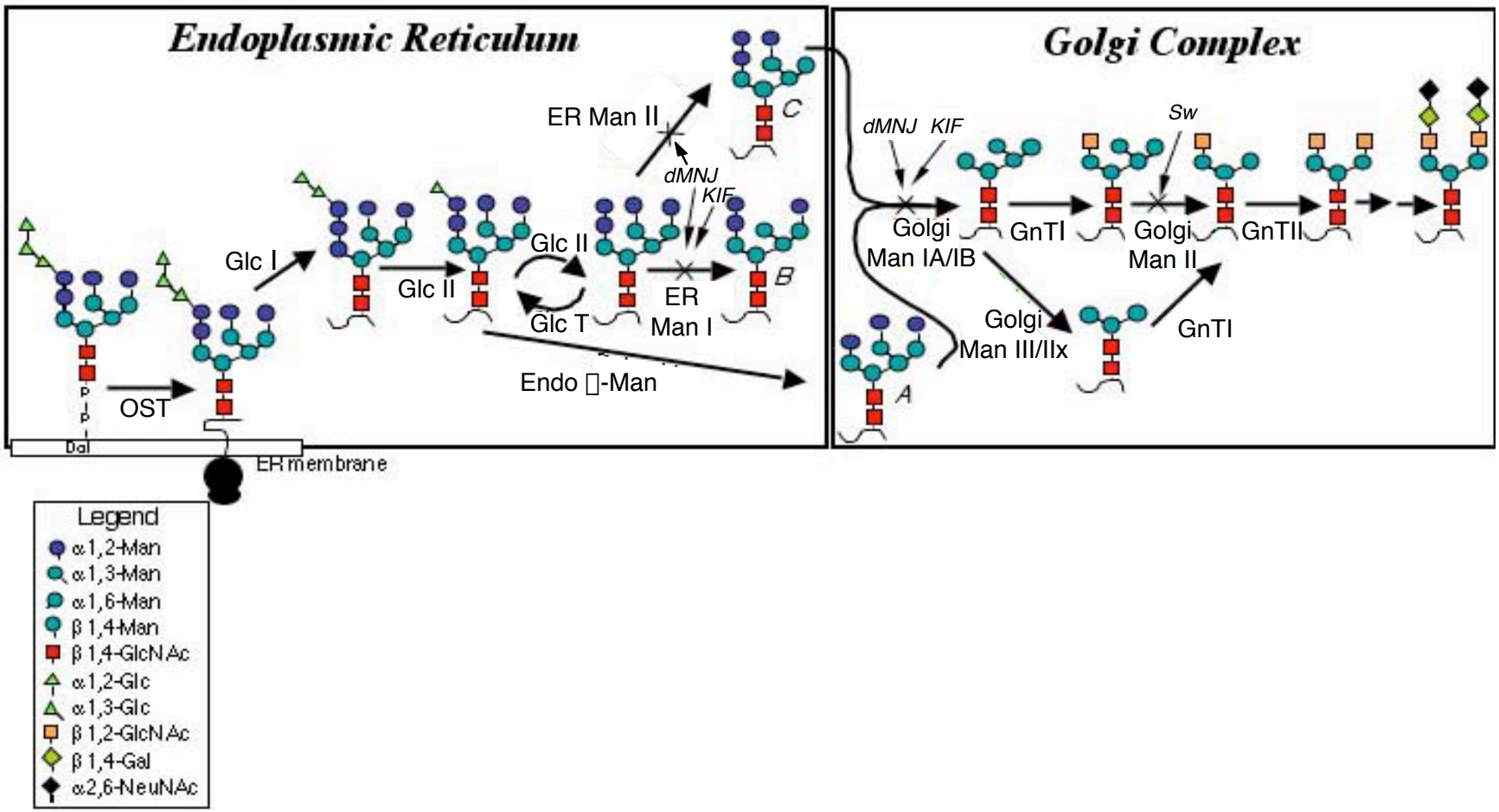
To investigate the novel Family 47 members, we chose two representative sequences, *HTM1* from *S. cerevisiae* and EDEM2 (c20orf31) from humans, in order to determine the function of this subgroup. The isolation and initial characterization of a full-length clone encoding Htm1p is covered in the first part of the dissertation, and the isolation and characterization of a full-length clone encoding EDEM2 is described in the second part. The results indicated that Htm1p and EDEM2, as well as the other EDEM members, do not have α1,2-mannosidase activity that the other Family47 members have. Instead, this subfamily appears to act in the ER-associated degradation pathway, targeting misfolded proteins for

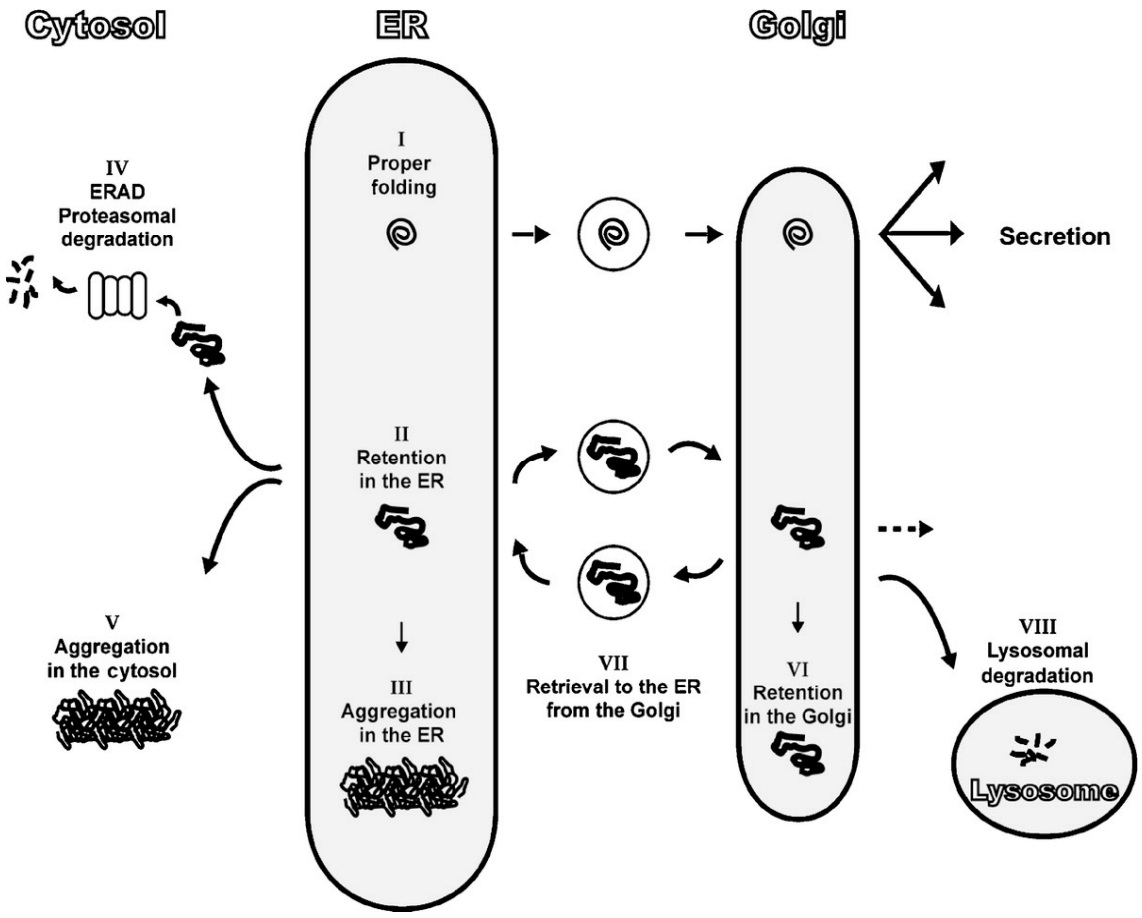
degradation.

Table 1. Table of abbreviations

AT	□1-antitrypsin
CPY*	misfolded carboxypeptidase Y
EDEM	ER degradation ehancing □-mannosidase like protein
ER	Endoplasmic Reticulum
ERAD	ER-associated degradation
ER Man I	ER mannosidase I
GT	(UDP)-glucose:glycoprotein glucosyltransferase
HBS	Hepes buffered saline
HTM1	homologous to mannosidase I (alias of MNL1)
MNL1	mannosidase-like protein (alias of HTM1)
MNS1	ER Man I
OST	Oligosaccharryl transferase

Figure 1. Asn-linked glycoprotein maturation in the ER and Golgi. The abbreviations for the enzymes are as follows: OST, oligosaccharryl transferase; Glc I, glucosidase I; Glc II, glucosidase II; Glc T, UDP-Glc:glycoprotein glucosyltransferase; ER Man I, ER mannosidase I; ER Man II, ER mannosidase II; Endo □□man, endo □-mannosidase; Golgi ManIA/B, Golgi mannosidases IA and IB; GnTII, GlcNAc transferase II; Golgi Man III/IIx, either □□mannosidase III or Golgi mannosidase IIx. Positions where processing inhibitors can act to block enzyme reactions are indicated by a thin arrow with the following abbreviations: dMNJ, 1-deoxymannojirimycin; Kif, kifunensine; Sw, swainsonine. The legend for the oligosaccharide structure displayed in the figure is indicated in the lower left.

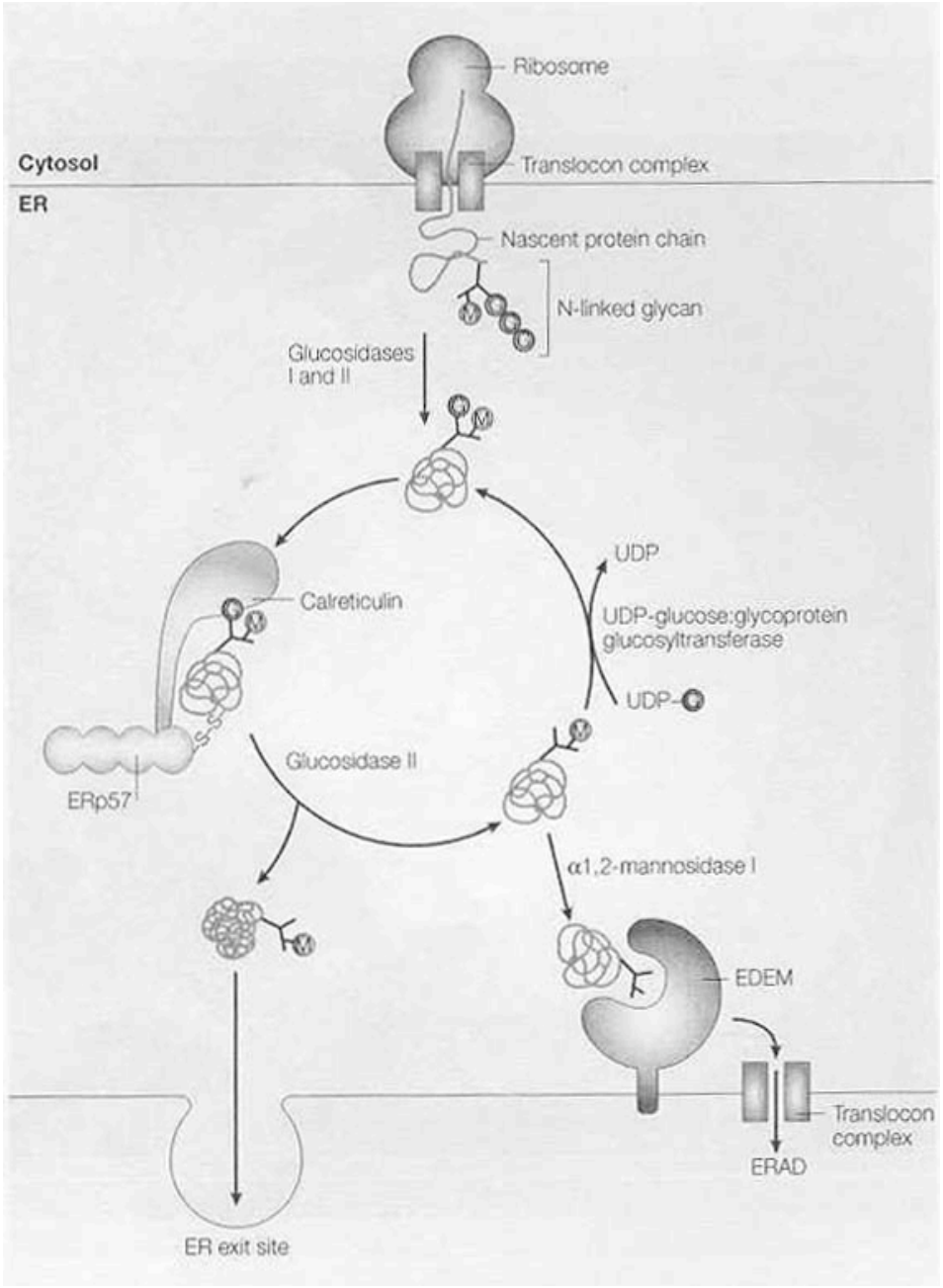

Endoplasmic Reticulum
Golgi Complex
ER Man II
dMNJ
KIF
C
B
Glc I
Glc II
Glc II
Glc T
ER
Man I
Endo -Man
OST
Dol
ER membrane
dMNJ KIF
Sw
Golgi
Man IA/IB
GnTI
Golgi
Man II
GnTII
Golgi
Man III/IIx
GnTI
A
Legend
α1,2-Man
α1,3-Man
α1,6-Man
β 1,4-Man
β 1,4-GlcNAc
α1,2-Glc
α1,3-Glc
β 1,2-GlcNAc
β 1,4-Gal
α2,6-NeuNAc



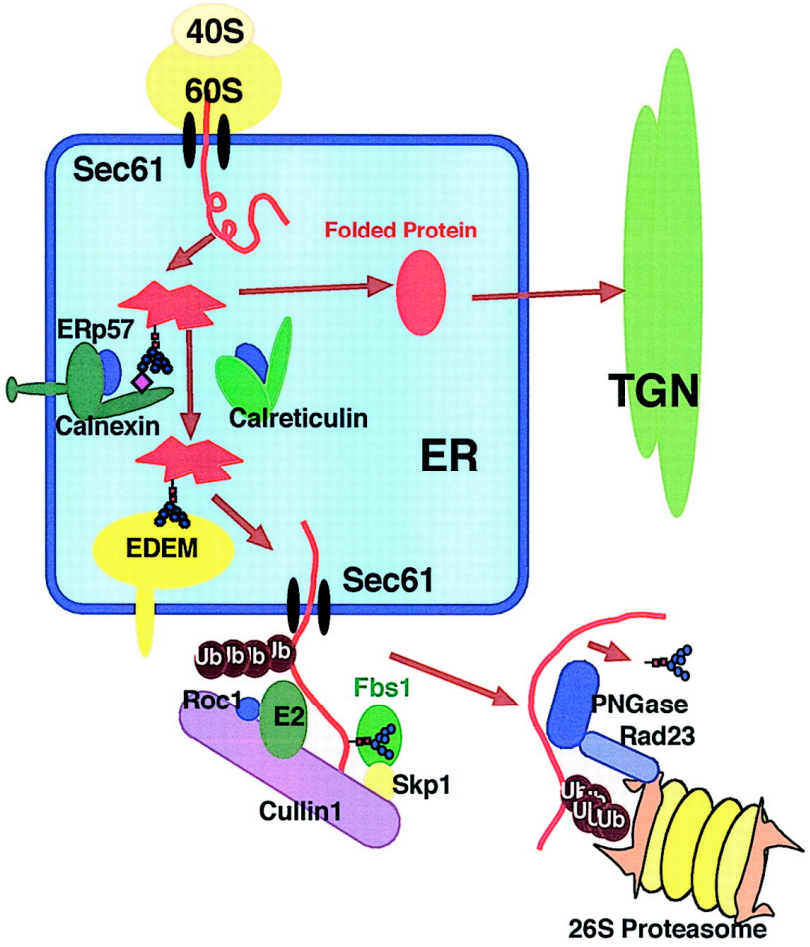
From Trombetta et al. (2003) Annu Rev Cell Dev Biol **19**, 649-676

Figure 2. Summary of quality control scenarios in the secretory pathway. Proteins that fold properly are transported out of the ER (I). Proteins that fail to fold correctly are generally retained in the ER and exposed to ER resident chaperones to attempt folding (II). They may aggregate (transiently or permanently) (III), or they may be retro-translocated to the cytosol resulting in proteasomal degradation (ERAD) (IV) or aggregation (V). Some incompletely folded proteins venture into the Golgi apparatus, where they may be retained (VI), or they can be retrieved to the ER (VII), or diverted to lysosomes for degradation (VIII).

Figure 3. The calnexin/calreticulin cycle. After transfer of the N-linked glycan to the protein, the oligosaccharide is trimmed by glucosidases I and II. The monoglucosylated glycoprotein can interact with the lectin chaperones, calnexin or calreticulin (calnexin omitted for simplicity). Both chaperones interact with the thiol-dislulfide oxidoreductase ERp57. Trimming of the last glucose by glucosidase II promotes the release from calnexin/calreticulin. At this stage, correctly folded glycoproteins can exit the ER. Misfolded glycoproteins can re-enter the cycle by re-addition of glucose by the UDP-glucose:glycoprotein glucosyltransferase. Chronically misfolded glycoproteins can be removed from the cycle after trimming by ER mannosidase I, which leads to recognition by EDEM. EDEM is a putative lectin, which targets the glycoproteins for degradation.



From Ellgaard et al. (2003) Nat Rev Mol Cell Biol. **3**, 181-91

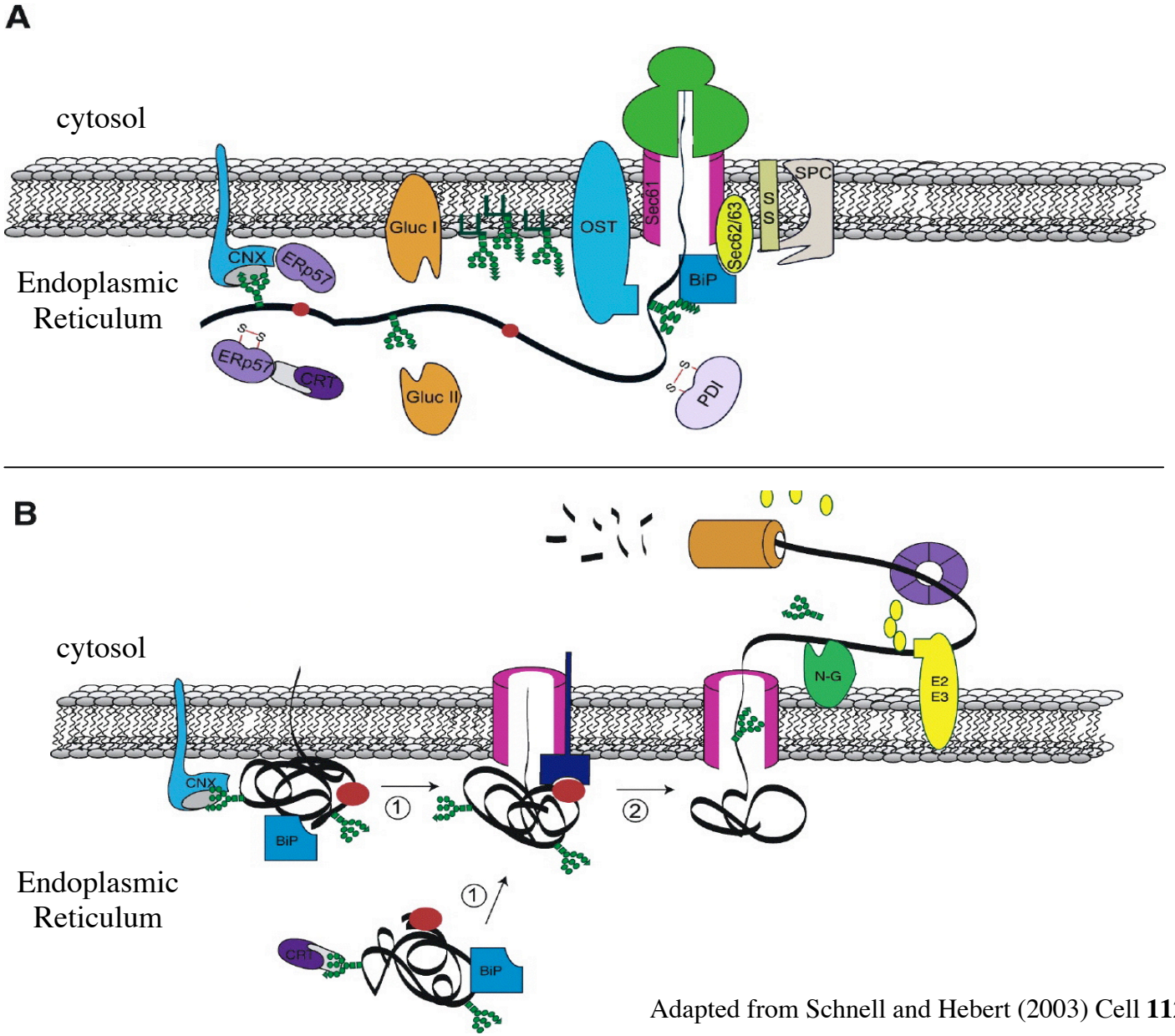


From Yoshida (2003) J. Biochem, Vol. **134**, (2) 183-190

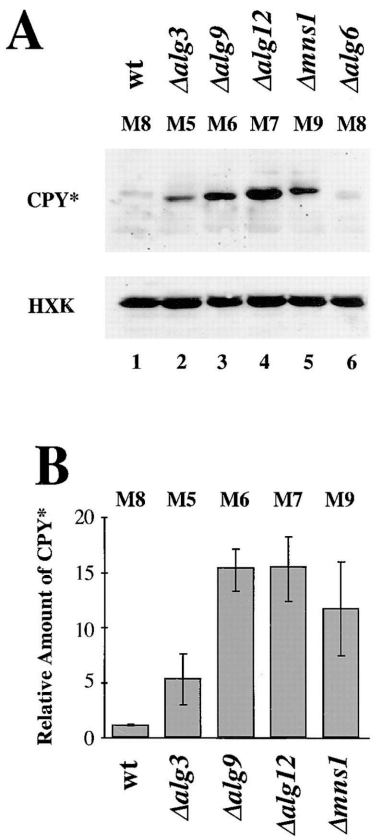
Figure 4. Glycoprotein ERAD. In this model for ERAD, misfolded glycoproteins are recognized as substrates for proteasomal degradation by a combination of ER mannosidase I trimming and interaction with EDEM. The misfolded protein is then retro-translocated to the cytosol through the Sec61 translocon with the assistance of the AAA-ATPase cdc48/p97 (not shown). In the cytosol, the misfolded protein receives ubiquitin modification by the ubiquitin ligase complex (represented here by Cullin1, Skp1, Fbs1, Roc1, and E2). The 26S proteasome degrades the protein after de-glycosylation by PNGase (cytosolic N-glycanase).

Figure 5. ER translocation and retro-translocation. (A) In cotranslational translocation several proteins interact with the polypeptide while it is still being synthesized by the ribosome. These proteins include the signal peptidase complex (SPC) that removes the signal sequence (SS), BiP, PDI, OST, glucosidase I (Gluc I), glucosidase II (Gluc II), calnexin (CRT), calreticulin (CRT), and Erp57.

(B) In this model for retro-translocation, (1) the quality control receptor (red ball) recognizes misfolded proteins and in turn is recognized by its cognate membrane bound receptor (dark blue protein). (2) Retro-translocation of a relatively unfolded protein is mediated by p97/cdc48 (purple wheel) which is recruited to the membrane by an unknown receptor. p97/cdc48 initially recognizes a non-ubiquitinated substrate. Ubiquitin (yellow ovals) are added to the polypeptide by the ubiqutin ligase complex (E2/E3). The oligosaccharides are removed by N-glycanase (N-G), and the polypeptide is degraded by the proteasome (orange cylinder). This model of ERAD is for recognition of the polypeptide. The model for glycoprotein ERAD follows a slightly different mechanism for targeting proteins for degradation, as shown in figure 4.



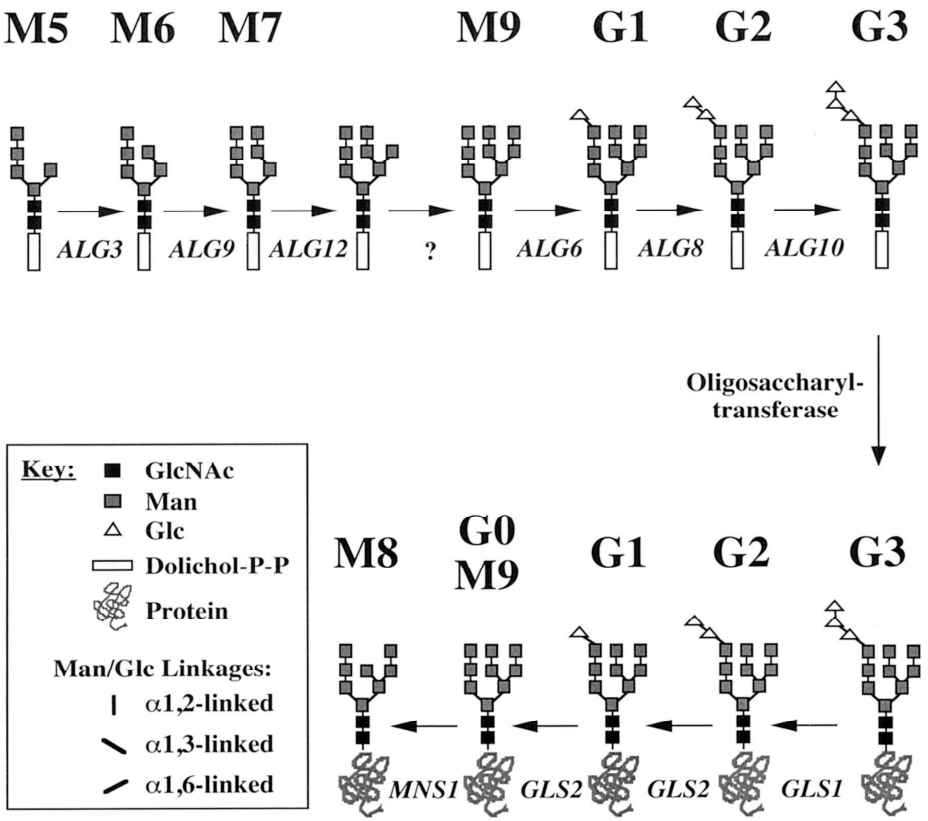
Adapted from Schnell and Hebert (2003) Cell **112**, 491-505



From Jakob et al. (1998) J Cell Biol **142**, 5, 1223-1233

Figure 6. Degradation of CPY* is determined by a specific oligosaccharide structure. Panel A, CPY immunoblot indicating the amount of CPY* accumulating in cells with mutations in various points along the biosynthesis and processing pathway of N-linked oligosaccharide structures. The membrane was reprobed with antibody against hexokinase, which served as an internal standard. Panel B, Quantification of degradation of CPY*. Panel C, biosynthesis and processing of oligosaccharides in the ER lumen of *S. cerevisiae*. The genes which encode the transferases and glycosidases are shown underneath the arrow, and the abbreviation for the oligosaccharide structure is shown above the structure.

C

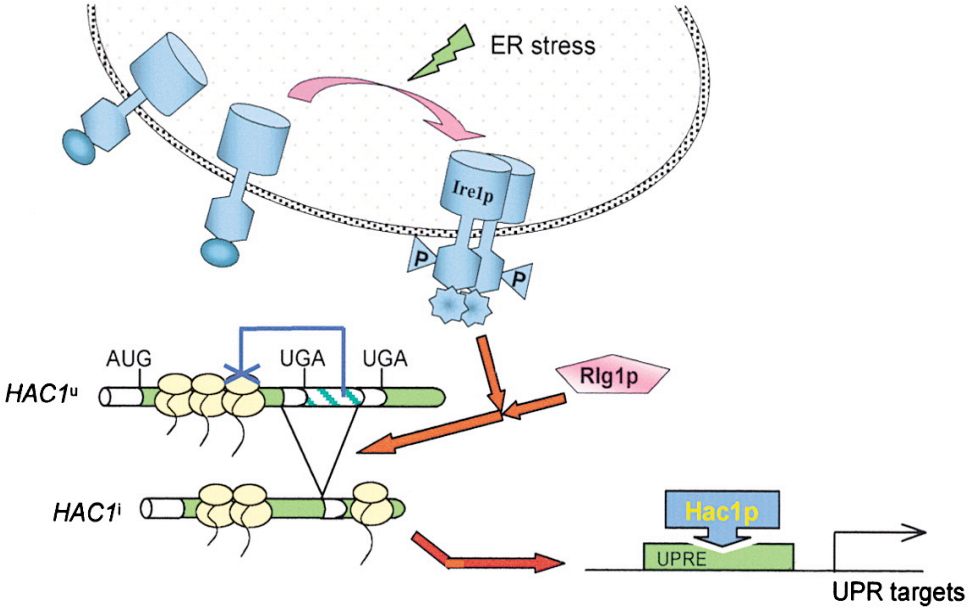


From Jakob, et al. (1998) J Cell Biol **142**, 5, 1223-1233

Table 2. ERAD requirements for substrates in yeast

Soluble			Integral membrane		
Substrate	**Required**	**Dispensible**	**Substrate**	**Required**	**Dispensible**
CPY*	BiP	Ssa1p	Ste6p*	Ssa1p	BiP
	Png1p	Cne1p		Ubc6/7p	
	Der1p	PDI			
	Hsp90		CFTR	Ssa1p	BiP
	Hrd1p/Der3p			Ubc6/7p	Cne1p
	Hrd3p				
	Sec61p		Vph1p	Ssa1p	BiP
	Cue1p				
	Pmr1p		Hmg2p	Ubc7p	Der1p
	Sec63p			Hrd1p/Der3p	Ubc6p
	Mns1p				
	Htm1p		Sec61p	Hrd1p/Der3p	Scj1p/Jem1p
	Scj1p/Jem1p			Ubc6/7p	
	Ubc6/7p			Cue1p	
	Cdc48p				
	Ufd1p		Pdr5p*	Hrd1p/Der3p	BiP
	Npl4p			Hrd3p	Der1p
				Ubc6/7p	
PI Z	BiP	Ssa1p		Sec61p	
P□F	BiP	Ssa1p			
	PDI	Ubiquitination			
	Sec61p	Sec63p			
	Cne1p	Scj1p			
	Scj1p/Jem1p	Cer1p/Lhs1p/ Ssi1p			
		Hsp90			
		Eug1p			

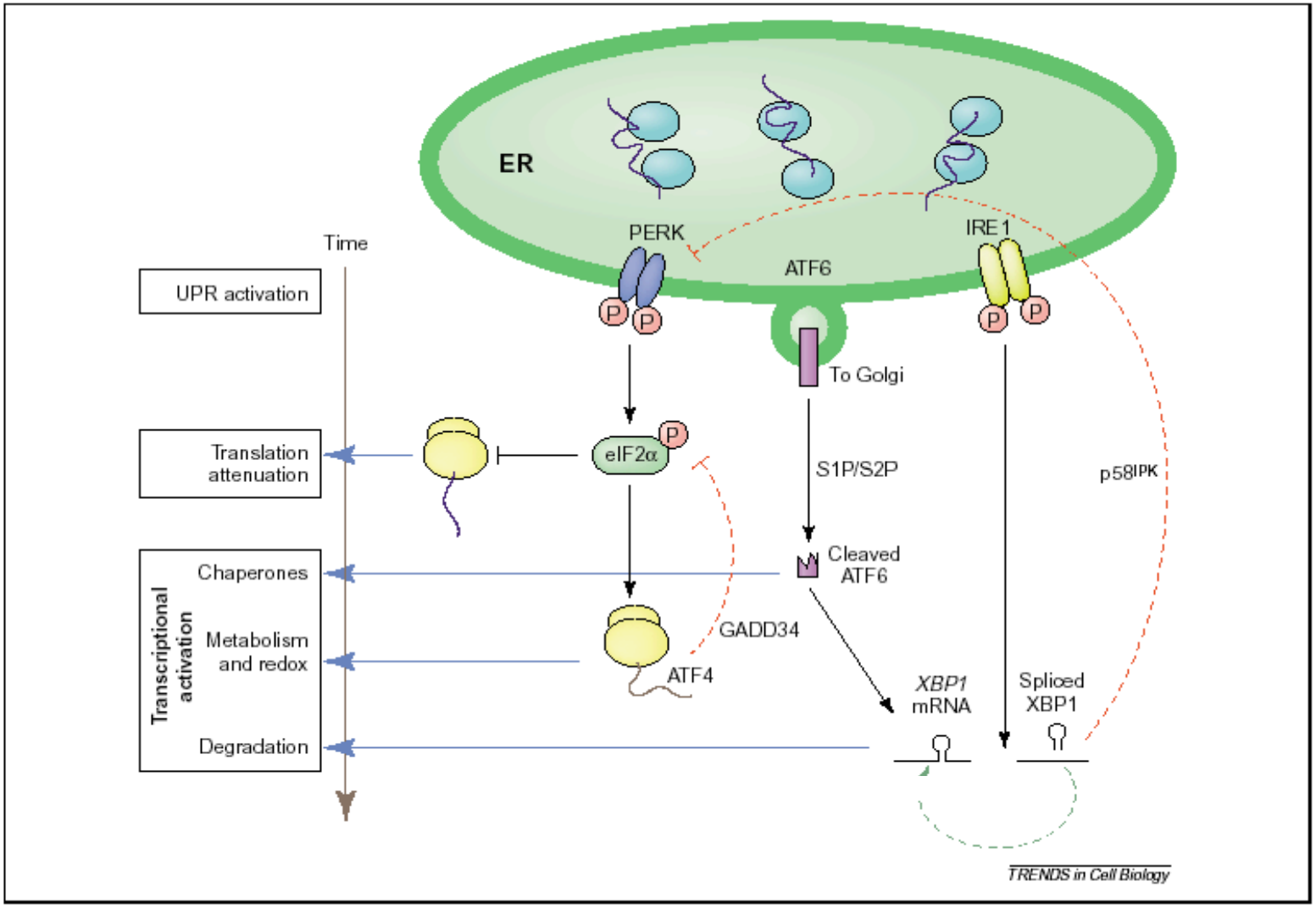
Updated from Fewell et al. (2001) Annu. Rev. Genet. **35**, 149-91



From Ma and Hendershot (2001) Cell **107**, 827-830

Figure 7. Unfolded Protein Response in *Saccharomyces*. During ER stress, Kar2p (not shown) dissociates from Ire1p, allowing oligomerization and activation of Ire1p. The endoribonuclease activity of Ire1p removes an intron from HAC1 mRNA, which inhibits its translation. Rlg1p re-ligates HAC1 mRNA in a form that is efficiently translated. Hac1p transcription factor binds to the UPREs and leads to the transcriptional upregulation of these genes.

Figure 8. The Unfolded Protein Response in mammalian cells. The ER sensors, PERK, ATF6, and IRE1 mediate distinct components of the UPR. PERK phosphorylates eIF2 to modulate translation. ATF6, when cleaved, activates transcription of chaperones and XBP1. IRE1 splices XPB1, which transcriptionaly activates factors in the degradation pathway. The broken arrows represent feedback loops, both positive (green) and negative (red).



From Rutkowski and Kaufman (2004) *Trends Cell Biol* **14**(1) 20-28

Table 3. Select ERAD substrates of medical relevance

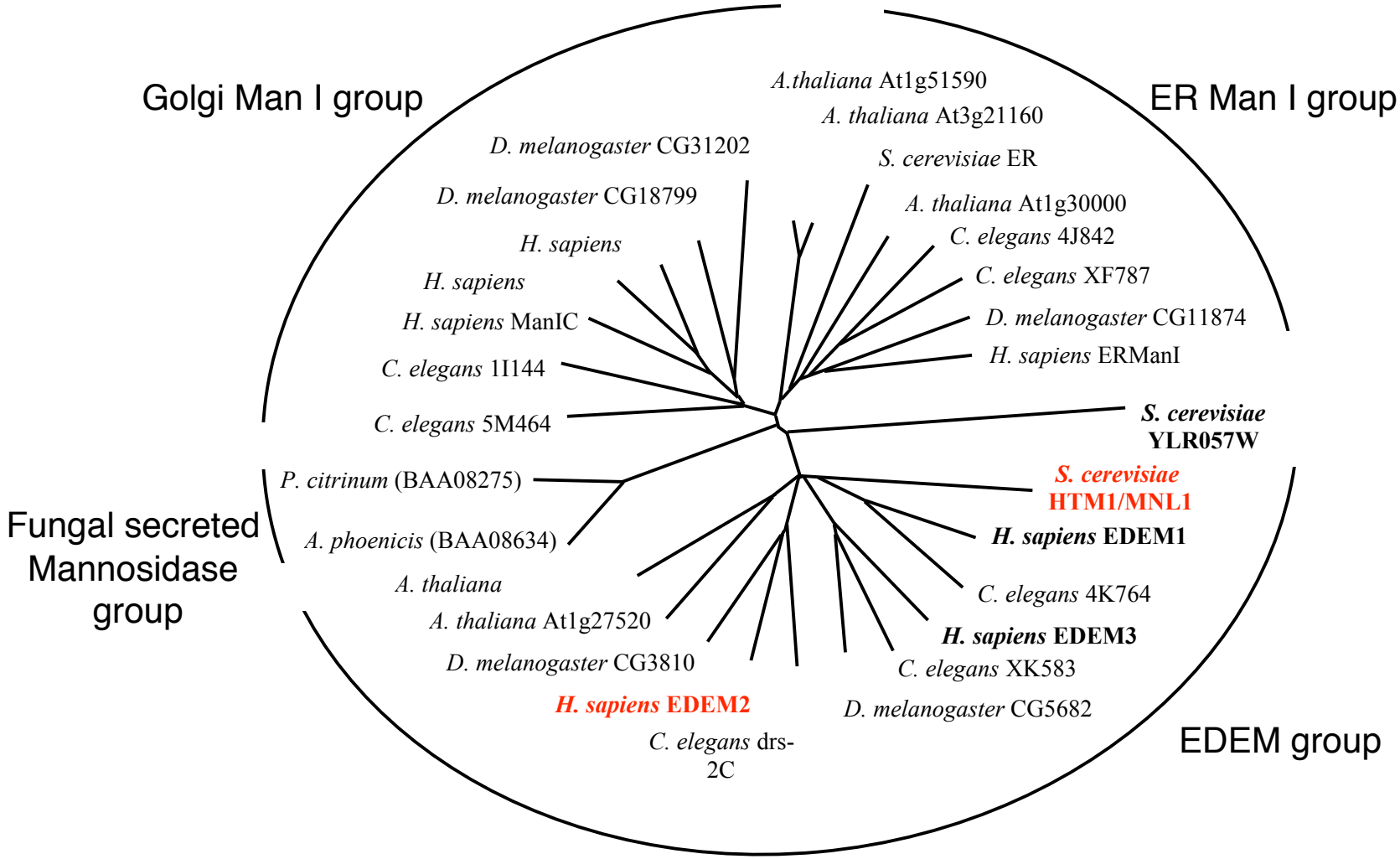
Select ERAD substrates of medical relevance

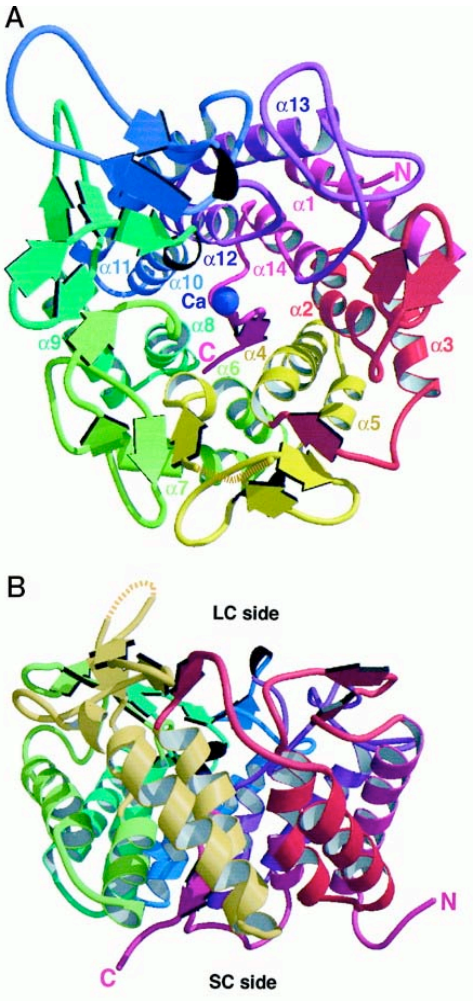
ERAD substrate	Associated disease
α1-protease inhibitor	Emphysema and liver disease
Aquaporin-2	Nephrogenic diabetes insipidus
Beta-hexosaminidase	Tay-Sachs disease
CD4	AIDS
Collagen	Osteogenesis Imperfecta
Connexin	Charcot-Marie-tooth disease
ΔF508 CFTR	Cystic fibrosis
Fibrinogen	Familial hypofibrinogenemia
HMG-CoA reductase	Heart disease
Insulin receptor mutants	Type A insulin resistance
LDL receptor class II mutants	Hypercholesterolemia
MHC class I; HEF	Infantile (CMV) hepatitis
	Hereditary hemochromatosis
MPO Y173C	Myeloperoxidase deficiency
Pael-R	Parkinson's Disease
Pro-PTHrP	Hypercalcemia assoc. malignancy
Rhodopsin	AD retinitis pigmentosa
Thyroglobulin	ongential hypothyroid
Tyrosinase	Amelanotic melanomas
Wilson protein	Wilson disease

From McCracken, A. and Brodsky, J. (2003) BioEssays 25, 868-877

Figure 9. Dendrogram for selected members of the Class 1 mannosidase family displayed as a radial unrooted tree. Where complete genome sequences were present, all known members from human, *Drosophila*, *C. elegans*, *Arabidopsis*, and selected fungal sources were chosen. The sequences are divided into four groups, ER mannosidase I, Golgi mannosidase I, fungal secreted mannosidase, and EDEM. Sequences were edited to compare the conserved □-mannosidase catalytic core domain of the family. The GenPept numbers are noted in parentheses.

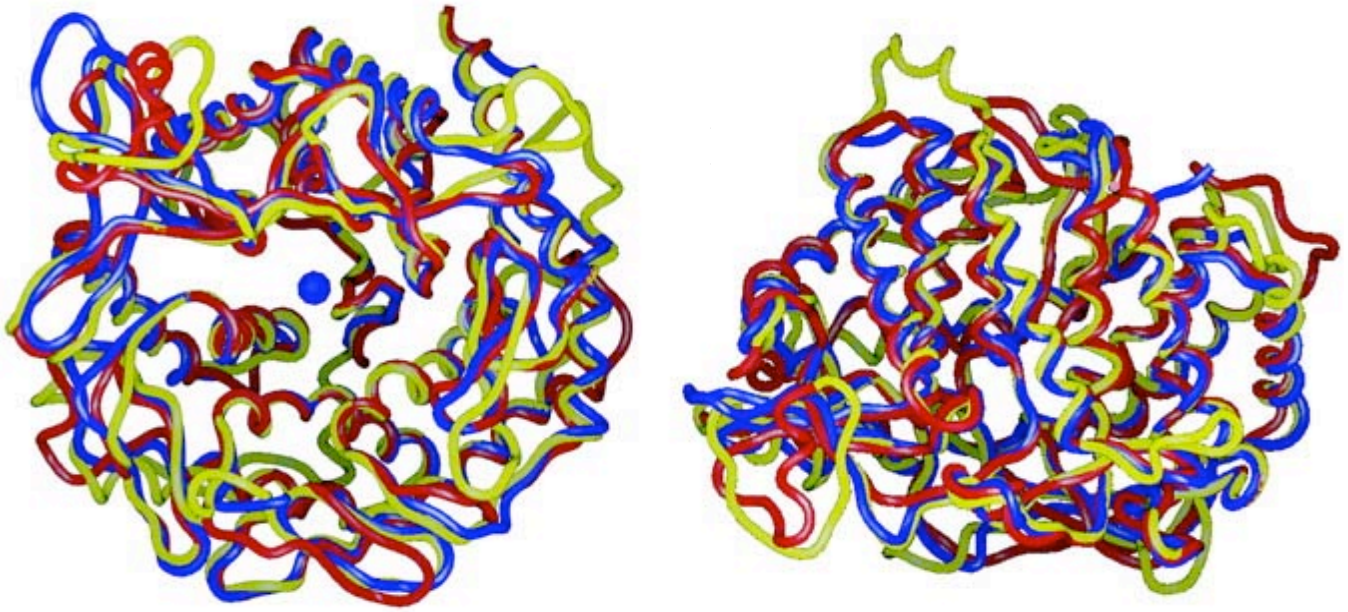
Glycosylhydrolase Family 47





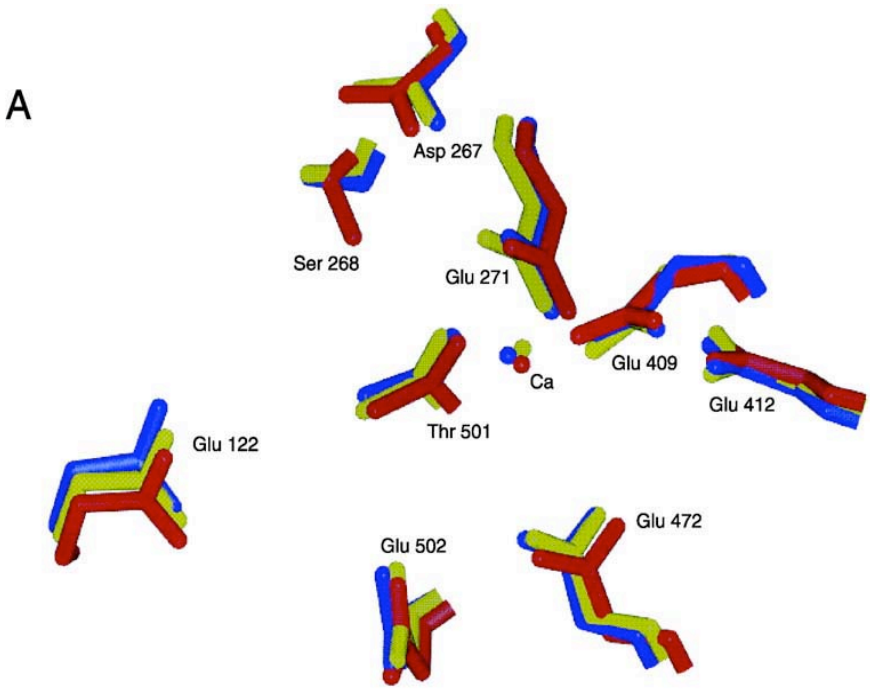
From Vallee et al. (2000) J Biol Chem. **275** (52), 41287-98

Figure 10. Structure of human ER mannosidase I. (Panel A) A ribbon representation of human ER mannosidase I viewed down the $(\beta\alpha)_7$-barrel axis from the long connection (*LC side*). The color scheme pairs the antiparallel β helices and is the same for both panels. (Panel B) Structure rotated 90°. The calcium ion (Ca) is represented as a dark blue sphere. *SC side*, short connection.



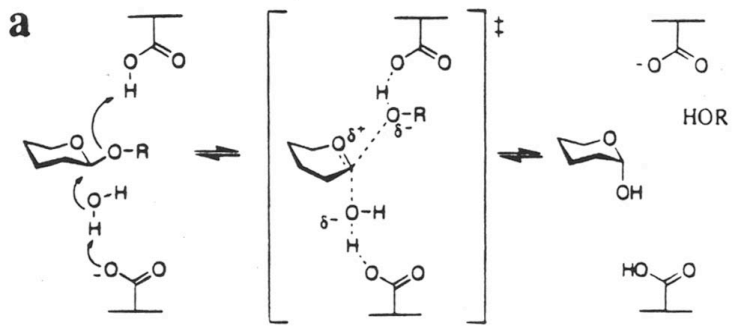
From Lobsanov et al. (2002) J Biol Chem. **277,** 5620-30

Figure 11. Comparison of fungal, yeast, and human □1,2-mannosidases. Overlay of the carbon backbone of *P. citrinum* □1,2-mannosidase (red), *S. cerevisiae* ER mannosidase I (yellow), and human ER mannosidase I (blue). (A) The structure is viewed down the $(\square\square)_7$-barrel axis from the long connection side, and (B) at 90° to the first orientation.



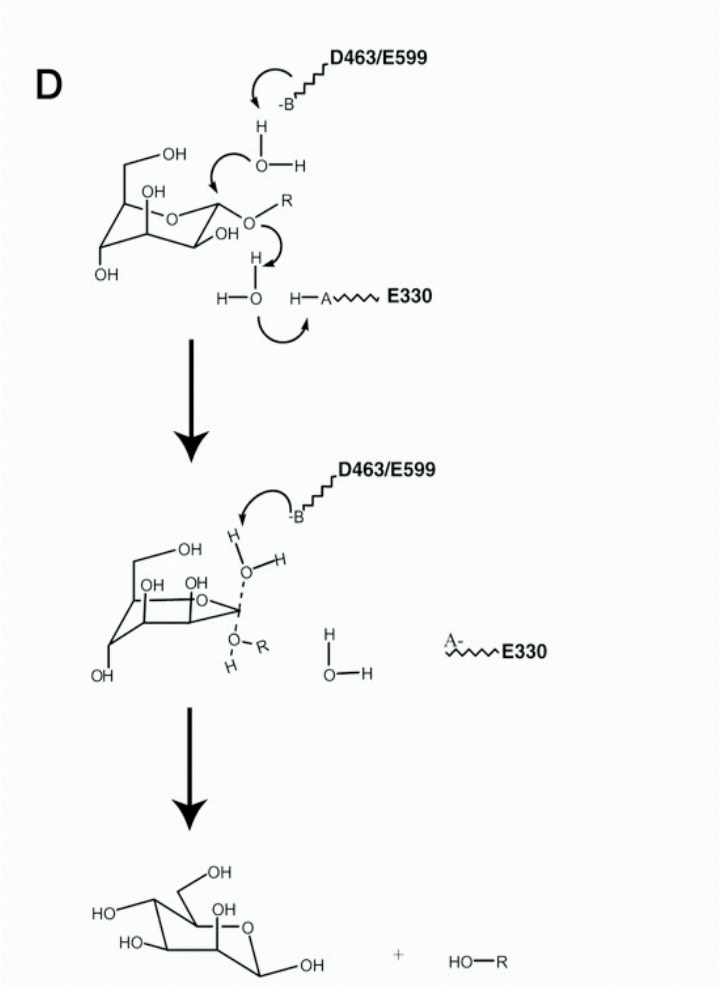
From Lobsanov et al. (2002) J Biol Chem. **277**, 5620-30

Figure 12. Structurally conserved catalytic site. Overlay of the conserved residues and calcium at the active site of *P. citrinum* □1,2-mannosidase (red), *S. cerevisiae* ER mannosidase I (yellow), and human ER mannosidase I (blue). For clarity, only a subset of the 11 conserved residues are shown.



From Zechel and Withers (2000) Acc. Chem. Res. **33**, 11-18

Figure 13. General mechanism of action for inverting glycosylhydrolases. Glycosylhydrolases hydrolyze the glycosidic bond by either retention or inversion of the anomeric configuration at the site of hydrolysis. Family 47 glycosylhydrolases are inverting glycosidases. The classic model contains two carboxyl groups placed ~10 Å apart and occurs via a direct displacement of the leaving group by water.



From Vallee et al. (2000) J Biol Chem **275** (52), 41287-41298

Figure 14. Proposed catalytic mechanism for human ER mannosidase I. D463 or E599 are the most likely candidates for the catalytic base. The catalytic base extracts a proton from water, which then attacks the C-1 atom. E330 acts as the catalytic acid, which also utilizes a water molecule in the reaction (numbering corresponds to human ER mannosidase I).

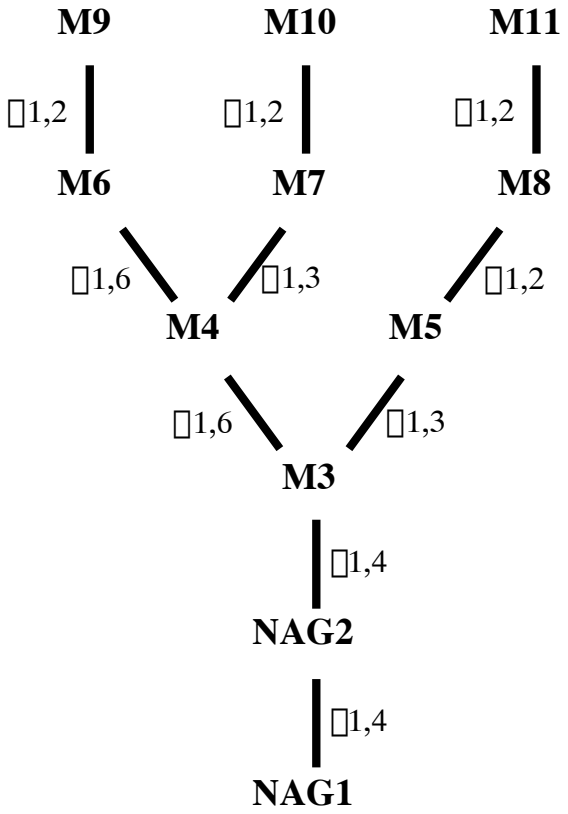


Figure 15. $Man_9GlcNAc_2$. Representation of the $Man_9GlcNAc_2$ oligosaccharide showing the linkages and numbering of the saccharide units. Mannose (M), GlcNAc (NAG).

CHAPTER 3

MATERIALS AND METHODS

I. Materials

Restriction enzymes were purchased from New England Biolabs, Roche, or Promega. *Taq* polymerase was from Perkin Elmer and *Pfu* polymerase was from Stratagene. The Sephaglas DNA band purification system was from Pharmacia. The QIAquick columns, Ni^{+2}-NTA resin, and the pQE32 vector were from Qiagen. The Benchmark pre-stained protein ladder and the pCR2.1, pPCR-Blunt, pCDNA3.1, pPICZ, and pPICZ vectors were from Invitrogen. The pBluescript SK^- vector was from Stratagene. The pEAK vector was from Edge Biosystems . The p416 vector was a gift of Zarmik Moqtaderi (Harvard Medical School, MA). Zeta probe membrane was from Bio-Rad. PVDF membrane was from Millipore. Anti-myc antibody was from Invitrogen. Monoclonal anti-mouse HA antibody clone 16B12 was from Covance. Goat anti-human 1-antitrypsin antibody and HRP-conjugated anti-goat antibody were from ICN. Mouse monoclonal CPY antibody 10A5-B5 was from Molecular Probes. Rabbit polyclonal CPY antibody was from Research Diagnostics, Inc. Rabbit -CPY antibody was a gift from Markus Aebi. Rabbit anti-yeast actin antibody, Y-act 486, was a gift of Dr. Richard Meagher (University of Georgia, Athens, GA). Rabbit antibody to Kar2p was a gift of Randy Schekman (UC Berkeley, CA). Calnexin antibody was from Stressgen. Calreticulin antibody was from ABR (Affinity BioReagents). Cdc2 antibody was from Oncogene. Alkaline phosphatase conjugated anti-mouse and anti-rabbit antibodies were from Promega. Alexa Fluor 488 and

Alexa Fluor 594 antibodies were from Molecular Probes. SuperSignal Western Blotting Kit was from Pierce**.** Immun-Star AP Substrate was from Bio-Rad. [^{35}S]Met/Cys was from ICN. DMEM was from Sigma. Complete EDTA free Protease Inhibitor cocktail and Endoglycosidase H were purchased from Roche. N-Glycanase was a gift of Dr. J. Michael Pierce (University of Georgia, Athens, GA). Kifunensine, Deoxymannojirimycin, and castanospermine were from Toronto Research Chemicals. α1-Antitrypsin cDNA was provided by Dr. Rick Sifers (Baylor College of Medicine, TX).

II. Methods

II-1. Isolation and characterization of Htm1p

cDNA Isolation of *HTM1*:

Genomic DNA was prepared from *S. cerevisiae* strain YPH499. Cells were grown in 40 ml YPD (1% yeast extract, 2% peptone, 2% glucose) overnight and centrifuged at 3,000 x g for 5 min. Cells were resuspended in 3 ml 1.1 M sorbitol, 0.1 M EDTA, pH 7.5 and then incubated with 0.1 ml of 2 mg/ml lyticase (Sigma) for 1h at 37ºC. The sample was then centrifuged (3,000 x g, 5 min) and resuspended in 5 ml 50 mM Tris, 20 mM EDTA (pH7.4). To the resuspended sample, 0.5ml of 10% SDS was added, the sample was incubated for 30 min at 65ºC, and then 1.5 ml of 5 M potassium acetate was added, followed by a 1 h incubation on ice. The sample was then centrifuged (12,000 x g, 10 min), and the supernatant was transferred to a new tube. Two volumes of ethanol were added, and the sample was mixed and centrifuged (3,000 x g, 15 min). The pellet was dried and resuspended in 10 mM Tris, 1 mM EDTA (pH 7.4). The sample

was then centrifuged (12,000 x g, 15 min). 150 μl of 1 mg/ml RNAse was added to the supernatant and incubated for 30 min at 37ºC. One volume isopropanol was added and the sample was gently mixed. The precipitate was removed and then resuspended in 10 mM Tris, 1 mM EDTA (pH 7.4). The DNA sample was adjusted to 0.2 M NaCl, and then two volumes absolute ethanol was added and incubated at –80ºC for several min. The sample was microfuged at maximum speed for 10 min, and then the DNA pellet was air dried and resuspended in 10 mM Tris, 1mM EDTA (pH 7.4). Using the sequence obtained from the Saccharomyces Genome Database, primers were designed to amplify the open reading frame (2391 bp):

5' primer: 5'-CCCCTCGAGGGAAATAGAATGGTTTGCTGCTTA-3'

3' primer: 5'-CCCGCGGCCGCTCATACAATAAATAAGTTGATG-3'.

The 5' primer and 3'primer respectively contain an Xho I site and a Not I site. Nine base pairs upstream of the initiation codon were included to retain the Kozak sequence. The PCR product was resolved on a 1% agarose gel containing 0.5μg/ml ethidium bromide, purified using the Sephaglas DNA band purification kit (Amersham Biosciences), and subcloned into the pCR-Blunt vector (Invitrogen). The sequence of the cloned PCR product, HTM1/pPCR-Blunt, was confirmed by the Molecular Genetics Instrumentation Facility, University of Georgia.

Expression and purification of a Htm1p peptide in *E. coli* :

The 5' primer was designed to insert a Sma I site after the first 60 bp of *HTM1*, and the 3'primer contained a Pst I site after the stop codon:

5' primer: 5'-CCTAGTCTAGACCCGGGGAAGACGATGCGTACTCATTCACTTCTAAA-3'

3' primer: 5'-GGCTGCAGTCATACAATAAATAAGTTGATGATG-3'.

Using genomic DNA sequence as the template, the PCR product was amplified and subcloned into the pCR-Blunt vector. To generate a truncated construct, a Kpn I digested fragment was excised and the construct was religated to itself, excising the COOH-portion of *HTM1*. A Sma I/Hind III fragment, which contains the NH_2-terminal sequence of *HTM1*, was then inserted into Sma I/Hind III digested pQE32 bacterial expression vector (HTM1/pQE32); this appended the vector encoded sequence, MRGS$(His)_6$GIRMRARYPG, to the NH_2-terminus of the HTM1 construct. This truncated construct (HTM1/pQE32) expressed amino acids 21 to 385. HTM1/pQE32 was transformed into the *E. coli* strain SG13009 (pREP4), and positive transformants were selected by antibiotic resistance to 25 μg/ml kanamycin and 100 μg/ml ampicillin.

For expression, 6 ml LB ampicillin (100μg/ml) kanamycin (25μg/ml) was inoculated with a single colony and grown 8 h at 37°C; this culture was used to inoculate 1 L of LB (Amp/Kan) which was grown for ~3.5 h and reached an OD_{600} of 0.8. One millileter of 1M IPTG was added to induce expression and the culture was incubated for 4 h more. The cells were harvested by centrifugation, and the cell pellet was frozen in dry ice/alcohol. The cells were then resuspended in 50 ml 50 mM NaH_2PO_4, 300 mM NaCl, pH 8. The suspension was sonicated and then centrifuged (12,000 x g, 25 min). The pellet was resuspended in 30 ml buffer A (6 M GuHCl, 0.1 M NaH_2PO_4, 10 mM Tris-Cl, 20 mM imidazole, pH 8) and recentrifuged. The supernatant was incubated with 3.5 ml packed Ni^{+2}-NTA resin for 4 h.

The Ni^{+2}-NTA column was washed with buffer A plus 1% Tween-20 and 10 mM imidazole at 0.5 ml/min until the OD_{280} trace became constant. The column was washed with 40 ml buffer B (6.5 M Urea, 0.1 M NaH_2PO_4, 10 mM Tris-Cl, pH 6.3), followed by 6 ml buffer B

with 100 mM imidazole (pH6.3), 6 ml buffer B, 12 ml buffer B with 1M NaCl, 15 ml buffer B, 75 ml buffer B with a pH gradient from pH 6.3 – 4, ending with 18 ml buffer B (pH 4). Samples from each fraction were run on SDS-PAGE. Relevant fractions were pooled (60 ml) and dialyzed (MWCO 6-8,000) overnight against 1.5 L 1 M Urea, 12.5 mM NaH_2PO4, 75 mM NaCl, pH 7, and then dialyzed twice against 1.5 L PBS. Precipitate was resuspended in PBS, 0.1% SDS.

Rabbit immunization was performed by the Animal Resources Facility, University of Georgia. The ~45 kDa purified Htm1p fragment (1 mg/ml) was provided for injection into New Zealand White rabbits (rabbit #330). Booster immunizations were given 21 days after the initial injection, with subsequent injections given at 3-6 weeks. Serum was tested on the purified peptide at dilutions ranging from 1:2,500 to 1:20,000; pre-immune serum was used for the negative control.

Expression of Htm1p in *Pichia pastoris:*

The coding region was transferred from HTM1/pPCR-Blunt into Xho I/Not I cut pPICZ B expression vector (HTM1/pPICZ). To generate an in-frame fusion with the vector-encoded COOH-terminal myc-His tag, HTM1/pPICZ was altered by site-directed mutagenesis using the QuikChange site-directed mutagenesis kit (Stratagene) to change the stop codon, TGA, to a Ser codon, TCG, and remove the Not I site.

primer 1: 5'-TTCTAGAAAGCTGGCGGCCGATACAATAAATAAGTTGATGATGG-3'

primer 2: 5'-CATCAACTTATTTATTGTATCGGCCGCCAGCTTTCTAG-3'.

The resulting construct (HTM1-myc-His/pPICZ) encodes a polypeptide containing 6 amino acids (SAASFL) between the coding region and the vector encoded COOH-terminal myc epitope and 6x His peptide.

An HA epitope was also added to the COOH-terminus of HTM1. The myc-2xHA-6xHis tag was removed from EDEM2-myc-HA-His/pEAK by digestion with Xba I, and then this tag was inserted into Xba I cut HTM1-myc-His/pPICZ, replacing the myc-His tag with a myc-HA-His tag (HTM1-myc-HA-His/pPICZ).

For the αHA-HTM1/pPICZ construct, the predicted signal peptide cleavage site of *HTM1* was replaced with the α-factor signal sequence and 2 x HA epitope. PCR primers were designed to amplify the α-factor signal sequence from the pPICZα A vector and introduce an Asc I site for insertion of the HA tag. The following restriction enzyme sites were designed into the primers: Hind III and Xba I sites were inserted into the 5' primer and Kpn I and Asc I sites were inserted into the 3' primer.

5' primer: 5'-AAGCTTCTAGACAACTAATTATTCGAAACGATGAG-3'

3' primer: 5'-GGTACCGGCCCTCAGCTTCAGCCTCTCTTTTCTCGAGA-3'.

The amplimer obtained using these primers was inserted into Kpn I/Hind III cut pPICZα. An Asc I/Sma I fragment from TCM-EDEM2/pEAK contained the HA tag, and this fragment along with *HTM1* (Sma I/Not I from NHis-HTM1/pEAK) was inserted into Asc I/Not I cut modified pPICZα expression vector

A truncated construct containing only the mannosidase homology domain was generated from the αHA-HTM1/pPICZ construct. The truncated-HTM1/p416 construct contained a stop codon introduced by site directed mutagenesis (see *S. cerevisiae* expression section), and this

segment was swapped into □-HA-HTM/pPICZ using BamHI and Sal I restriction sites, creating □HA-truncated-HTM1/pPICZ.

The pPICZ constructs were linearized by SacI, and 5 μg each was used for transformation of the X-33 and KM71H *Pichia* host strains by the EasyComp transformation kit; protocol described in Invitrogen manual. To screen transformants, colonies were inoculated into 100 ml BMGY and grown 1 to 2 days to an OD_{600} of 2-6. The cells were centrifuged (2800 x g, 10 min) and resuspended in 20 ml BMMY for induction. The cultures were grown for 3 to 4 days and every 24 h methanol was added to a concentration of 0.5%. The cells were collected by centrifugation (2800 x g, 10 min) and lysed using YPER-S solution (Pierce). The media was further centrifuged (28,000 x g, 30 min). The entire cell lysate and 5ml media were immunoprecipitated with antibodies to myc or HA.

Expression of Htm1p in mammalian cells:

To generate a full-length construct without the putative signal sequence or transmembrane sequence, HTM1/pPCR-Blunt was used as a template for PCR. First, the primers were designed to amplify a portion of the HTM1 coding region from amino acid 21 to 185, with Xba I and Sma I sites inserted at the 5' end of the coding region.
5' primer: 5'-CCTAGTCTAGACCCGGGGAAGACGATGCGTACTCATTCACTTCTAAA-3'
3' primer (A5): 5'-GGTAGACGTTAAATATGCAGGTAAGAGCCTATCTG-3'.
The resulting 535 bp amplimer was subcloned into the pPCR-Blunt vector and sequenced. Second, the 535 bp amplimer was used to replace the front end of HTM1/pPCR-Blunt construct using Xba I and Pfl MI. Third, this construct was inserted into Sma I/Not I cut pPROTA+ expression vector generating a construct of HTM1 without the putative signal sequence.

A second construct with a modified NH_2-terminus was also created. The putative signal sequence of HTM1 was replaced by the following sequence: Ig κ signal sequence (METDTLLLWVLLLVPGSTGDAA), followed by 9 x His tag, and a Sma I site at the 3' end of the sequence. The Ig κ signal and His tag had been previously inserted into pBSSK which placed a Hind III site upstream of the signal sequence and Sma I site at the 3' end. A Hind III/Sma I fragment containing the Ig k signal sequence and His tag sequence and a Sma I/Not I fragment from HTM1/pPROTA+ (containing *HTM1* residues 21 to 796) were both inserted into a Hind III/Not I cut pEAK 10 vector creating NHis-HTM1/pEAK.

Two COOH-tagged constructs were also made. HTM1-myc-His and HTM1-myc-HA-His were each excised from pPICZ by digestion with Xho I and Age I and ligated into pcDNA3.1+. HEK293 cells were transfected by the calcium phosphate method, described in EDEM2 section.

Expression of Htm1p in *S. cerevisiae*:

The full-length coding region was excised from HTM1/pCR-Blunt using the Xba I and Spe I sites encoded in the vector's polylinker and ligated into Xba I/Spe I cut p416 TEF vector (HTM1/p416). A COOH-terminal myc-His tagged version was created by digestion of HTM1-myc-His/pcDNA3.1 with Spe I and Pme I and ligation into a Spe I/Sma I cut p416 TEF vector (HTM1-myc-His/p416). The other COOH-tagged sequence, HTM1-myc-HA-His was inserted into Not I/Age I cut p416 TEF after it had been removed from the pcDNA3.1+ vector (HTM-myc-HA-His/p416).

S. cerevisiae cells were transformed by the lithium acetate method (161). Briefly, 100 ml YPD cultures were inoculated and grown to an OD_{600} of 1, then the cells were washed with water

and resuspended in 1.5 ml LiAc solution (100 mM LiAc, 1 mM EDTA, 10 mM Tris-Cl, pH7.5). The yeast suspension (200 μl) was added to 5μg DNA and 10 μg carrier DNA (keeping DNA volume less than 10% of total volume). A PEG solution (1.2 ml of 40% PEG3350, 100 mM LiAc, 1 mM EDTA, 10 mM Tris-Cl, pH 7.5) was added and the solution was incubated at 30ºC for 30 minand at 42ºC for 15 min. The cells were spun down, resuspended in TE buffer (1 mM EDTA, 10 mM Tris-Cl, pH7.5), and plated on selective media (uracil deficient SC plates).

For expression, 3 L YPD or SC (-ura) inoculated with HTM1-Myc-His/p416 or HTM1-myc-HA-His/p416 was harvested at an OD_{600} of 1. The cell pellet (12 ml) was washed in water, recentrifuged (2,000 x g, 5 min), and resuspended in 3 ml YPER-S solution (Pierce) per gram of cells. PMSF (1.5 mM final concentration) and 1x EDTA free Complete Protease Inhibitor cocktail were also added, and the cells were incubated on a rocking platform for 20 minat room temperature. The cells were centrifuged (12,000 x g, 20 min, 4ºC) and the supernatant was loaded on a 4 ml Ni^{+2}-NTA column. The column was washed with 40 ml wash buffer (0.5% Triton-X 100, 375 mM NaCl, 1 mM $CaCl_2$, 25 mM imidazole, 75 mM Na-MES, pH 7), followed by a 20 ml wash (150 mM NaCl, 1 mM $CaCl_2$, 50 mM imidazole, 75 mM Na-MES, pH 7), and eluted using a 35 ml imidazole gradient (50-250 mM) containing 150 mM NaCl, 1mM $CaCl_2$, 75 mM Na-MES, pH 7.

Expression of the Htm1p point mutations in *S. cerevisiae*:

Three conserved residues (Asp 279, Arg 407, and Thr 495) were changed to alanine by site directed mutagenesis of HTM1/pPCR-Blunt.

D279A primer1: 5'-GGAATCGGCGCCTCCATCGCCTCCCTTTTCGAATACGC-3'

D279A primer2: 5'-GCGTATTCGAAAAGGGAGGCGATGGAGGCGCCGATTCC-3'

R407A primer1: 5'-GAATGGTACCCATTAGCACCGGAATTTTTTGAGTC-3'

R407A primer2: 5'-GACTCAAAAAATTCCGGTGCTAATGGGTACCATTC-3'

T495A primer1: 5'-GTGACGTAATTTTCAGTGCAGAAGCCCATCCAATGTGG-3'

T495A primer2: 5'-CCACATTGGATGGGCTTCTGCACTGAAAATTACGTCAC-3'.

The mutated *htm1* sequences (D279A/pCR Blunt, R407A/ pCR Blunt, and T495A/ pCR Blunt) were swapped into Xba I/Spe I cut p416 TEF vector. *S. cerevisiae* was transformed by the LiAc method using 5 μg DNA and selected by growth on uracil deficient SC plates.

N-glycanase digest of Htm1p:

Glycosidase digests were carried out on Ni^{+2}-NTA protein, expressed from HTM1-Myc-His/p416 in *Saccharomyces*. The digest was done as described in the EDEM2 section. The samples were resolved by SDS-PAGE, immunoblotted with rabbit anti-HTM1 pQE antibody (1:1000) followed by the anti-rabbit–alkaline phosphatase-conjugated antibody (1:5000), and developed with Immun-star substrate.

Substrate assays and HPLC methods:

Oligosaccharide substrates were obtained, and assays were performed essentially as described previously (149). Pyridylamine-tagged $Man_9GlcNAc_2$, $Man_8GlcNAc_2$ (isomers A, B, and C), $Man_6GlcNAc_2$, or $Man_5GlcNAc_2$ were each incubated for 2 – 10 h at 37ºC with Ni^{+2}-NTA purified NHis-HTM1 in a buffer of 50 mM Na-MES, 1mM $CaCl_2$ (pH 7). The digestion products were resolved on a Hypersil APS-2 NH_2 HPLC column.

Deletion of *HTM1*:

PCR based gene disruption was used to generate a knockout of *HTM1* in the *Saccharomyces strains* YPH499 and YPH501 (162). PCR primers were designed to amplify the *kanMX* sequence from the pFA6a-kanMX2 vector (162). The 5' primer contained 47 bp homologous to *HTM1* and 19 bp homologous to *kanMX* , while the 3' primer had 45 bp homologous to *HTM1* and 21 bp homologous to *kanMX:*

5' primer: 5'-ATGGTTTGCTGCTTATGGGTGCTTCTAGCATTGTTGCTTCATCTCGACA GCTGAAGCTTCGTACGC-3'

3' primer: 5'-TCATACAATAAATAAGTTGATGATGGGCGTGCATTCTAACATGAGCAT AGGCCACTAGTGGATCTG-3'.

The amplimer contained a *kanMX* selectable marker with flanking regions homologous to *HTM1*, and it was used to replace the coding region of *HTM1*, by deleting bp 48 to 2346.

Competent *S. cerevisiae* cells were prepared according to a published protocol (163). Competent cells (50 μl) were transformed by adding 5 μl carrier DNA (10 μg/ml sheared herring sperm DNA), vortexing briefly, and incubating 20 min at 30ºC. Then 300 μl 40% PEG3350/100 mM LiAc was added, the solution was briefly vortexed, and incubated for 20 min at 30ºC. The cells were heat shocked at 42ºC for 20 min, followed by centrifugation at 5000 rpm in a microfuge, and resuspended in 1 ml YPD for incubation at 30ºC for 2 h. Aliquots of the suspension were plated on YPD plates containing 200 mg/ml G418. Colonies that grew were restreaked on YPD G418 plates. PCR and a Southern blot were performed to confirm the deletion. For Southern blot analysis, genomic DNA (5 μg) from *S cerevisiae* strains YPH499, YPH501, □*htm1* (YPH499), and □*htm1* (YPH501) was digested with the indicated restriction enzymes creating the following DNA fragments: 5.1 kb and 361 bp Hind III fragments, 4.2 kb

Pvu II fragment, 5.1 kb Nhe I/Xba I fragment, and 3.8 kb Bst XI fragment. The digestion products were separated on 1% agarose gel and transferred to a Zeta probe nylon membrane by capillary blotting, described previously (151). Blots were prehybridized and hybridized at 42ºC, and washed at 65ºC, essentially as described previously (164). A full-length probe was generated by digestion of HTM1/pCR-Blunt with Xho I and Not I. The labeled probe was generated using ^{32}P-dCTP with the Ready-to-go labeling system (Pharmacia). Blots were visualized using a Phosphorimager (Molecular Dynamics).

For PCR verification, primers outside of the open reading frame, primers in the *HTM1* coding region (A5 and S2), and one primer in *kanMX* were used:

Primer upstream of *HTM1* coding region (HTM1 S9): 5'-TTCCGTTGTGCTTTAACGAC-3'

Primer downstream of *HTM1* coding region: 5'-CATCATTTATTACTGGTGCCA-3'

HTM1 Primer A5: 5'-GGTAGACGTTAAATATGCAGGTAAGAGCCTATCTG-3'

HTM1 Primer S2: 5'-GGATATCATCAGTTGAATCCATTAATGTTAGAAAA-3'

kanMX primer: 5'-CCTCGACATCATCTGCCC-3'.

A single colony of yeast was transferred into the bottom of a PCR tube, microwaved for 1 minute, and then placed on ice. PCR reaction mix (25 μl of 10 mM Tris (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 0.2 mM dNTP, 1 μM primers, 1-2 U Taq DNA polymerase). Two to three primers in various combinations were added to the PCR mix. The initial denaturation step was 120 sec, 94ºC. The sequence of denaturation (30 sec, 94ºC), annealing (30 sec, 50ºC), and elongation (90 sec, 72ºC) was repeated 30 times. The reactions were analyzed on an agarose gel.

Deletion of *YLR057w*:

PCR based gene disruption was used to generate a knockout of *YLR057w* (Figure 27) in the *S. cerevisiae* isogenic strains (YG618, Δ*htm1* (SS328), YG778, and Δ*htm1* Δ*mns1* (SS328)). PCR primers were designed to amplify the yeast selectable marker gene, *HIS3*, from pRS303 (165). The 5' primer and 3' primers each contained 40 bp homologous to *YLR057w* and 20 bp homologous to pRS303:

5' primer: 5'-CTACTCTATAAGCAAAACCTTATGTCCATAGCGCGATTAGGATTGT ACTGAGAGTGCACC-3'

3' primer: 5'-ATGTATGTGACGATTTTTCTAACGTTAACTTCATTTCCTGTGCGGTA TTTCACACCG-3'.

Using 2.5μg of the PCR product, *S. cerevisiae* cells were transformed by the LiAc method and colonies were selected from histidine deficient SC plates. The gene deletion was verified by PCR analysis:

Primer upstream of *YLR057w* coding region: 5'-CTGGAAGCTTTAAAATGCCC-3'

Primer downstream of *YLR057w* coding region: 5'-CACAACGACACTTAATCCC-3'

Sense YLR057w specific primer (S5): 5'-CTGGAAGCTTTAAAAATGCCC

Antisense YLR057w specific primer (A5): 5'-CTTGAGTTAATAACTGGCAACCA-3'

Sense pRS303 specific primer: 5'-ACGCAGTTGTCGAAC-3'

Antisense pRS303 specific primer: 5'-AATCCCGCAGTCTTC-3'

Deletion of *DER1*:

PCR based gene disruption was used to generate a knockout of *DER1* in the *S. cerevisiae* strain, YG618. PCR primers were designed to amplify the *kanMX* sequence from the pFA6a-

kanMX2 vector. The 5' primer contained 40 bp homologous to *DER1* and 19 bp homologous to pFA6a-kanMX2, while the 3' primer contained 40 bp homologous to DER1 and 22 bp homologous to pFA6a-kanMX2.

5' primer: 5'-ATGGATGCTGTAATACTGAATCTCTTAGGCGACATTCCTTCAGCT GAAGCTTCGTACGC-3'

3' primer: 5'-TTAGGGTGTTTCAGTGTTGCGGAACCAGTCGTACGGAGATGCAT AGGCCACTAGTGGATCTG-3'.

Using 3 μg of the PCR product, *S. cerevisiae* cells were transformed and selected by the method described for the disruption of *HTM1*.

CPY* accumulation:

The following isogenic *S. cerevisiae* strains were used for the experiment: YG618, Δ*htm1* (YG618), Δ*htm1* (SS328), YG778, Δ*mns1* (SS328), Δ*ylr057w* (YG618), Δ*ylr057w* (YG778), Δ*ylr057w* Δ*htm1* (YG618), Δ*ylr057w* Δ*mns1* Δ*htm1* (YG618), and Δ*der1* (YG618). The cells were grown in YPD or SC media to stationary phase. Cells were centrifuged and washed once with water. The OD_{600} was measured and $9x10^8$ cells per sample were harvested and boiled with 750 μl SDS-Sample buffer. Samples were subject to SDS-PAGE and Western blot analysis. The blots were probed with the monoclonal or polyclonal anti-CPY antibody at a 1:2000 dilution, followed by the alkaline phosphatase conjugated anti-mouse antibody (1:5000) or HRP-conjugated-anti-rabbit antibody (1:20,000). The proteins were visualized with the ImmunStar chemiluminescence detection substrate or Supersignal chemiluminescence detection kit and developed on X-ray film. The next day the blot was probed

with an anti-actin antibody, followed by HRP-conjugated anti-rabbit antibody and developed with Supersignal chemiluminescence detection kit .

In a similar type of experiment, recombinant Htm1p was expressed in *S. cerevisiae* strain, Δ*htm1* CPY*, using the following constructs: HTM1/p416, HTM1-myc-His/p416, HTM1-myc-HA-His/p416, D279A/p416, R407A/p416, and T495A/p416.

Label and immunoprecipitation of recombinant Htm1p from *S. cerevisiae*:

Saccharomyces [YCJ1, W303CQ, W3031C, Δ*der1* (YG618)] transformed with HTM1-Myc-HA-His/p416 were grown to an OD_{600} of 0.5 to 2 in SC (-ura). $8x10^7$ cells per sample was harvested, resuspended in fresh media, and incubated for 1 h. The cells were resuspended in 2ml SC (-met/-cys), incubated for 30 min, and then 250μCi [^{35}S]-Met/Cys was added. The cells were labeled for 30 minto 3 h and then washed in PBS. Cells were lysed in YPER-S as described or lysed with glass beads. 2% CHAPS, HBS (200 mM NaCl, 50 mM Na-Hepes, pH 7.6), 1 mM PMSF, 1x EDTA free protease inhibitor cocktail and acid washed glass beads (400-600 μm) were added to the cells and the sample was vortexed for 30 second pulses for a total of 2 min. The samples were centrifuged (17,000 x g, 15 min, 4ºC) and immunoprecipitated with antibody to HA, Kar2p, or CPY*, as described in EDEM2 section.

II-2. Isolation and characterization of EDEM2

cDNA isolation of EDEM2:

Mouse Golgi mannosidase IA (GenBank™ U042999) was used to query human EST sequences using the BLAST algorithm in the NCBI database. Several partial, novel sequences

were identified. Based on sequence comparison to other family 47 hydrolases, the sequence of EDEM2 was not completed at the 5' end. 5'-RACE was used to isolate the remainder of the coding region, using a cDNA preparation from human spleen mRNAs the template. Primers for the RACE were based on the sequence from the human EST R67182.

Primer matched with AP1 (A3): 5'-GATTGTTGCAAATCCGCACTCCACCT-3'

Primer matched with AP2 (A1): 5'-TGTCGTAGGCGTGGTAGAACATGG-3'

The first amplification employed 5 μg of ligation-anchored cDNA, and a gene-specific primer (A3) was matched with the adaptor primer complementary to the ligation-anchored adaptor (primer AP-1; CLONTECH). The primary amplification conditions employed primers at a concentration of 2.5 μM in a reaction containing components of the Expand Long PCR Kit (Roche Molecular Biochemicals) essentially as described by the manufacturer, in a thermal cycler programmed for a preincubation at 94°C, (1 min) followed by a temperature step cycle of 94°C (30 s) and 72°C (2.5 min) for 5 cycles followed by 94°C, (30 s) and 70°C (2.5 min) for 5 cycles, followed by a temperature step cycle of 94°C (30 s) and 68°C (2.5 min) for 25 cycles. The 1.1 kb product of the first round of PCR was diluted 1:10 in water, and 5 μl was used in a secondary round of PCR using a nested adaptor primer (AP-2, CLONTECH) and an EST-specific primer (A1). The conditions for the secondary round of PCR were identical to the first round. The resulting 270 bp RACE product was isolated, subcloned into the pCR2.1 cloning vector, and sequenced.

To generate a construct without the putative signal sequence or transmembrane sequence, human pancreas or spleen cDNA sources were used as templates for PCR. The primers were designed to amplify the EDEM2 coding region from amino acid 18 to the stop codon, with a Sma I site at the 5' end of the coding region.

5' primer: 5'-CCCGGGCAGCACCATGGTGCGCCAGGTCCCGACGG-3'

3' primer: 5'-TTAATTATTTATGAGGAGTCTAGGAAAACCTGTCC-3'.

The resulting 1.7kb amplimer was subcloned into the pCR2.1 vector and sequenced (-ss EDEM2/pCR2.1). Insertion into this vector placed a Spe I site downstream of the stop codon. The full-length coding sequence of EDEM2 was generated by inserting the Sfc I/Hind III cut 1.7 kb amplimer (-ssEDEM2) and Xho I/Sfc I cut 270 bp RACE product into Xho I/Hind III cut pBSSK.

Expression and purification of a portion of EDEM2 in *E. coli*:

Primers were designed from the R67182 EST sequence. The 5' primer contained a Sma I site, and the 3'primer contained a Sal I site after the stop codon:

5' primer: 5'-CCCGGGATCCCGCTCGGCCTCCTGTGCGCGCTGCTG-3'

3' primer: 5'-GTCGACGGAGTCTAGGAAAACCTGTCCCAGTAATGC-3'.

The PCR product was amplified from the EST and subcloned into the pCR-Blunt vector where a Hind III site was added to the 3' end. A Kpn I fragment was then excised from the vector, cutting off the last ~500 amino acids of EDEM2. This construct was inserted into Sma I/Hind III cut pQE32, where a 6x His tag was appended to the NH_2-terminus of EDEM2 (pQE32 described in HTM1 section).

EDEM2/pQE32 was transformed into *E. coli* SG13009 strain. LB media (500 ml) with 100 μg/ml ampicillin and 25 μg/ml kanamycin was inoculated and grown to an OD_{600} of 0.63. The induction, expression, and lysis follow the same method as described for HTM1/pQE32. The recombinant protein was purified from the cell lysate by Ni^{+2}-NTA chromatography. A 2 ml column was used to bind the recombinant protein, followed by washing at 0.5 ml/minute with

7.5 ml 6 M GuHCl, 0.1 M NaH_2PO_4, 10 mM Tris-Cl, 20 mM imidazole (pH 6.3). The buffer was changed to 50 ml Buffer C (8 M urea, 0.1 M NaH_2PO_4, 10 mM Tris-Cl, pH 6.3) plus 30 mM imidazole and 1% Tween 20. The imidazole concentration was increased to 100 mM, in Buffer C (14 ml). The next step included a pH gradient of pH 6.3 to 4, using Buffer C (60 ml). The column was then washed Buffer C, pH 4 (21 ml). Peptide eluted during the pH gradient was pooled and dialyzed (MWCO 6-8,000) in successive steps:

1) 6 M Urea, 0.1 M NaH_2PO_4, 10 mM Tris-Cl, pH 7
2) 3 M Urea, 0.1 M NaH_2PO_4, 10 mM Tris-Cl, pH 7
3) 1 M Urea, 25 mM NaH_2PO_4, 2.5 mM Tris-Cl, pH 7
4) PBS.

The sample was concentrated by Centricon 10. Precipitation occurred during the concentration and the precipitate was resuspended in 3ml PBS, 0.1% SDS and incubated at 55ºC for 10 min. Immunization of rabbits has been described previously (Rabbit 310).

MAP peptide:

Two multiple antigen peptides (MAP) were synthesized by Molecular Genetics Instrumentation Facility (University of Georgia). The 506 MAP antibody was generated against a peptide from the C-terminus of EDEM2 (FSPENHDQAREKPAKQK). The NH_2-506 MAP antibody was generated against a peptide from the NH_2-terminus of EDEM2 (GAPGDGSAPDPAHYRERVK). Immunization of rabbits is the same as described for HTM1/pQE32 expressed protein.

CPY* complementation assay in *S. cerevisiae*:

Full-length EDEM2 was inserted into a yeast expression vector. EDEM2/pcDNA3.1 was swapped into Sma I/Hind III cut p416 vectors (series ADH, CYC, GPD, and TEF). Transformation of *S. cerevisiae* Δ*htm1* CPY* yeast and CPY* complementation experiments were done as described in HTM1 section.

Full-length EDEM2 constructs:

Full-length EDEM2 was inserted into Xho I/Hind III cut pCDNA3.1. A full-length EDEM2 construct with a COOH-tag was also made. To generate an in-frame fusion with the vector-encoded COOH-terminal myc epitope tag, full-length EDEM2 in pcDNA3.1/myc-His vector was altered by site-directed mutagenesis using the QuikChange site-directed mutagenesis kit to change the stop codon, UAA, to a Leu codon, CTA.

primer 1: 5'-TGAGTTTTTGTTCTAGAAAGCTTAGTGAGGAGTCTAGGAAAACC-3'

primer 2: 5'-TTCCTAGACTCCTCACTAAGCTTTCTAGAACAAAAACTCATC-3'

The resulting construct (EDEM2-myc-His/pcDNA3.1) encodes a polypeptide containing 4 amino acids (LSFL) between the coding region and the vector encoded COOH-terminal Myc epitope and 6x His peptide.

A 2x HA epitope (YPYDVPDYAAYPYDVPDYA) was also added to the COOH-terminus of the EDEM2-myc-His construct at a Sal I restriction site between the Myc epitope and His tags. Sense and antisense oligonucleotides containing a HA epitope sequence flanked by Sal I sites were synthesized, phosphorylated with T4 DNA kinase, mixed at 95ºC, and allowed to cool gradually.

oligonucleotide 1: 5'-TCGACTACCCCTACGACGTCCCCGACTACGCCGCGTACCCCTA

CGACGTCCCCGACTACGCCG -3'

Oligonucleotide 2: 5'-TCGACGGCGTAGTCGGGGACGTCGTAGGGGTACGCGGCGTAG TCGGGGACGTCGTAGGGGTAG-3'.

The annealed oligonucleotides were ligated with the Sal I-cut EDEM2-myc-His/BSSK vector and sequenced to confirm the proper insertion of the HA tag (EDEM2-myc-HA-His/BSSK). EDEM2 was inserted into Not I cut pEAK 10 vector (Edge Biosystems) (EDEM2-myc-HA-His/pEAK).

For a construct without the putative signal sequence or transmembrane sequence, EDEM2 was transferred from the -ss EDEM2/pCR2.1 construct into Sma I/Spe I digested mut5 N His BSSK vector (figure 22), whereby the putative signal sequence of EDEM2 was replaced by the following vector encoded sequence: the Ig signal sequence and 10 x His tag. Using restriction sites from the pBSSK polylinker, the coding region was then subcloned into Hind III/Not I cut pEAK 10 expression vector creating NHis-EDEM2/pEAK. A second construct without the putative signal sequence or transmembrane sequence was generated, replacing the putative signal sequence of EDEM2 by the following sequence: a 5' Hind III restriction site followed by the 5'UTR and NH_2-terminal signal sequence of *T. cruzi* lysosomal mannosidase (TCM) (GenBank™ AF077741), a 10x His tag, and an Asc I site at the 3' end of the sequence. An Asc I/Not I fragment containing the HA tag and EDEM2 (residues 18-579, see following section) was inserted into an Asc I/Not I cut TCM-His-pEAK 10 vector (Moremen lab, unpublished) creating TCM-EDEM2/pEAK.

A 2x HA epitope tag was added onto the NH_2-terminus of EDEM2. Sense and antisense oligonucleotides containing an HA epitope sequence flanked by Sal I sites were synthesized (Molecular Genetics Instrumentation facility, University of Georgia), phosphorylated with T4

DNA kinase, mixed at 95°C, and allowed to cool gradually.

Oligonucleotide 1: 5'-CGCGCCTGTACCCCTACGACGTCCCCGACTACGCCGCGTACC CCTACGACGTCCCCGACTACGCCC-3'

Oligonucleotide 2: 5'-CCGGGGGCGTAGTCGGGGACGTCGTAGGGGTACGCGGCGTA GTCGGGGACGTCGTAGGGGTACAGG-3'.

The annealed oligosaccharide was ligated with Asc I/Xma I cut Nhis-EDEM2/AscI BSSK (figure), creating HA-EDEM2.

Truncated-EDEM2 construct:

The EDEM2 coding region was altered using the site directed mutagenesis kit to introduce a stop codon and Not I site after amino acid 492, 10 amino acids after the end of the mannosidase homology domain.

Primer1: 5'-GCCAGAGGCTGAAGTGAGCGGCCGCGAAGAGCAGTGGGAG-3'

Primer2: 5'-CTCCCACTGCTCTTCGCGGCCGCTCACTTCAGCCTCTGGC-3'.

A restriction fragment (EcoRV/Hind III) containing the mutated region was swapped into the TCM-EDEM2/pEAK construct to create a construct that expressed the TCM-His-HA tag and the mannosidase homology domain sequence but not the COOH-terminal extension beyond the homology region (truncated-EDEM2/pEAK). This construct was transfected into HEK 293 cells and expressed as described for the full-length construct.

α_1-antitrypsin construct:

The nonglycosylated form of PI Z, PIZ(NOG), was generated by site directed mutagenesis of PI Z variant of α1-antitrypsin in the pBSSK vector. Using the Quick-change site

directed mutagenesis kit, the 3 N-linked glycosylation sites were removed by altering NX(S/T) sites to QX(S/T) using the following primers:

Site 1 primer: 5'-GCACACCAGTCCCAGAGCACCAATATCTTC-3'

Site 1 primer complement: 5'-GAAGATATTGGTGCTCTGGGACTGGTGTGC-3'

Site 2 primer: 5'-GAGGGCCTGAATTTCCAGCTCACGGAGATTCCG-3'

Site 2 primer complement: 5'-CGGAATCTCCGTGAGCTGGAAATTCAGGCCCTC-3'

Site 3 primer: 5'-GCTGATGAAATACCTAGGCCAGGCCACCGCCATCTTC-3'

Site 3 primer complement: 5'-GAAGATGGCGGTGGCCTGGCCTAGGTATTTCATCAGC-3'.

For easier identification of the mutants, Site 1 primers contained a Bsp I restriction site, and site 3 primers contained an Avr II site. Nonglycosylated □1-antitrypsin, PI Z(NOG), was subcloned into pcDNA3.1/Zeo

Generation of stably and transiently transfected HEK293 cells:

For stable expression, HEK293 cells were grown to 40% confluency in T-75 flasks (Corning) and transfected by the calcium-phosphate method (166). Briefly, EDEM2 DNA (15 μg) was mixed with calcium phosphate and added to the cell monolayer in fresh DMEM/10% FCS (Sigma). Cells were incubated for 6 h at 37ºC. The medium was then replaced, and cells were grown an additional 24 h followed by selection with 0.5 μg/ml puromycin. During subsequent passages, the puromycin was increased to a final concentration of 2 μg/ml puromycin. Mock transfected HEK293 cells were used as a negative control for puromycin selection.

N_{HK} (167), PI Z (168) and PI Z(NOG) variants of □1-antitrypsin (GenBank™ NM 000295) were expressed from the pcDNA3.1/Zeo vector (Invitrogen). For stable expression, PI Z/pcDNA3.1 or N_{HK}/pcDNA3.1 was transfected into HEK293 cells in the same manner as the pEAK vector constructs and selected with 200 μg/ml zeocin.

For transient transfection, cells were grown to 70% confluency in 100 mm dishes prior to transfection by the Lipofectamine 2000 method (Invitrogen) according to the method of Wu et al (86). Each transfection cocktail consisted of 30 μg DNA and included combinations of PI Z/pCDNA3.1, N_{HK}/pCDNA3.1, pSV-□-galactosidase Control Vector (Promega), PI Z(NOG)/pCDNA3.1, TCM-EDEM2/pEAK, EDEM2-Myc-His/pEAK, and truncated-EDEM2/pEAK. Cells were collected forty-eight hours post-transfection, mixed, and subcultured at identical cell densities in either 60 mm or 100 mm dishes to eliminate differences in transfection efficiency.

Expression in suspension culture and purification by Ni^{+2}-NTA:

Suspension cultures were grown at 37ºC with 8% CO_2 using stirred bioreactor flasks. For purification of EDEM2, stably transfected pEAK cells were grown in 1L suspension culture ($1x10^6$ cells/ml) using SFM II media (Gibco) supplemented with 0.2μg/ml puromycin. The cells were harvested and washed in 50 mM Na-MES (pH 7). Cells were resuspended in lysis buffer (50 mM Na-MES, 375 mM NaCl, 1% TritonX-100, 20 mM imidazole, 1 mM $CaCl_2$, pH 6.3) supplemented with 1 mM PMSF (Sigma) and 1x EDTA free Complete Protease Inhibitor cocktail (Roche) . Cells were lysed in a glass-teflon homogenizer for 1 minute. The lysate was centrifuged at 31,000 x g for 15 min. The supernatant was loaded onto a 4 ml column of Ni^{+2}-NTA resin (Qiagen), and the column was washed with 60 ml of lysis buffer, followed by elution

using a 100 ml imidazole gradient (50-400 mM) containing 100 mM Na-MES, 150 mM NaCl, 1 mM $CaCl_2$ (pH 6.3). The samples were resolved by SDS-PAGE and immunoblotted with mouse anti-myc antibody (1:5000) or anti-HA antibody (1:5000) followed by the goat anti-mouse–alkaline phosphatase-conjugated antibody (1:5000). The relevant fractions were pooled and dialyzed against PBS.

Gel filtration of EDEM2:

EDEM2 (TCM-EDEM2 and NHis-EDEM2) purified by Ni^{+2}-NTA was concentrated by ultrafiltration on a YM-50 membrane (Amicon). A 250 μl sample of concentrated protein was loaded onto a 46 ml Superdex-200 column. The column was equilibrated in a buffer containing 50 mM Na-MES (pH 7), 200 mM NaCl, 1 mM $CaCl_2$, 250 mM NDSB, and the gel filtration was run at a flow rate of 0.25 ml/min. Fractions (1ml) were collected throughout the run. The column was calibrated with the standards thyroglobulin, ferritin, catalase, aldolase (Pharmacia), and BSA (Pierce).

N-glycanase and endoglycosidase H digest of EDEM2:

Glycosidase digests were carried out on Ni^{+2}-NTA purified protein (NHis-EDEM2 and EDEM2-myc-His). EDEM2 was heated at 90ºC for 5 minin SDS (0.25%) and □-ME (0.1M). The solution was cooled to room temperature and a 10-fold excess of NP-40 (2.5%) was added, and followed by a 14 h incubation at 37ºC with ~2 U N-glycanase (gift from Dr. Michael Pierce, University of Georgia) or 4 mU Endoglycosidase H (Roche). A mock digest was also performed. The samples were resolved by SDS-PAGE and immunoblotted with mouse anti-myc antibody (1:5000) followed by the goat anti-mouse–alkaline phosphatase-conjugated antibody

(1:5000) (Promega) or immunoblotted with rabbit anti-MAP antibody (1:2000), followed by the anti-rabbit–HRP antibody (1:20,000).

Substrate assays and HPLC methods:

Oligosaccharide substrates were obtained, and assays were performed essentially as described previously (149). Pyridylamine-tagged $Man_9GlcNAc_2$ (1000 pmol), $Man_8GlcNAc_2$ isomer B (1000 pmol), $Man_8GlcNAc_2$ isomers A and C (100 pmol), $Man_6GlcNAc_2$ (1000 pmol), or $Man_5GlcNAc_2$ (1000 pmol) were each incubated for 2 – 10 h at 37ºC with Ni^{+2}-NTA purified EDEM2 (NHis-EDEM2 and EDEM2-myc-His) in a buffer of 50 mM Na-MES, 1mM $CaCl_2$ (pH 7). As a positive control, recombinant human ER Mannosidase I was added into a sample of EDEM2. The digestion products were resolved on a Hypersil APS-2 NH_2 HPLC column.

Immunofluorescence localization of EDEM2 in mammalian cells:

Stably transfected HEK293 cells were grown in DMEM/10% FCS on eight-well chamber slides (Lab-Tek, Nalge Nunc) coated with 0.01% poly-L-lysine (Sigma). Cycloheximide (20 μg/ml) was added to selected chambers for 1, 3, or 6 hours. Slides were washed in PBS and fixed with 3.5% formaldehyde in 100 mM potassium phosphate (pH 7) for 15 min at 37ºC. Following fixation, the cells were washed with PBS and permeabilized by incubation with 0.2% saponin and 10% FCS in PBS (buffer A) for 15 minat 37ºC. Primary antibodies (mouse anti-myc monoclonal antibody (Invitrogen), 1:1000; rabbit anti-calreticulin polyclonal antibody (Affinity Bioreagents), 1:1000; mouse anti-HA monoclonal antibody (Covance), 1:1000; rabbit anti-human Golgi Man II polyclonal antibody (generated in Moremen lab to recombinant human Golgi Man II), 1:1000; rabbit anti-cytosolic mannosidase polyclonal antibody (generated in

Moremen lab to recombinant rat cytosolic mannosidase), 1:1000; or rabbit anti-MEK-1 polyclonal antibody (Santa Cruz Biotech.), 1:1000) diluted in buffer A were incubated with the chamber slides for 1 h at room temperature, followed by several washes with PBS. Secondary antibody (Alexafluor anti-mouse 594 (Molecular Probes), 1:1000); and Alexafluor anti-rabbit 488 (Molecular Probes), 1:1000) incubations were also for 1 h at room temperature followed by several washes with PBS. The slides were then mounted in Permafluor mounting media (Immunon Thermo Shandon) and viewed with a Bio-Rad MRC-600 laser scanning confocal microscope or Leica DM IRE2 microscope.

Immunoprecipitation of recombinant EDEM2 from HEK293 cells:

Cells were incubated in Met/Cys free DMEM/5% dialyzed FCS for 30 minand labeled with [^{35}S]-Met/Cys (ICN) from 30 min to 2 h. Approximately 5x10^6 cells (1/3 of 150mm x 25 mm plate) were used for a single immunoprecipitation. The following drugs were added to selected samples: 10 μg/ml tunicamycin (Sigma) was added 6 h before the label and also during the label, castanospermine (1-2 mM) was added to cells for the indicated time, and DTT (2 mM) was added for the last 20 min of a 2 h label. Cells were harvested and collected by centrifugation and lysed by vortexing in a buffer containing 2% CHAPS, Hepes buffered saline (HBS), pH 7.6. Before immunoprecipitation, cell lysates were centrifuged at 22,000 x g for 20 min, followed by incubation with 60 μl Protein A-agarose (Repligen) for 1 h. The primary antibodies for the immunoprecipitation were prebound to a separate aliquot of Protein A-agarose resin, and before immunoprecipitation unbound antibody was removed by washing with PBS. Controls containing Protein A-agarose with non-specific antibody were included for each cell line. The precleared cell lysates were added to the Protein A-agarose-antibody complex and incubated at 4ºC

overnight. The resin was washed 3 times with 0.5% CHAPS, HBS (pH 7.6) and once in 50 mM Na-Hepes (pH 7) before elution by boiling in SDS sample buffer. Samples were run on SDS-PAGE. The gel was incubated in 50% methanol/10% acetic acid for 30 min, followed by an incubation in 7% acetic acid/7% methanol/1% glycerol for 30 min, and an incubation in Amplify (Amersham Pharmacia Biotech) for 30 min. The gel was visualized by autoradiography.

Metabolic radiolabeling and immunoprecipitation of α1-antitrypsin:

In collaboration with Dr. Richard Sifers, metabolic radiolabeling was preceded by a 30 min incubation at 37ºC in Met-free DMEM (ICN Biomedical Research Products) containing 10% fetal calf serum prior to a 15-min pulse with [^{35}S] Met (Easy Tag Express Mix, Perkin Elmer Life Sciences) as described (168). Monolayers were washed with Dulbecco's phosphate-buffered saline (GIBCO-BRL Life Sciences) to remove unincorporated radiolabel, and then chased for the desired period at 37ºC with Met-free starvation medium supplemented with a 10-fold excess of unlabeled methionine. Cells were lysed and collected by scraping in buffered Nonidet P-40 (NP-40) detergent (Calbiochem) at 4ºC. Media were collected at each time point and adjusted to 0.5% NP-50. For immunoprecipitation, all samples were cleared by centrifugation (3000 x g, 5 min, 4ºC) prior to receiving an excess of goat anti human α1-antitrypsin (immunoglobulin fraction, ICN Biomedical Research Products), and then incubated for 2 hr, with constant rotation, at 4ºC. Immunocomplexes were washed as previously described (168) prior to resolution by SDS-PAGE and detection by fluorographic enhancement of the vacuum-dried gels (95,168) . Quantitative analysis was performed by densitometric scanning (National Institutes of Health IMAGE).

CHAPTER 4

RESULTS

I. HTM1

Characteristics of the coding regions of *S. cerevisiae HTM1* and *YLR057w*:

In BLAST searches of the *Saccharomyces* Genome database using the coding regions of family 47 α-mannosidases as query sequences, *HTM1* (2,388 bp) and *YLR057w* (2,547 bp) were found. In the phylogenetic analysis, *HTM1* falls within the group we have classified as EDEM. *YLR057w* does not appear to be as closely related (Figure 9), and we have designated *YLR057w* as an EDEM-like sequence. We isolated the coding region of one of the homologs, *HTM1*. Using *S. cerevisiae* genomic DNA as a template, a PCR amplimer was obtained that was identical to the sequence in the database. The open reading frame of *HTM1* encodes 796 amino acids (M_r 91,245 Da) (Figure 16), and the open reading frame of *YLR057w* encodes 849 amino acids (M_r 96,996 Da) (Figure 16). *HTM1* and *YLR057w* have 5 and 1 consensus N-linked glycosylation sites, respectively. Hydropathy plots obtained using the Pepplot subroutine of the GCG program indicated the presence of hydrophobic sequences at the NH_2-terminus of *HTM1* (Figure 17), which suggested the presence of the signal peptide sequence that causes proteins to enter the secretory pathway; if not cleaved, this sequence has the potential to anchor the protein to the ER membrane. *YLR057w* also has a hydrophobic sequence at the NH_2-terminus, which is predicted to be a transmembrane domain (according to PSORT II (169) and tmap (170)).

Direct sequence comparison within the EDEM group and also comparison of EDEM members to the catalytic members of the Family 47 glycosylhydrolases indicated that the sequence similarity among these proteins is restricted to the regions that have been shown to contain the $(\square\square)_7$ barrel catalytic domain of the Class I mannosidases (156,157) (Figure 18). In contrast, outside of the catalytic or mannosidase-homology domain there is limited similarity within the Family 47 glycosylhydrolases and also limited similarity to other protein sequences in the NCBI sequence database. The sequence relationship shared by the EDEM homologs with other Family 47 glycosylhydrolases is shown in Figure 9. The sequence relationship between *HTM1* and *YLR057w* is quite low. In the mannosidase homology domain, it shares 12% identity with *YLR057w*. By comparison, sequence identity between *HTM1* and the human EDEM homologs ranges from 33-39% identity (Table 4).

The EDEM members contain an extended COOH-terminal domain, which follows the mannosidase-homology domain. Outside of the mannosidase-homology domain these sequences do not share sequence similarity (Figure 19). *HTM1* has only a few residues at the NH_2-terminus but has a long COOH-extension following the mannosidase homology domain. At the C-terminus of *HTM1*, a putative GPI signal peptide cleavage site is present (Figure 16). The other *S. cerevisiae* homolog, *YLR057w*, has long NH_2-terminal and COOH-terminal extensions (Table 5).

Generation of an anti-Htm1p antibody:

A fragment of Htm1p was expressed in *E. coli* to prepare an antibody for immunodetection of Htm1p. The sequence encoding *HTM1* was modified by PCR to introduce a Sma I site at the predicted signal cleavage site and cloned into the pQE32 vector so that the

recombinant protein would contain a 6 x Histidine tag at the NH_2-terminus. For greater protein expression levels, the protein was truncated at residue 385. Expression of this construct was induced with IPTG and produced a protein with an apparent molecular mass of 45 kDa (Figure 20). For purification, the cells were lysed in guanidinium-HCl, and Ni^{+2}-NTA resin was used to capture the His-tagged protein. During chromatography, the buffer was switched to a urea buffer, and elution of the protein from the column was accomplished by lowering the pH to 4. In dialysis from urea to PBS, the protein precipitated out of solution but was able to be resuspended in SDS (0.1%). After purification, the major band seen by SDS-PAGE was at 45 kDa, although faint, contaminating bands could be seen at 30 kDa, 66 kDa, 80 kDa, and 98 kDa (Figure 20). One purification run yielded ~4 mg of protein. In Western blots, the anti-HTM1 antibody generated against the *E. coli* expressed protein was able to detect down to 20 ng of the antigen at 1:20,000 dilution, but 1:2500 dilution gave optimal results.

Expression of Htm1p in *Pichia pastoris*:

In order to generate protein for characterization studies, constructs were prepared for large scale expression in *Pichia*. The first construct, HTM1-myc-His/pPICZ contained the full-length coding region along with 9 bp of 5' UTR and a myc-His tag at the COOH-terminus. In the □HA-HTM1/pPICZ construct, the putative NH_2-terminal signal sequence was replaced with the □-factor signal sequence from the pPICZ vector. This construct contained an HA epitope at the NH_2-terminus between the □-factor signal sequence and the *HTM1* coding region. A third construct was also generated. This construct is a truncated version of □HA-HTM1/pPICZ, corresponding to amino acids 21-552. The truncated construct does not contain the COOH-terminal extension beyond the mannosidase homology domain.

Transformants were generated in each of the *Pichia* host strains (GS115, KM71H, X33) by homologous recombination of the construct in the *Pichia* genome at the 5' region of the alcohol oxidase gene (AOX1). Transformants were selected by antibiotic selection to the vector-encoded zeocin resistance gene. Expression was induced by the methanol-inducible promoter of AOX1. No expression was observed using the COOH-tagged constructs or the truncated-HTM1 construct. For HA-HTM1/pPICZ, a protein with an apparent molecular mass of over 100 kDa was detected by immublotting with antibody to HA in both crude cell extracts and immunoprecipitates of cell extracts using HA antibody, but it was not detected in the media (Figure 21). As no protein was secreted and low levels of expression were observed, *Pichia* was not a useful expression system for Htm1p.

Expression of Htm1p in mammalian cells:

Mammalian cells were tested as a system for protein expression. For expression in HEK293 cells, the pEAK 10 episomal vector was used. In the construct, Nhis-HTM1/pEAK, the yeast putative signal sequence was replaced with the Ig signal sequence and 5'UTR (Figure 22), and a His tag was placed between the new signal sequence and *HTM1* coding region. Cells were stably transfected and grown in suspension culture. Protein was purified from cell lysates by Ni^{+2}-NTA chromatography (Figure 23), yielding a protein with an apparent molecular mass of 97 kDa. The levels of expression observed indicated mammalian cells could be used for the purpose of protein expression.

Expression of Htm1p in *S. cerevisiae:*

Htm1p was expressed in *S. cerevisiae* to generate protein for characterization studies and also for complementation assays (discussed later). Untagged and COOH-terminally tagged constructs in the p416 vector were prepared to examine the function of Htm1p in yeast. The p416 TEF centromeric vector is a constitutive expression vector, which uses the strong TEF promoter (translation elongation factor 1 promoter of TEF2) (171). The untagged construct (HTM1/p416) contained the full-length *HTM1* coding region. There were two COOH-tagged constructs. The first contained the full-length *HTM1* coding region with a myc-His tag as a fusion protein at the COOH-terminus, and the second contained a myc-HA-His tag. *S. cerevisiae* strains (Table 6) were transformed with one of the constructs, and the *URA3* selectable marker in the vector was used to select cells on uracil-deficient plates. The expression cultures were grown in either YPD media or uracil-deficient SC media. The recombinant protein was either purified from cell extracts by Ni^{+2}-NTA chromatography (Figure 24) or characterized by biosynthetic labeling with [^{35}S]-Met/Cys. By either of these methods, a protein with an apparent molecular mass of 97 kDa was observed by SDS-PAGE for HTM1-myc-His.

Glycosylation of Htm1p:

A glycosidase digest was performed in order to examine the N-linked glycosylation state of Htm1p. Ni^{+2}-NTA purified Htm1p expressed from HTM1-myc-His/p416 was digested with ~2 units N-glycanase. The samples were run on SDS-PAGE, blotted, and probed with anti-HTM1 antibody (Figure 25). Digestion with N-glycanase resulted in a shift in apparent molecular mass from 97 kDa to 92 kDa, indicating that Htm1p receives N-linked glycosylation modifications.

Gene disruption in *S. cerevisiae:*

To determine the function of *HTM1*, the coding region was disrupted. Gene disruption was accomplished by PCR targeting (172), which involved a PCR-generated product of the *S. cerevisiae* selection marker (*KanMX*) flanked by 35-50 bp of DNA homologous to the 5' and 3' ends of *HTM1* (Figure 26). The *KanMX* marker is a hybrid of the TEF promoter and of the coding sequence of the kan^r gene of transposon *Tn*903 that encodes an aminoglycoside transferase, which provides resistance to geneticin (G418) (162). The PCR product was transformed into *S. cerevisiae* strains YPH499 and YPH501 (Table 6), and transformants were selected by G418. The gene disruption removed amino acids 17-782 of *HTM1*. In the Southern blot analysis (Figure 27), the full length *HTM1* sequence was used as a probe. From wild-type YPH499 cells, genomic DNA was digested with restriction enzymes, and the labeled probe was able to detect DNA fragments of the expected size, indicating that the *HTM1* coding region was present. However, using Δ*htm1* (YPH499) cells, the corresponding bands were not detected, indicating the disruption was successful. In the PCR analysis from wild-type cells, the S9/A5 primer pair and A9/S2 primer pair generated a 700 bp band and 490 bp band, respectively, indicating the presence of the *HTM1* coding region, while in the knockout line, the S9/kanMX primer pair generated a 275 bp band, indicating the presence of the *KanMX2* cassette (Figure 28). The PCR analysis data indicated that *HTM1* had been deleted. The cells were viable after the knockout, and there was no obvious change in the rate of growth of the cells on different carbon sources or at elevated or lowered temperature, and no obvious change in morphology was detected.

Gene disruption of *YLR057w* was performed in a similar fashion. The disruption was done in order to determine whether the role of *YLR057w* was similar to that of *HTM1* (discussed

later). PCR primers were designed to remove amino acids 7-844. The PCR amplimer of the *HIS3* selectable marker was flanked with short sequences homologous to *YLR057w*. *S. cerevisiae* cells were transformed with the amplimer and selected for on histidine-deficient plates. The gene disruption was verified by PCR analysis. The YLR057w sense/A5 primer pair generated an 1170 bp band if the coding region was present, while the YLR057w sense/His A1 primer pair generated a 737 bp band if the *HIS3* marker was present (Figure 29). The cells were viable after *YLR057w* disruption.

HTM1 involved in glycoprotein degradation:

In collaboration with Dr. Markus Aebi (ETH Zurich), a bank of mutants of the N-linked glycosylation pathway was crossed with a Δ*htm1* strain (SS328 background). A phenotype was detected after a cross with a *S. cerevisiae* strain containing the *stt3-7* mutation. The phenotype of the cross is a suppression of the *stt3-7* phenotype. Stt3p is a component of the oligosaccharyl transferase (OST) and is essential for vegetative growth in yeast. Of the *STT3* mutations that result in an underglycosylation of proteins, *stt3-7* is the most severe (173). The Stt3-7p mutant has a reduced OST activity and a temperature-sensitive, synthetic lethal phenotype.

Prior work on Δ*mns1* (ER Man I) had shown that its cross with *stt3-7* produced the same phenotype as Δ*htm11* x *stt3-7* (personal communication, Markus Aebi), and previous work on *MNS1* had also demonstrated a role for *MNS1* in glycoprotein degradation using a misfolded carboxypeptidase Y (CPY) as a marker protein (1). CPY is a vacuolar proteinase. In the ER, the proCPY receives four N-linked oligosaccharides (67 kDa), which are modified in the Golgi (69 kDa). When CPY reaches the vacuole, proteolytic cleavage forms the mature protein (63 kDa). CPY* is a misfolded version of CPY, which is caused by a point mutation (174). When

expressed at normal levels, CPY* does not reach the vacuole, but is largely retained in the ER, followed by degradation by the ubiquitin-proteasomal pathway (174,175).

As Δ*htm1 stt3-7* and Δ*mns1 stt3-7* strains share the same phenotype, *HTM1* was examined to see if it played a role in the degradation of CPY*. The following isogenic *S. cerevisiae* strains (SS328 or YG618 background) were used to study *HTM1*: Δ*mns1* CPY, Δ*mns1* CPY*, Δ*htm1* CPY, Δ*htm1* CPY*, Δ*htm1* Δ*mns1* CPY*. In cell extracts from equivalent amounts of cells, CPY or CPY* from the various strains was examined in immunoblots using anti-CPY antibody. Like Δ*mns1* CPY*, Δ*htm1* CPY* resulted in an accumulation of CPY*; however, wild-type CPY was processed normally in the Δ*htm1* strain (Figure 30). The CPY* which accumulated was the p1 form (ER); additionally, some appears to be in the p2 form (Golgi) (Figure 30). In the double knockout (Δ*htm1* Δ*mns1* CPY*) CPY* accumulated, but the effect was not additive (Figure 34). *HTM1* appears to be involved in the degradation of CPY*, but not of wild-type CPY.

***HTM1* expression restores CPY* degradation:**

The full-length, untagged (HTM/p416) and COOH-tagged constructs (HTM-myc-His/p416 and HTM1-myc-HA-His/p416) were each transformed into the Δ*htm1* CPY* strains to determine whether they could restore CPY* degradation to wild-type levels (Figure 31). Vector encoded *HTM1* was able to restore degradation, and the addition of the COOH-tags did not alter the ability to complement. The p416 vector control was not able to restore complementation. Protein levels were compared by immunoblotting with an anti-actin antibody.

The *HTM1* point mutations were also tested for their ability to restore CPY* degradation (Figures 32 and 33). T495A/p416 was able to complement, indicating this residue is not

essential for function. D279A/p416 and R407A/p416 did not restore complementation, suggesting that these residues are necessary for function.

***YLR057w* is not involved in CPY* degradation:**

YLR057w is a protein with similarity to the Class I mannosidases and the EDEM proteins. In *S. cerevisiae*, both *MNS1* and *HTM1* play a role in CPY* degradation; the other *S. cerevisiae* homolog, *YLR057w,* was also tested to see if it played a role in CPY* degradation. Similar to the Δ*htm1* CPY* strain, the following isogenic Δ*ylr057w* strains (Table 6) were examined for CPY* accumulation: Δ*ylr057w* CPY*, Δ*ylr057w* Δ*htm1* CPY*, Δ*ylr057w* Δ*mns1* CPY*, and Δ*ylr057w* Δ*htm1* Δ*mns1* CPY*. In these strains, the levels of CPY* were not affected by disruption of *YLR057w* (Figure 34).

Δ*der1*:

A Δ*der1* strain was generated for immuoprecipitation studies, as its disruption is known to result in the ER accumulation of soluble, misfolded proteins (5,176). Gene disruption of *DER1* was also performed by PCR targeting using the *KanMX* marker. In the amplimer, the *KanMX* cassette was flanked by 40 bp homologous to the 5' and 3' ends of *DER1*. The PCR disruption was designed to remove amino acids 13-199. Der1p has 211 residues and is an ER membrane protein. In our disruption, CPY* also accumulated when *DER1* was knocked out in the YG618 strain (Figure 35).

Immunoprecipitation of Htm1p from *S. cerevisiae* cell extracts:

To determine whether Htm1p is able to interact with the ERAD substrate, CPY*, HTM1-myc-HA-His/p416 was transformed into the following yeast strains: Δ*ubc7* CPY*, CPY*, Δ*der1* CPY*, and Δ*htm1* CPY*. Immunoprecipitation from [^{35}S] methionine-labeled cell lysates with the antibody to the HA epitope resulted in the immunoprecipitation of Htm1p and an unidentified band of ~80 kDa (Figure 36). Co-immunoprecipitation of CPY* could not be detected.

Immunoprecipitation of Kar2p from [^{35}S]-labeled extracts resulted in the co-immunoprecipitation of a protein similar in size to Htm1p (Figure 37). Elution from the resin and subsequent sequential immunoprecipitation with antibody to HA confirmed that the corresponding band was indeed Htm1p.

II. EDEM2

Characteristics of the human EDEM cDNA sequences:

Three human EDEM homologs were found by BLAST searches of EST sequences using the coding regions of Family 47 α-mannosidases as query sequences. Based on the nomenclature for mouse and human EDEM1 (96), we have termed the remaining human homologs EDEM2 and EDEM3.

The open reading frame of EDEM2, also known as C20orf31, encodes a 578 amino acid (M_r 64,752 Da) protein (Figure 38) that predicts an NH_2-terminal cleavable signal sequence. Kyte-Doolittle hydropathy analysis (Figure 17) and analysis of the EDEM2 sequence using the SOSUI program suggest EDEM2 is a soluble protein (177,178). Four putative Asn-linked

glycosylation consensus sites were identified. With the exception of two polymorphisms, the protein sequence matches the NCBI protein reference sequence (GenPept NP_060687) (Figure 38). In the nucleotide sequence (GenBank NM_018217) there are two silent polymorphisms (bp 222 and 270) and two nonconservative changes (bp 1366 and 1531), which result in protein sequence changes of A456T and S511C, respectively. Our protein sequence is identical to GenPept BAB14731.1/Swiss-Prot Q9BV94; however, in comparison to the nucleotide sequence (GenBank AK023931), there is one silent polymorphism at bp 270.

The open reading frame of EDEM1 encodes a 657 amino acid (M_r 73,768 Da) protein that is predicted to be a type-II ER transmembrane protein. Five putative Asn-linked glycosylation consensus sites were identified. The open reading frame of EDEM3 (c1orf22) encodes a 889 amino acid (M_r 100,304 Da) protein. EDEM3 is predicted to be a soluble protein. Seven putative Asn-linked glycosylation consensus sites were identified.

The sequence relationship shared by the EDEM homologs is shown in Figures 9 and 18. Comparison of the entire sequence indicated that human EDEM1 is 92% identical to mouse EDEM and human EDEM2 is 93% identical to its mouse homolog (GenPept AAH08268). A mouse homolog of EDEM3 was not found. Comparing the mannosidase homology domain of the human EDEM homologs, EDEM2 shares 42% and 43% identity with EDEM1 and EDEM3, respectively. Comparatively, EDEM2 shares 32% identity with human ER Man I. Outside of the mannosidase homology domain, the human EDEM homologs are surprisingly not conserved. EDEM1 has a much longer N-terminus before the mannosidase homology domain compared to EDEM2 and 3, which have only a few residues before the mannosidase homology domain (Figure 19 and Table 5). The sequences following the mannosidase homology domain are also

divergent. The length of the tail is comparable between EDEM1 and 2, while EDEM3 has an extensively longer tail.

Generation of anti-EDEM2 antibodies:

A fragment of EDEM2 (residues 5 – 297) was expressed in *E. coli* to prepare an antibody for immunodetection. In the EDEM2/pQE32 construct, the predicted signal sequence was replaced with a 6x His tag and a vector-encoded start codon. The method for expression and purification (Figure 40) was similar to HTM1/pQE32. After purification, the major band seen by SDS-PAGE was at ~30 kDa. One purification run yielded 1 mg. The anti-EDEM2 antibody generated against the *E. coli* expressed protein was able to detect 10 ng of the antigen at 1:20,000 dilution.

Two MAP antibodies (179) were generated against EDEM2. The 506 MAP antibody was generated against a peptide at the COOH-terminus, residues 535 - 552. The NH_2-MAP antibody was generated against a peptide corresponding to residues 21 - 40. Using Ni^{+2}-NTA purified EDEM2, both antibodies detect EDEM2, but the NH_2-MAP antibody also detects a cross reacting band at ~95 kDa.

Recombinant protein expression in HEK293 cells:

Epitope-tagged constructs were prepared (Figure 41) to examine the localization and function of full-length EDEM2 in mammalian cells. The first construct (EDEM2-myc-HA-His/pEAK) contained the full-length EDEM2 coding region with a myc-HA-His tag as a fusion protein at the COOH-terminus. Constructs were also designed with modifications at the NH_2-terminus. In the first construct (NHis-EDEM2/pEAK), the putative NH_2-terminal signal

sequence was replaced with the cleavable signal sequence of IgG □, and a 10 x His tag was inserted between the IgG □-signal sequence and EDEM2 coding region (Figure 22). In the second construct (TCM-EDEM2/pEAK), the putative NH_2-terminal signal sequence was replaced with the cleavable signal sequence of the *T. cruzi* □□mannosidase (180). This construct also contained a His tag and HA epitope at the NH_2-terminus between the TCM-signal sequence and the EDEM2 coding region. HEK293 cells were independently transfected with the constructs and stable transformants were generated by antibiotic selection. The recombinant protein was either purified by Ni^{+2}-NTA chromatography from detergent cell extracts, or characterized by biosynthetic labeling with [^{35}S]-Met/Cys and immunoprecipitation from the labeled cell extracts. In radiolabeled cell extracts or unlabeled, Ni^{+2}-NTA purified material of either NH_2- or COOH-tagged forms of EDEM2, an apparent molecular mass of approximately 70 and 71 kDa was observed by SDS-PAGE, respectively. These size differences are consistent with the differences in lengths of the epitope tags appended to the coding regions.

A truncated-EDEM2 construct, corresponding to amino acid 18 to 492 of EDEM2, was also generated from the TCM-EDEM2/pEAK construct. The truncated construct does not contain the COOH-terminal extension beyond the mannosidase homology domain. In immunoprecipitates from radiolabeled HEK293 cell extracts (Figure 42) or from Ni^{+2}-NTA purified unlabeled material, a 58 kDa protein was detected. Truncated-EDEM2 was found both intracellularly and in the media, in contrast to the predominantly intracellular location of full-length EDEM2 (Figure 42). Thus, the COOH-terminal polypeptide extension beyond the □-mannosidase homology domain plays a role in ER retention of EDEM2.

Glycosylation of EDEM2:

In immunoprecipitates from lableled cell extracts, digestion with N-glycanase or endoglycosidase H resulted in a shift in the size of full-length EDEM2 from ~70 kDa to 64 kDa, while truncated-EDEM2 shifted from 58 kDa to ~46 kDa (Figure 43). Treatment of the cells with tunicamycin produced similar results. Sensitivity of the N-glycans to cleavage by endoglycosidase H indicates that the oligosaccharides are still high mannose-type and that the majority of the glycoprotein population of full-length EDEM2 has not reached the Golgi complex.

EDEM2 that was able to be secreted into the media exhibited a slightly larger size than the intracellular form, as determined by SDS-PAGE, suggesting that glycan modification occurred in the Golgi during transit. The secreted EDEM2 was insensitive to digestion by endoglycosidase H (Figure 43).

Gel filtration:

Ni^{+2}-NTA-purified TCM-EDEM2 or NHis-EDEM2 was resolved on an analytical Superdex 200 column where it eluted at approximately the same position as the thyroglobulin standard, 669 kDa (Figure 44). When the zwitterionic compound, NDSB-201, was added to prevent the formation of aggregates, its presence or absence had no observed effect on the elution profile of EDEM2 from the column. The NH_2-tagged EDEM2 protein is approximately 70 kDa by SDS-PAGE. Whether the large size stems from oligomerization or formation of a complex with other proteins has not been determined.

Assay for α-1,2 mannosidase activity:

To determine if EDEM2 has α-1,2 mannosidase activity, the COOH-terminally tagged form of EDEM2 was purified from cell extracts by Ni^{+2}-NTA chromatography and was incubated with $Man_{9-5}GlcNAc_2$-PA tagged substrates. The oligosaccharides were analyzed by NH_2-HPLC, and no change in retention time was detected (Figure 45). In contrast, a similar amount of recombinant ER mannnosidase I (145) or Golgi mannosidase IA (149) each could trim the substrates to their expected digestion products. These data are consistent with the lack of α-1,2 mannosidase activity for other EDEM homologs, *S. cerevisiae* Htm1p (2,3) and murine EDEM (96).

To determine the detectable limit of PA-tagged oligosaccharides, a dilution series of $Man_9GlcNAc_2$-PA was run on the NH_2-HPLC column (Figure 46). The detection limit was 1 pmol. Additionally, a small amount of $Man_8GlcNAc_2$, is usually present in the $Man_9GlcNAc_2$ standards. The $Man_8GlcNAc_2$ peak, which represents 1.5% of the area of the $Man_9GlcNAc_2$ peak, can be seen when 100 pmol of $Man_9GlcNAc_2$ is loaded.

EDEM2 is localized to the ER:

The subcellular localization of EDEM2 was determined by the stable transfection of the full-length COOH-terminally tagged protein (EDEM2-Myc-His/pcDNA3.1) into HEK293 cells followed by detection of the Myc epitope tag by indirect immunofluorescence. Double staining of the cells was accomplished by the use of a mouse monoclonal anti-Myc antibody and a rabbit polyclonal anti-calreticulin antibody, detected by the fluorphore-conjugated secondary antibodies. The immunofluorescence pattern of EDEM2-Myc-His was broadly distributed throughout the cytoplasm in a reticular pattern (Figure 47). Untransfected cells showed no

detectable fluorescence. Co-localization of the anti-myc immunofluorescence pattern with the antibody to the ER marker protein, calreticulin, was observed.

Using the NH_2-terminally tagged TCM-EDEM2/pEAK construct (Figure 48) with anti-HA and anti-calreticulin antibodies, a reticular pattern was present and a similar co-localization of EDEM2 and calreticulin was observed. Comparing TCM-EDEM2 with the Golgi marker protein, □-mannosidase II, the immunofluorescence pattern was not similar, although some limited overlap could be detected. No similarity in localization was detected between TCM-EDEM2 and the cytosolic proteins MEK-1 or cytosolic □□mannosidase. When the cell line expressing truncated-EDEM2 was examined (Figure 49), a reticular pattern was observed and a similar co-localization with calreticulin was revealed. The overlay was not completely identical, likely as a result of the intensity of the signal. Truncated-EDEM2 did not co-localize with the Golgi or cytosolic markers. In the cell line not transfected with an EDEM2 construct, no signal was detected with the HA antibody (Figure 50). Other controls included cells not incubated with primary antibody or cells not permeabilized (Figures 47 and 51). These data indicate that the full-length NH_2- and COOH-tagged forms of EDEM2 are localized in the ER and that the portion of truncated-EDEM2 that had not yet been secreted was also retained within the ER.

Immunoprecipitation of EDEM2 with □1-antitrypsin variants:

To determine whether EDEM2 is able to interact with ERAD substrates, HEK293 cells were co-transfected with EDEM2 and mutant forms of □1-antitrypsin. The N_{HK} and PI Z variants were each retained in the ER prior to their degradation by the proteasome, although some PI Z was able to be secreted. Immunoprecipitation of TCM-EDEM2 from [^{35}S]-methionine labeled extracts with the antibody to the HA epitope resulted in the co-immunoprecipitation of proteins

of similar size as PI Z (Figure 52). However, in cells not transfected with α1-antitrypsin, this band was still present. In sequential immunoprecipitations, elution from the resin after immunoprecipitation with antibody to HA and subsequent immunoprecipitation with antibody to α1-antitrypsin confirmed that a portion of the corresponding bands (50 kDa) were indeed α1-antitrypsin (Figure 53). In pulse label experiments, direct detection of EDEM2 in α1-antitrypsin immunoprecipitates were complicated by non-specific background bands in the vicinity of EDEM2. Sequential immunoprecipitations, where α1-antitrypsin immunoprecipitations were followed by HA immunoprecipitations, EDEM2 was readily detected (Figure 53). Control immunoprecipitations containing no primary antibodies resulted in no specific secondary immunoprecipitation products, and cells not expressing epitope tagged EDEM2 also resulted in no detection of secondary immunoprecipitation products. Additionally, the data for EDEM2-myc-HA-His gave similar results as TCM-EDEM2 (Figures 52 and 54).

In support of the data, immunoprecipitation from unlabeled cell extracts followed by immunoblotting also detected an association between EDEM2 and mutant forms of α1-antitrypsin. Thus, α1-antitrypsin could be detected in immunoblots following immunoprecipitation of EDEM2 from cell extracts using the HA antibody (Figure 54). Using TCM-EDEM2 that had been partially purified from extracts of cells expressing TCM-EDEM2 and PI Z by Ni^{+2}-NTA chromatography, immunoprecipitation with the antibody to α1-antitrypsin followed by immunoblotting with the HA antibody resulted in the detection of EDEM2 (Figure 54). These data indicate that a complex between the α1-antitrypsin variants and recombinant EDEM2 was formed, and the location (NH_2- or COOH-) of the epitope tag did not interfere with the interaction.

In a similar experiment, cells were transfected with either TCM-EDEM2 or truncated-EDEM2, labeled with [^{35}S]-Met/Cys and treated with 2 mM DTT for 30 min, and cell lysates were immunoprecipitated with the antibody to HA. Compared to the full length EDEM2, the interaction between the 50 kDa protein and truncated-EDEM2 was greatly reduced (Figure 55).

EDEM2 accelerates the degradation of □1-antitrypsin:

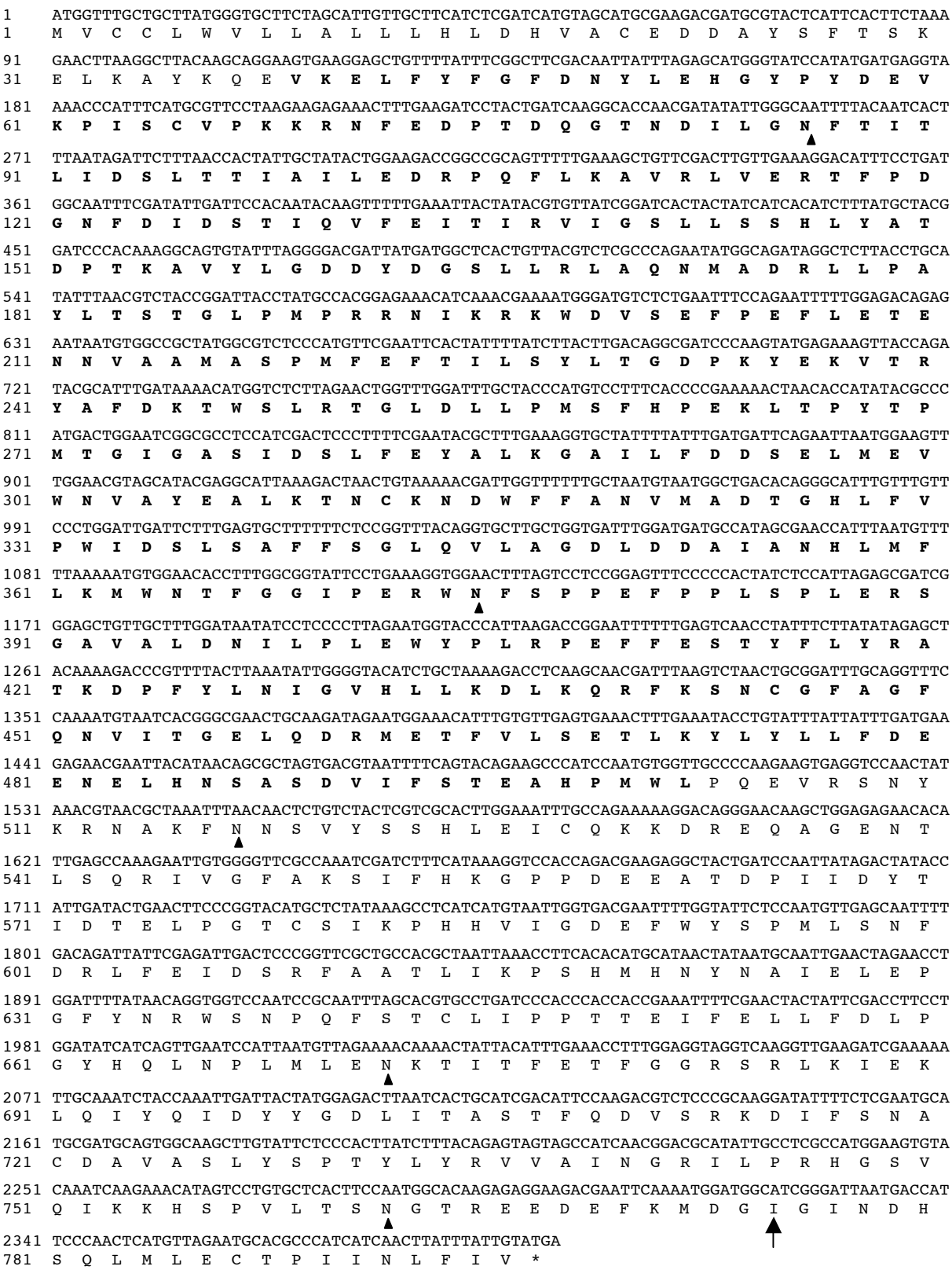
In HEK293 cells, TCM-EDEM2 or truncated-EDEM2 was transiently co-transfected with N_{HK}, PI Z, unglycosylated PI Z, or a □-galactosidase control construct. To examine the effect of EDEM2 on degradation, HEK293 cells were pulse labeled with [^{35}S]-methionine, chased with unlabeled media for up to 5 h, and □1-antitrypsin was immunoprecipitated from the media and cell lysates. During the chase, there is a slight shift in the mobility of □1-antitrypsin variants to a lower apparent molecular mass indicating that glycan trimming had occurred (Figure 56). No shift in mobility was detected for the nonglycosylated PIZ, as expected (Figure 56).

In cells that were not transfected with EDEM2, N_{HK} was retained in the cell for 3 h and then decreased in amount by the 5 h time point. Upon overexpression of TCM-EDEM2, N_{HK} was stable at the 1 h chase time point but was mostly degraded by 3 h (Figure 56). Overexpression of truncated-EDEM2 did not accelerate degradation of N_{HK}, compared to the □-gal control. For the PI Z variant in cells not transfected with EDEM2, some □1-antitrypsin was able to fold properly and be secreted, while the majority was retained intracellularly and was not degraded. Upon overexpression of TCM-EDEM2, PI Z degradation was accelerated, resulting in reduced radiolabeled protein by the 3 h time point and reduced secretion (Figure 56). Overexpression of truncated-EDEM2 did not accelerate degradation of PI Z, compared to the □-gal control (Figure 56).

In cells that were transfected with the nonglycosylated PI Z variant, overexpression of TCM-EDEM2 did not alter the rate of disposal of nonglycosylated PI Z compared to the β-gal control (Figure 56).

Figure 16. Nucleotide and amino acid sequences of *HTM1* and *YLR057w*. In *HTM1*, the sequence that is predicted to be the cleavable signal sequence is underlined. The predicted transmembrane domain of *YLR057w* is also underlined. Triangles below the peptide sequence indicate the potential N-glycosylation sites. The putative GPI-signal cleavage site is indicated by the arrow. The mannosidase homology domain of *HTM1* is in bold.

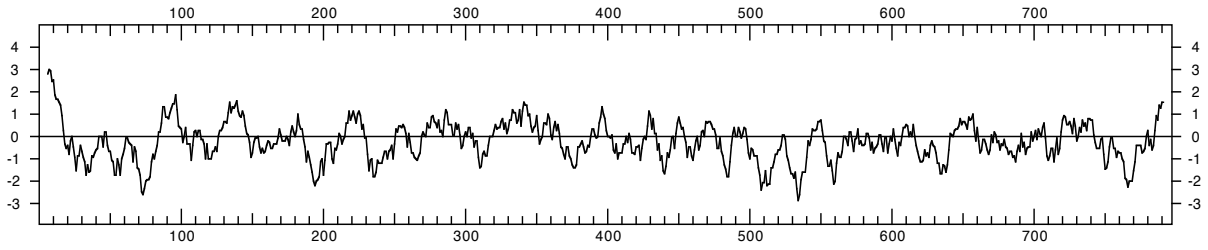
HTM1


1 ATGGTTTGCTGCTTATGGGTGCTTCTAGCATTGTTGCTTCATCTCGATCATGTAGCATGCGAAGACGATGCGTACTCATTCACTTCTAAA
1 M V C C L W V L L A L L L H L D H V A C E D D A Y S F T S K
91 GAACTTAAGGCTTACAAGCAGGAAGTGAAGGAGCTGTTTTATTTCGGCTTCGACAATTATTTAGAGCATGGGTATCCATATGATGAGGTA
31 E L K A Y K Q E V K E L F Y F G F D N Y L E H G Y P Y D E V
181 AAACCCATTTCATGCGTTCCTAAGAAGAGAAACTTTGAAGATCCTACTGATCAAGGCACCAACGATATATTGGGCAATTTTACAATCACT
61 K P I S C V P K K R N F E D P T D Q G T N D I L G N F T I T
271 TTAATAGATTCTTTAACCACTATTGCTATACTGGAAGACCGGCCGCAGTTTTTGAAAGCTGTTCGACTTGTTGAAAGGACATTTCCTGAT
91 L I D S L T T I A I L E D R P Q F L K A V R L V E R T F P D
361 GGCAATTTCGATATTGATTCCACAATACAAGTTTTTGAAATTACTATACGTGTTATCGGATCACTACTATCATCACATCTTTATGCTACG
121 G N F D I D S T I Q V F E I T I R V I G S L L S S H L Y A T
451 GATCCCACAAAGGCAGTGTATTTAGGGGACGATTATGATGGCTCACTGTTACGTCTCGCCCAGAATATGGCAGATAGGCTCTTACCTGCA
151 D P T K A V Y L G D D Y D G S L L R L A Q N M A D R L L P A
541 TATTTAACGTCTACCGGATTACCTATGCCACGGAGAAACATCAAACGAAAATGGGATGTCTCTGAATTTCCAGAATTTTTGGAGACAGAG
181 Y L T S T G L P M P R R N I K R K W D V S E F P E F L E T E
631 AATAATGTGGCCGCTATGGCGTCTCCCATGTTCGAATTCACTATTTTATCTTACTTGACAGGCGATCCCAAGTATGAGAAAGTTACCAGA
211 N N V A A M A S P M F E F T I L S Y L T G D P K Y E K V T R
721 TACGCATTTGATAAAACATGGTCTCTTAGAACTGGTTTGGATTTGCTACCCATGTCCTTTCACCCCGAAAAACTAACACCATATACGCCC
241 Y A F D K T W S L R T G L D L L P M S F H P E K L T P Y T P
811 ATGACTGGAATCGGCGCCTCCATCGACTCCCTTTTCGAATACGCTTTGAAAGGTGCTATTTTATTTGATGATTCAGAATTAATGGAAGTT
271 M T G I G A S I D S L F E Y A L K G A I L F D D S E L M E V
901 TGGAACGTAGCATACGAGGCATTAAAGACTAACTGTAAAAACGATTGGTTTTTTGCTAATGTAATGGCTGACACAGGGCATTTGTTTGTT
301 W N V A Y E A L K T N C K N D W F F A N V M A D T G H L F V
991 CCCTGGATTGATTCTTTGAGTGCTTTTTTCTCCGGTTTACAGGTGCTTGCTGGTGATTTGGATGATGCCATAGCGAACCATTTAATGTTT
331 P W I D S L S A F F S G L Q V L A G D L D D A I A N H L M F
1081 TTAAAAATGTGGAACACCTTTGGCGGTATTCCTGAAAGGTGGAACTTTAGTCCTCCGGAGTTTCCCCCACTATCTCCATTAGAGCGATCG
361 L K M W N T F G G I P E R W N F S P P E F P P L S P L E R S
1171 GGAGCTGTTGCTTTGGATAATATCCTCCCCTTAGAATGGTACCCATTAAGACCGGAATTTTTTGAGTCAACCTATTTCTTATATAGAGCT
391 G A V A L D N I L P L E W Y P L R P E F F E S T Y F L Y R A
1261 ACAAAAGACCCGTTTTACTTAAATATTGGGGTACATCTGCTAAAAGACCTCAAGCAACGATTTAAGTCTAACTGCGGATTTGCAGGTTTC
421 T K D P F Y L N I G V H L L K D L K Q R F K S N C G F A G F
1351 CAAAATGTAATCACGGGCGAACTGCAAGATAGAATGGAAACATTTGTGTTGAGTGAAACTTTGAAATACCTGTATTTATTATTTGATGAA
451 Q N V I T G E L Q D R M E T F V L S E T L K Y L Y L L F D E
1441 GAGAACGAATTACATAACAGCGCTAGTGACGTAATTTTCAGTACAGAAGCCCATCCAATGTGGTTGCCCCAAGAAGTGAGGTCCAACTAT
481 E N E L H N S A S D V I F S T E A H P M W L P Q E V R S N Y
1531 AAACGTAACGCTAAATTTAACAACTCTGTCTACTCGTCGCACTTGGAAATTTGCCAGAAAAAGGACAGGGAACAAGCTGGAGAGAACACA
511 K R N A K F N N S V Y S S H L E I C Q K K D R E Q A G E N T
1621 TTGAGCCAAAGAATTGTGGGGTTCGCCAAATCGATCTTTCATAAAGGTCCACCAGACGAAGAGGCTACTGATCCAATTATAGACTATACC
541 L S Q R I V G F A K S I F H K G P P D E E A T D P I I D Y T
1711 ATTGATACTGAACTTCCCGGTACATGCTCTATAAAGCCTCATCATGTAATTGGTGACGAATTTTGGTATTCTCCAATGTTGAGCAATTTT
571 I D T E L P G T C S I K P H H V I G D E F W Y S P M L S N F
1801 GACAGATTATTCGAGATTGACTCCCGGTTCGCTGCCACGCTAATTAAACCTTCACACATGCATAACTATAATGCAATTGAACTAGAACCT
601 D R L F E I D S R F A A T L I K P S H M H N Y N A I E L E P
1891 GGATTTTATAACAGGTGGTCCAATCCGCAATTTAGCACGTGCCTGATCCCACCCACCACCGAAATTTTCGAACTACTATTCGACCTTCCT
631 G F Y N R W S N P Q F S T C L I P P T T E I F E L L F D L P
1981 GGATATCATCAGTTGAATCCATTAATGTTAGAAAACAAAACTATTACATTTGAAACCTTTGGAGGTAGGTCAAGGTTGAAGATCGAAAAA
661 G Y H Q L N P L M L E N K T I T F E T F G G R S R L K I E K
2071 TTGCAAATCTACCAAATTGATTACTATGGAGACTTAATCACTGCATCGACATTCCAAGACGTCTCCCGCAAGGATATTTTCTCGAATGCA
691 L Q I Y Q I D Y Y G D L I T A S T F Q D V S R K D I F S N A
2161 TGCGATGCAGTGGCAAGCTTGTATTCTCCCACTTATCTTTACAGAGTAGTAGCCATCAACGGACGCATATTGCCTCGCCATGGAAGTGTA
721 C D A V A S L Y S P T Y L Y R V V A I N G R I L P R H G S V
2251 CAAATCAAGAAACATAGTCCTGTGCTCACTTCCAATGGCACAAGAGAGGAAGACGAATTCAAAATGGATGGCATCGGGATTAATGACCAT
751 Q I K K H S P V L T S N G T R E E D E F K M D G I G I N D H
2341 TCCCAACTCATGTTAGAATGCACGCCCATCATCAACTTATTTATTGTATGA
781 S Q L M L E C T P I I N L F I V *

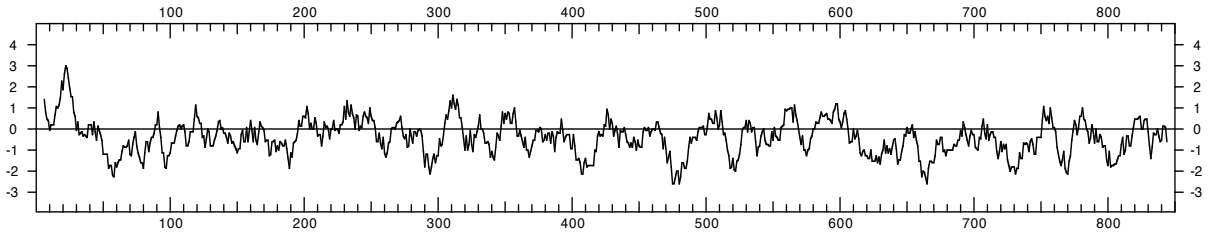
YLR057w

```
1    ATGTCCATAGCGCGATTAGTATACTCTCTTTTCAGAAGAGTACGATCGGTTCTCTTACTTTTCATTACCATATCGCTATTGTTTTACTAT
1    M  S  I  A  R  L  V  Y  S  L  F  R  R  V  R  S  V  L  L  L  F  I  T  I  S  L  L  F  Y  Y

91   ACGTTCCAGAATGAAATCGATATCTTGAATTCGTACGCTTTAAACGACTCACTACCGTCCATAAACAACTATGAGCACAATACAGAAGGT
31   T  F  Q  N  E  I  D  I  L  N  S  Y  A  L  N  D  S  L  P  S  I  N  N  Y  E  H  N  T  E  G

181  TCTTCGAAACTGGATCCTCCAGACCTATCGTCAACCGGTTCTGACCGAATAGCCACCGACAAGGAAAATGGCAACGTTGCAGTGGATTTG
61   S  S  K  L  D  P  P  D  L  S  S  T  G  S  D  R  I  A  T  D  K  E  N  G  N  V  A  V  D  L

271  TCAGATCCAGCTACATTGAGAGAAAAAAACAAGTACTTTCCCTTGCTGTTAAAGGGATCGTCTCACCAGATTGGTTCGAATTTACCAATT
91   S  D  P  A  T  L  R  E  K  N  K  Y  F  P  L  L  L  K  G  S  S  H  Q  I  G  S  N  L  P  I

361  TCGTCTCTTTTGACATATAAGGAAAAGTATCCCGTCTTGTTCGAATACTCATCACCTTCCTTAACTTCTATAAGCCAGAATGACGTTCAT
121  S  S  L  L  T  Y  K  E  K  Y  P  V  L  F  E  Y  S  S  P  S  L  T  S  I  S  Q  N  D  V  H

451  AAGATACAGCCGGCAATGCAATTGCCACCAGATGTTGATATGATAAAACAGATAAAGGACATTTTCATGAAATCATGGAACCAAGAGCAG
151  K  I  Q  P  A  M  Q  L  P  P  D  V  D  M  I  K  Q  I  K  D  I  F  M  K  S  W  N  Q  E  Q

541  TTGCTGCTGAAAAGTAACCTCAGAAGGGAATCTACTTGGCCCATAGATCTAATTGATTCTCTAGATACATTGTATCTATGTGGCGAGACG
181  L  L  L  K  S  N  L  R  R  E  S  T  W  P  I  D  L  I  D  S  L  D  T  L  Y  L  C  G  E  T

631  AAACTTTTTCAAGATTCTGTTAATATAATAGAAGATTTTGACTTTCGAGTACCTCCGTTAGCAATGGAAGTTATTGATATTCCTGACATC
211  K  L  F  Q  D  S  V  N  I  I  E  D  F  D  F  R  V  P  P  L  A  M  E  V  I  D  I  P  D  I

721  ACTACCCGTGTCCTAGAAGGGTTACTGTCTGCATACGAATTATCTATGGATAAACGATTACTTAACAAAGCAAAACATGTTGCAGATTTC
241  T  T  R  V  L  E  G  L  L  S  A  Y  E  L  S  M  D  K  R  L  L  N  K  A  K  H  V  A  D  F

811  ATATTAAGATCATTTGATACCCCTAATAGAATTCCTATTTTGAAATATTTTTGGAAATCTGATCTGAGAAATCGATTTCCAGACCGTACT
271  I  L  R  S  F  D  T  P  N  R  I  P  I  L  K  Y  F  W  K  S  D  L  R  N  R  F  P  D  R  T

901  GTTCCATCAGGTCAATTGACTACTATGGCACTCGCGTTTATCAGATTGTCACAGTTAACCCGTTTGAATAAGTATTTTGATGCTGTTGAA
301  V  P  S  G  Q  L  T  T  M  A  L  A  F  I  R  L  S  Q  L  T  R  L  N  K  Y  F  D  A  V  E

991  AGAGTTTTCACCACTATACGTCAATCATATAACGAATTTGATATGGAGTTTATGCTGCCTGACGTCGTTGATGCCTCTGGTTGCCAGTTA
331  R  V  F  T  T  I  R  Q  S  Y  N  E  F  D  M  E  F  M  L  P  D  V  V  D  A  S  G  C  Q  L

1081 TTAACTCAAGAGGAAATTGAGAATGGTGCACACCTAAAGGGATCCAGCATCATGAAAAGTATCAATGAAAATTTCAAGTTCGTACACTGT
361  L  T  Q  E  E  I  E  N  G  A  H  L  K  G  S  S  I  M  K  S  I  N  E  N  F  K  F  V  H  C

1171 AGCAGTTAGGGAAATTCTTAAATCCTCCAATAGATGATAATAGCCTCCAAGAACAATCTCAATACCAGGCTTATAGAATTAATGAAAAA
391  Q  Q  L  G  K  F  L  N  P  P  I  D  D  N  S  L  Q  E  Q  S  Q  Y  Q  A  Y  R  I  N  E  K

1261 ACAGTACCCATTCTAGAAAATCTATTCAAAATCAATGACCTCTTCCAATCTTCTTATGATATCCTGGACGGCTCAAGTAAAAATGCTAAT
421  T  V  P  I  L  E  N  L  F  K  I  N  D  L  F  Q  S  S  Y  D  I  L  D  G  S  S  K  N  A  N

1351 GCTGCAACGATGGACCCCAGTATTGGTTCAGAGGTAGAAGCAGTCGACGAAATAATTGAGAAAAGAAACTTTAAAGACGGTACTAAAAAA
451  A  A  T  M  D  P  S  I  G  S  E  V  E  A  V  D  E  I  I  E  K  R  N  F  K  D  G  T  K  K

1441 GATTCGACAAAAAACACCGTAGGTGATAAATCATTAATAGACTCGCAAACGTTTTTAACAAATTCCATCAGTAACATTTTTAAATTCATG
481  D  S  T  K  N  T  V  G  D  K  S  L  I  D  S  Q  T  F  L  T  N  S  I  S  N  I  F  K  F  M

1531 ACTTTTCGACCCATGTTGCCCAAGCAAACGGAAAACAAAAAGTTTAATTTTCTGAATTCCATTCTAACAAAATCTCAATTCATGCCCACG
511  T  F  R  P  M  L  P  K  Q  T  E  N  K  K  F  N  F  L  N  S  I  L  T  K  S  Q  F  M  P  T

1621 ACTAATGAGCTAGATGTCACGATAAGAAAATCTTACGATGTCTCGTTATACTCATGCAGACTTGGCGGCATTCTAGGTTTAAGTTCACGT
541  T  N  E  L  D  V  T  I  R  K  S  Y  D  V  S  L  Y  S  C  R  L  G  G  I  L  G  L  S  S  R

1711 GTTCCCCACCGAGGCGGTGTCAACACGAAATATATTTTACCATCGTCACTATTGGAAATGAGCGAAATAATAACAGAAAGTTGTTTTATG
571  V  P  H  R  G  G  V  N  T  K  Y  I  L  P  S  S  L  L  E  M  S  E  I  I  T  E  S  C  F  M

1801 CTAATGGAGGAGTTCGACGGGCTGCTCCCTCAAAAATTCGAGTTGGACCCTTGCACTGATGAGACTAATGGGAATTGCGAGTTTAACGGC
601  L  M  E  E  F  D  G  L  L  P  Q  K  F  E  L  D  P  C  T  D  E  T  N  G  N  C  E  F  N  G

1891 GAAACTAAATCACGAATGATTGCCAACGGCGAATATGAAACTTTTGAAAATGACCTAGATGTCGGGATCAAAGTTTCGAATTATGGGAAA
631  E  T  K  S  R  M  I  A  N  G  E  Y  E  T  F  E  N  D  L  D  V  G  I  K  V  S  N  Y  G  K

1981 GGTGGGAATGACCAGAAAGCTAAGAGAAATGTACTAAGCAAAGATGGGATAACAGAGACTCAAAACATAAAAGGTGATACTGTTGGAAGC
661  G  G  N  D  Q  K  A  K  R  N  V  L  S  K  D  G  I  T  E  T  Q  N  I  K  G  D  T  V  G  S

2071 TCGAAATCGATTGCTGAAATAGACGGCGACGAAGTAACGCAAATACGTAGAGTTTTTACGCTTGGTAAGGATATCAAACCGCATATTACA
691  S  K  S  I  A  E  I  D  G  D  E  V  T  Q  I  R  R  V  F  T  L  G  K  D  I  K  P  H  I  T

2161 ACGGACGATACCATGGGTTCTCAGTGGAAAAACCATCCAGACTGGCCATTTTGGGTCAACAAGGTAGAATCCAGAAGATTGTTGGATTCG
721  T  D  D  T  M  G  S  Q  W  K  N  H  P  D  W  P  F  W  V  N  K  V  E  S  R  R  L  L  D  S

2251 AATATCATTGAATCTATATTTTATATGTACAGAATCAGTGGAGAACAAAAATGGAGAAGCATGGGTAAGCAGTCTTTTGGAATCTTAATG
751  N  I  I  E  S  I  F  Y  M  Y  R  I  S  G  E  Q  K  W  R  S  M  G  K  Q  S  F  G  I  L  M

2341 CAAGAATTAATGGAACTAAACAGTGGTGCCAAAGGATTGTGGCAAATAAAGGAATTTTATGAAAACGGTGAAAAAGTGAACAATGATTTA
781  Q  E  L  M  E  L  N  S  G  A  K  G  L  W  Q  I  K  E  F  Y  E  N  G  E  K  V  N  N  D  L

2431 CCTAGCTATTGGTTCTCCAGAACATTAAAATACTATTTGCTACTATTCAGCGATGGTGACAAGGTTTCCCTGGACAAACACATCTTAACT
811  P  S  Y  W  F  S  R  T  L  K  Y  Y  L  L  L  F  S  D  G  D  K  V  S  L  D  K  H  I  L  T

2521 CAAGGAGGGCACATAATCAAAAAGAAATGA
841  Q  G  G  H  I  I  K  K  K  *
```

Htm1p



YLR057w



EDEM2

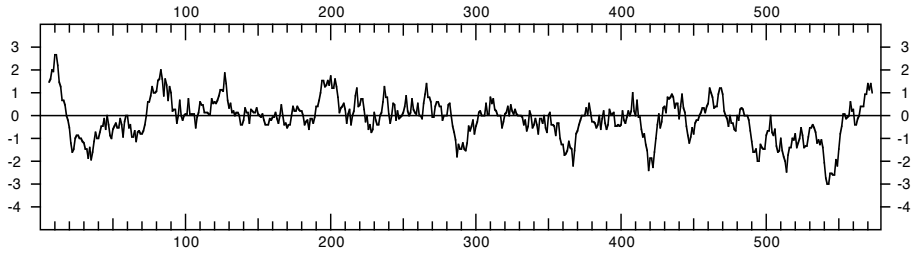
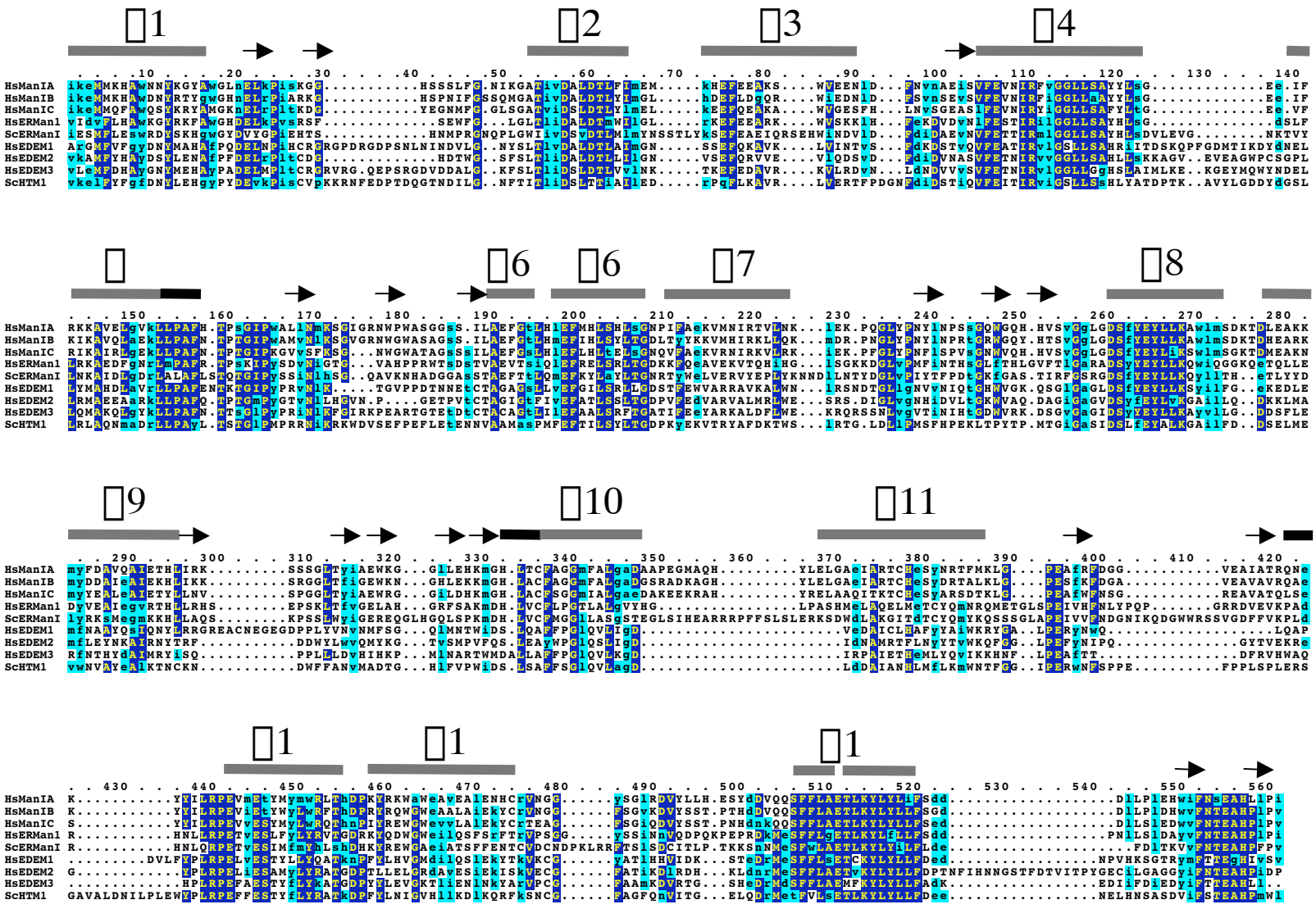


Figure 17. Kyte-Doolittle hydropathy plots for Htm1p, YLR057w, and EDEM2. The plots were obtained using the DNA Strider program (v1.2).

Figure 18. Alignment of human EDEM homologs with representative sequences of the catalytic domains of the Family 47 glycosylhydrolases. The multiple sequence alignment generated by Clustal W contains the catalytic core of selected members of the Class I mannosidases for comparison to the EDEM homologs: human EDEM1 (GenPept CAC69370); human EDEM2 (GenPept NP_060687); human EDEM3 (GenPept NP_079467); human ER mannosidase I (ER ManI) (GenPept NP_057303); human Golgi mannosidase IA (ManIA) (GenPept NP_005898); human Golgi mannosidase IB (ManIB) (GenPept NP_006690); human Golgi mannosidase IC (ManIC) (GenPept NP_065112); *S. cerevisiae* Htm1p (NP_012074); and *S. cerevisiae* ER mannosidase I (ScERManI) (NP_012665). Identical residues are shaded dark blue, while similar residues are shaded light blue. Using the crystal structure of human ER Man I, the positions of the α helices are represented by grey bars, 3_{10} helices are black bars, and β sheets are shown by arrows. The alignment includes amino acid residues 39-482 of EDEM2.


HsManIA
HsManIB
HsManIC
HsERMan1
ScERManI
HsEDEM1
HsEDEM2
HsEDEM3
ScHTM1

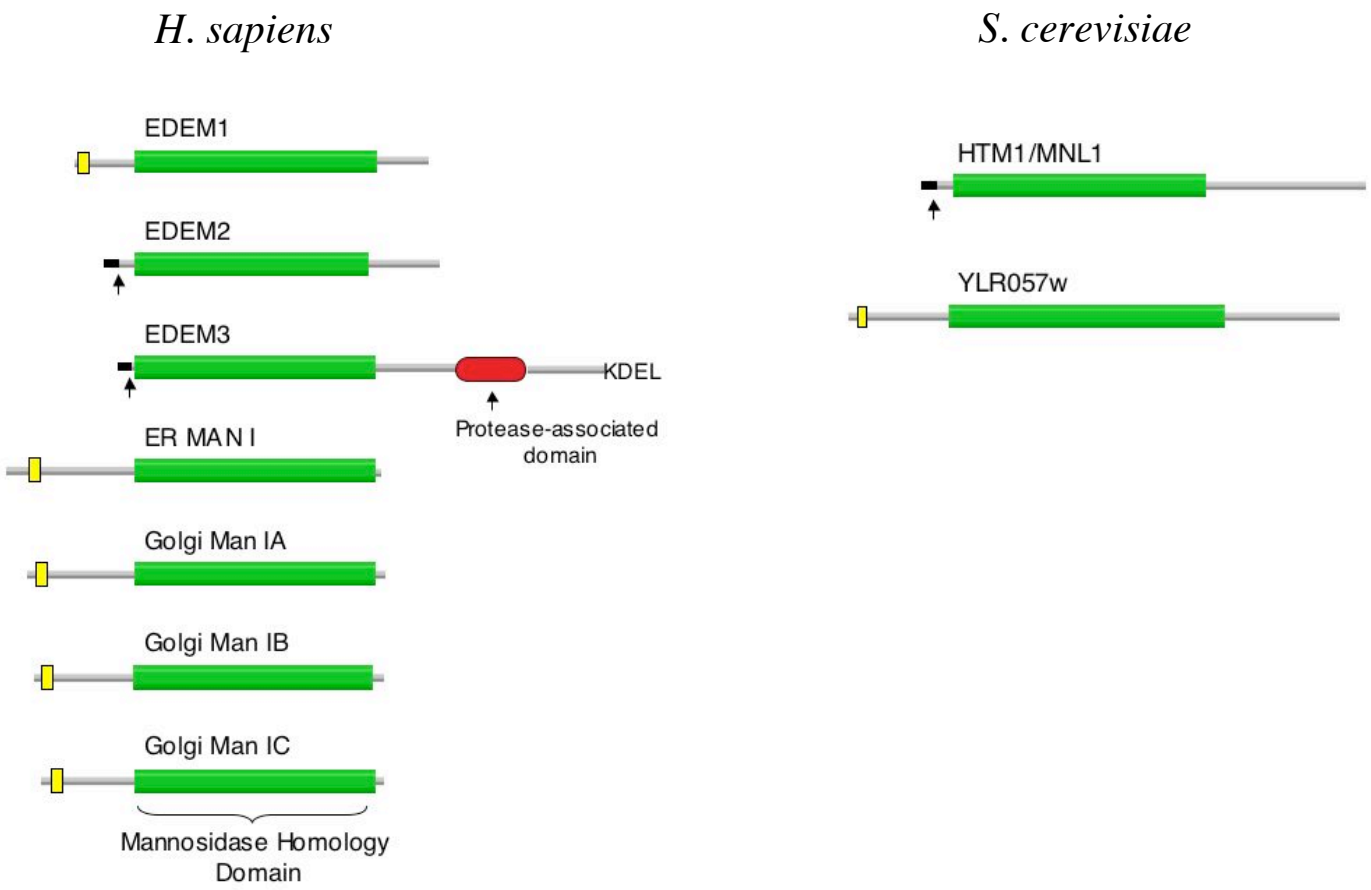


Figure 19. Cartoon representation of the full-length protein sequences of human and *S. cerevisiae* Family 47 members. The mannosidase homology domain is shown in green, the protease-associated domain is red, the signal sequence is black, the transmembrane domain is yellow, and other sequences are grey. Putative signal sequence cleavage site is indicated by the arrow.

Table 4. Sequence identity in Family 47 glycosylhydrolases

% similarity

		1	2	3	4	5	6	7	8	9	10
1	*H. sapiens* Golgi Man IA	---	84.2	77.7	55.4	44.6	41.7	42.1	40.9	38.4	23.1
2	*H. sapiens* Golgi Man IB	71.1	---	76.6	54.8	45.2	42.3	41.2	41.6	38.7	23.9
3	*H. sapiens* Golgi Man IC	64.6	62.6	---	52.1	43.2	39.5	41.2	40.5	37.7	22.9
4	*H. sapiens* ER Man I	36.5	36.7	36.1	---	52.7	42.7	42.7	44.3	38.6	24.6
5	*S. cerevisiae* ER Man I	30.9	32.3	29.9	38.1	---	35.4	35.4	35.5	32.5	22.2
6	*H. sapiens* EDEM1	27.7	28.9	27.7	31.1	24.7	---	54.5	57.7	54.4	25.5
7	*H. sapiens* EDEM2	27.9	29.9	29.6	29.1	23.3	41.9	---	52.3	46.5	23.0
8	*H. sapiens* EDEM3	29.4	29.2	29.4	30.2	23.4	44.7	42.5	---	48.6	23.8
9	*S. cerevisiae* HTM1/MNL1	23.5	25.1	23.9	26.2	21.5	40.6	34.0	35.3	---	22.3
10	*S. cerevisiae* YLR057w	13.4	14.1	11.6	13.5	13.1	14.2	13.3	13.2	13.1	---

% identity

Table 5. Comparison of domain lengths for Family 47 glycosylhydrolases

	NH_2 extension*	mannosidase-homology domain	COOH extension
H. sapiens			
EDEM1	134	452	71
EDEM2	39	443	96
EDEM3	13	442	434
ER Man I	253	442	4
Golgi Man1A	199	441	13
Golgi Man1B	184	442	15
Golgi Man1C	178	439	13
S. cerevisiae			
HTM1/MNL1	39	464	293
YLR057w	146	491	212

*including cytoplasmic tail, transmembrane domain, stem region, or cleaved signal sequences in the relevant protein

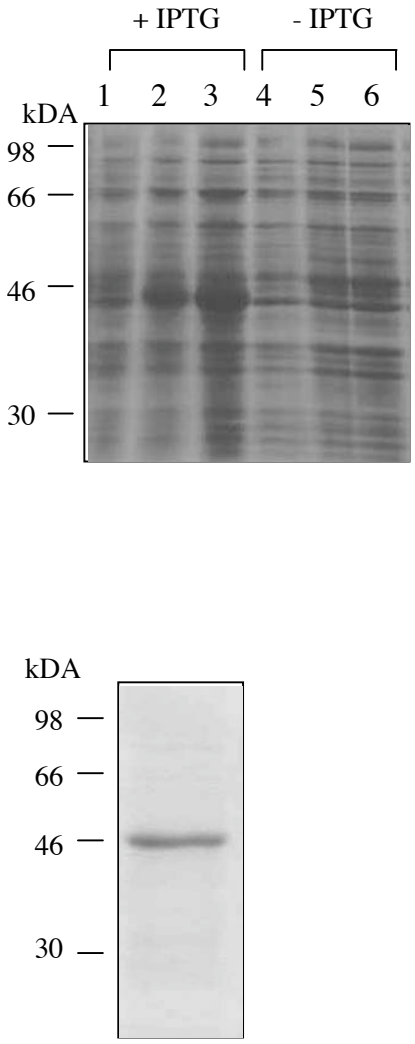


Figure 20. *E. coli* expression of a fragment of Htm1p. The truncated construct contained a 6x His tag at the NH_2-terminal end. (A) Expression was induced with 1 mM IPTG for the indicated times. Lanes 1-3 show *E. coli* lysates after induction for 0, 1,5, and 6 h respectively. Lanes 4-6 show *E. coli* lysates without IPTG for 0, 1.5, and 6 h, respectively. Protein was visualized by Coomassie staining. (B) After induction, recombinant protein was purified by Ni^{+2}-NTA chromatography, pooled, and dialyzed. This material was used for rabbit immunization. Protein was visualized by Coomassie staining.

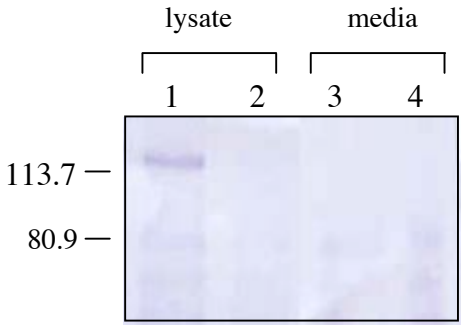


Figure 21. *P. pastoris* expression of Htm1p. Cell lysate and media from 20 ml cell cultures were immunoprecipitated using antibody to the HA epitope, as described in methods section. The samples were separated by SDS-PAGE and immunoblotted with anti-HA antibody. The immunoblot shows two *Pichia* transformants; Lanes 1 and 3 are from the 1st colony. Lanes 2 and 4 are from the 2nd colony. Only the colony in lane 1 has expression, and no secreted expression was detected.

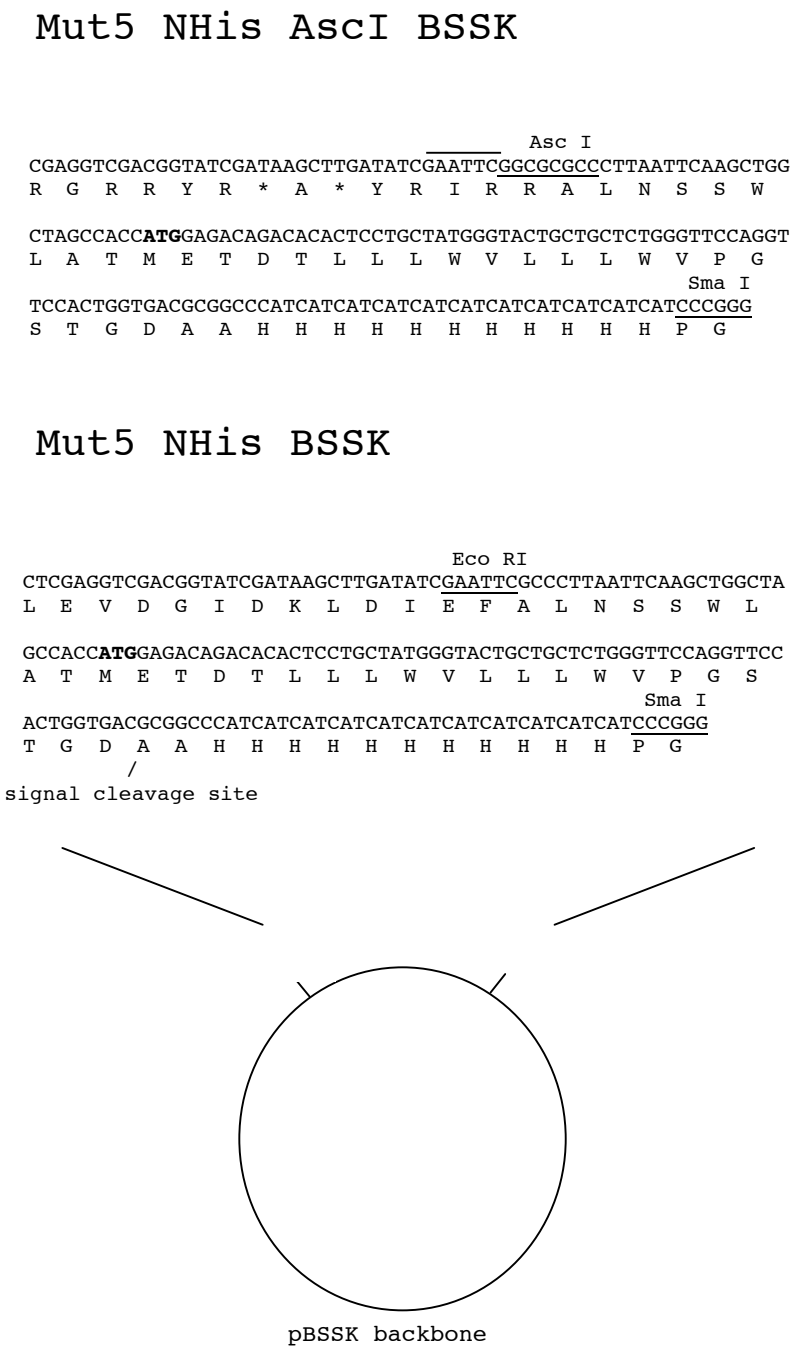


Figure 22. Ig κ vectors. The Mut5 NHis BSSK and Mut5 NHis AscI BSSK vectors were created by insertion of the the Ig κ signal cleavage sequence and His tag between the Eco RI and SmaI sites in pBSSK. The Asc I site was added by site directed mutagenesis.

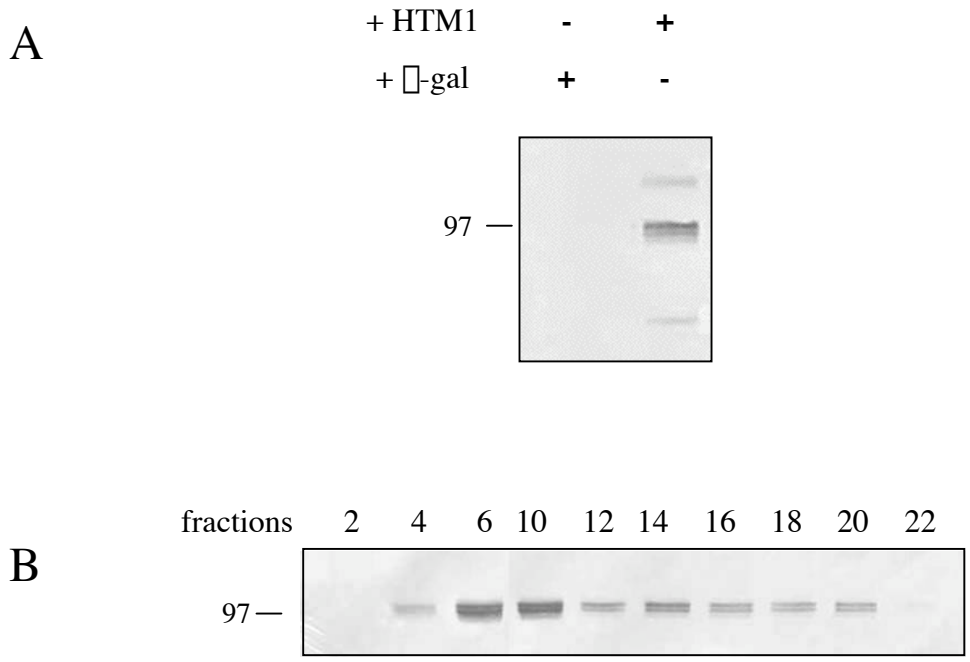


Figure 23. Mammalian expression of Htm1p. (A) Test of Htm1p pQE antibody. HEK293 cells transfected with NHis-HTM1/pEAK or mock transfected with β-gal/pEAK. The cell lysate was separated by SDS-PAGE, and Htm1p was detected by immunoblotting with anti-HTM1 pQE antibody (1/2000) followed by anti-rabbit alkaline phosphatase-conjugated secondary antibody (1/5000). (B) NHis-HTM1/pEAK was expressed from stably transfected HEK293 cells grown in suspension culture (0.5 L). Cell lysate from ~2.8×10^8 cells was purified by Ni^{+2}-NTA chromatography. After washing the resin, fractions were collected during elution by imidazole. Htm1p was detected after immunoblotting with HTM1 pQE antibody (1/2000) followed by anti-rabbit alkaline phosphatase-conjugated secondary antibody (1/5000). No bands were observed after Coomassie staining of a gel containing the same fractions as the immunoblot.

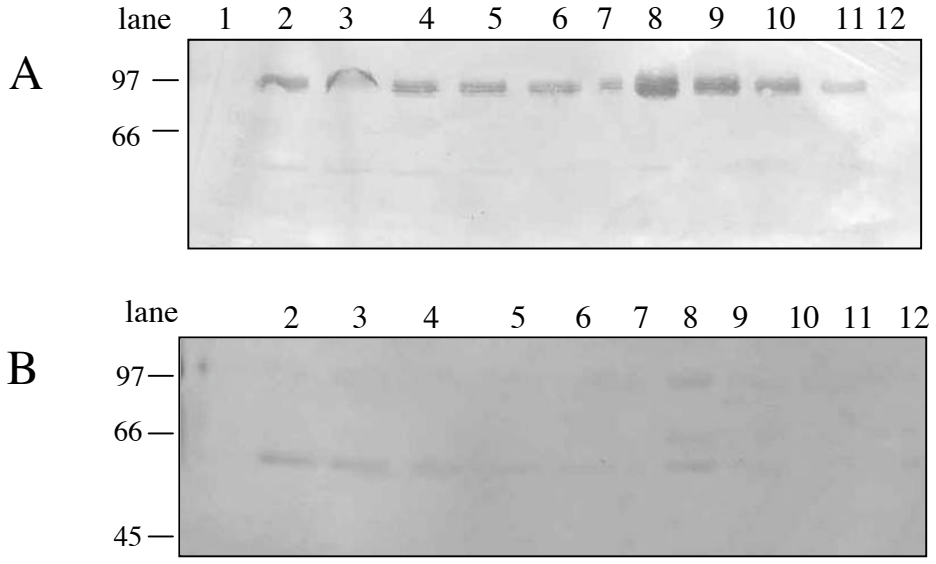


Figure 24. *S. cerevisiae* expression of Htm1p. HTM1-myc-HIS/p416 was expressed in *S. cerevisiae* cells, and cell lysate from a 3L culture was purified by Ni^{+2}-NTA chromatography. After washing the resin, fractions were collected during elution by imidazole; lanes indicate fractions collected over elution profile. In panel A, Htm1p was detected after immunoblotting with anti-myc antibody (1/5000) followed by anti-mouse alkaline phosphatase conjugated secondary antibody (1/5000). In panel B, a gel containing the same fractions was stained with Coomassie dye. Contaminating proteins are observed early and in the mid-range of the run. A band in lanes 8 and 9 corresponds to the size of Htm1p.

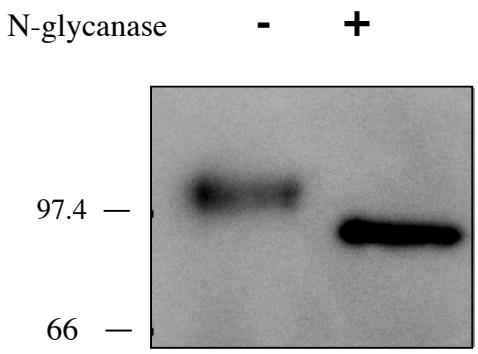
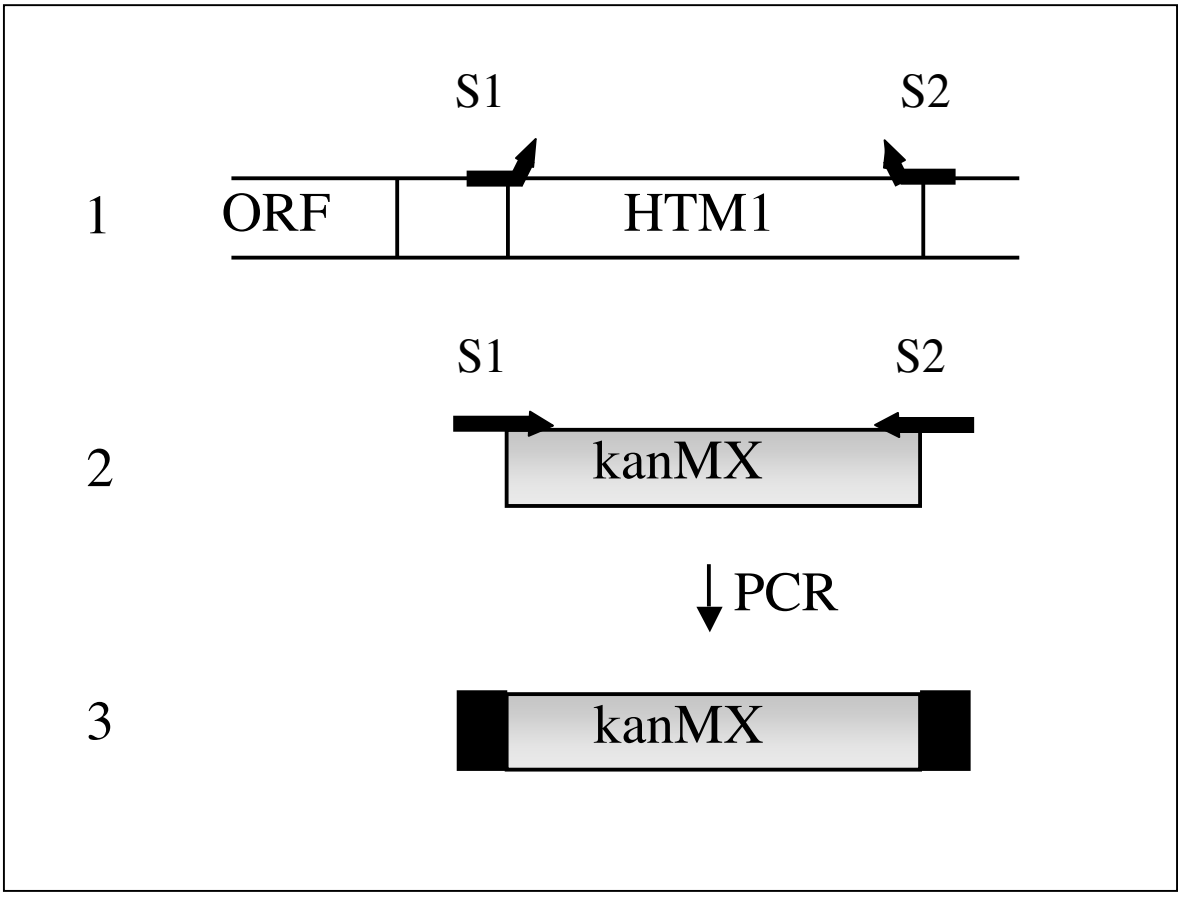


Figure 25. N-glycanase digest of Htm1p. Htm1p containing a myc-HIS tag was expressed in *S. cerevisiae*, and the Ni^{+2}-NTA purified protein was digested with N-glycanase. The sample was separated by SDS-PAGE. The immunoblot was probed with anti-HTM1 pQE antibody (1:1000) followed by the anti-rabbit–alkaline phosphatase-conjugated antibody (1:5000).



Modified from unpublished application protocol of Achim Wach, Arndt Brachat, and Peter Philippsen

Figure 26. Short flanking homology PCR. (1) The primers have 35-45 nucleotides homologous to *S. cerevisiae* genomic DNA target locus (*HTM1*) followed by 18-19 nucleotides of sequence derived from the kanMX marker. (2) The oligonucleotide primers are annealed to the marker, and (3) a PCR amplimer is generated, which contains the kanMX marker flanked by short regions homologous to the genomic DNA.

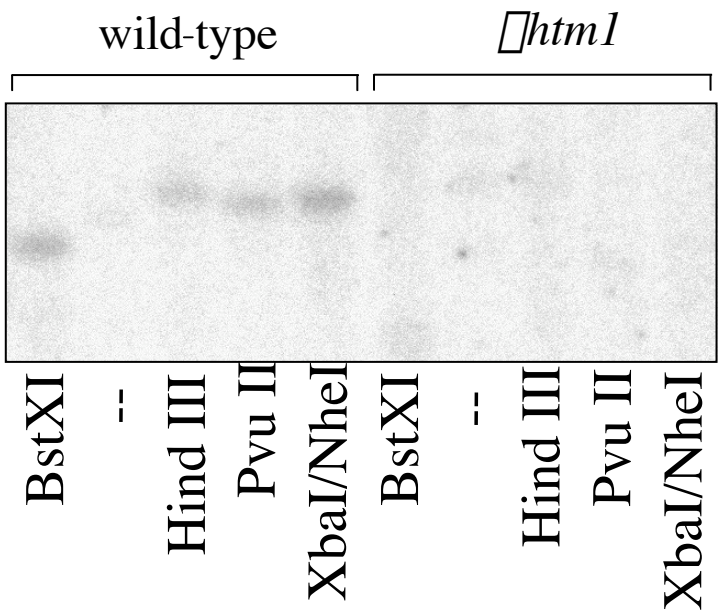
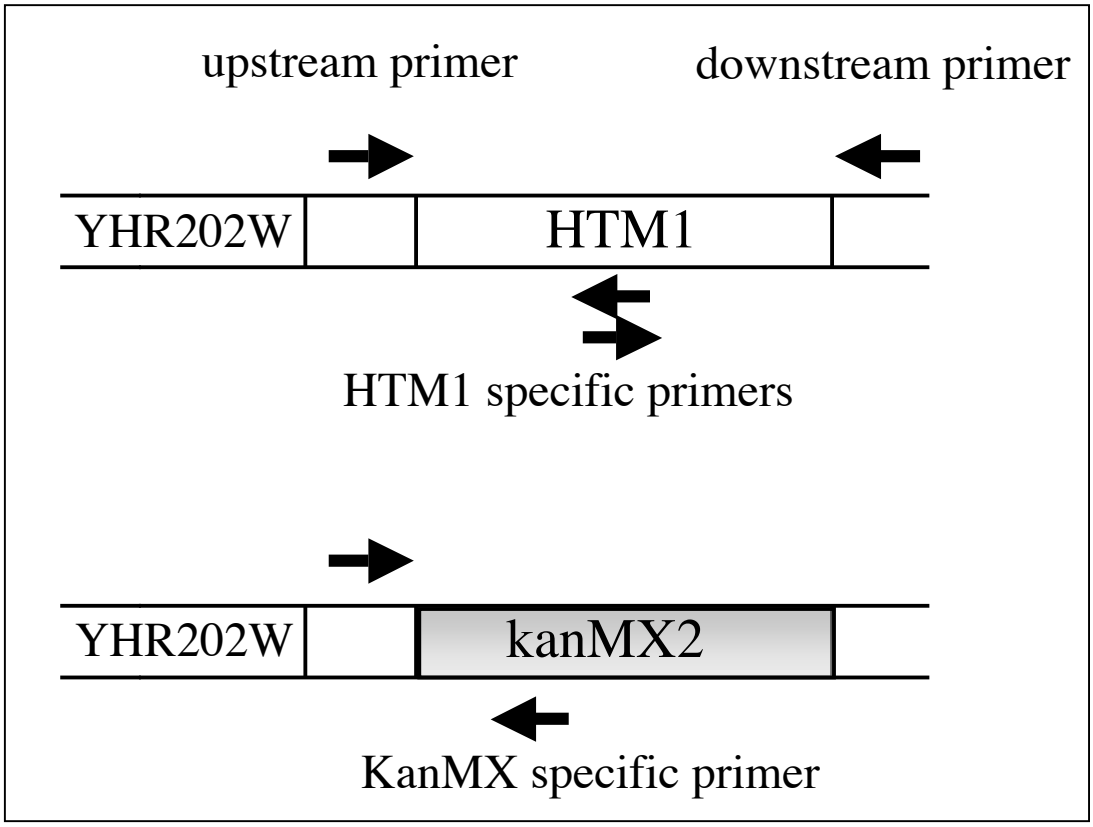


Figure 27. Southern blot analysis. For the Southern blot, genomic DNA from *S. cerevisiae* was digested with BstXI, Hind III, Pvu II, or Xba I/Nhe I as indicated (described in methods). Digestion products were separated on a 1% agarose gel, blotted onto a nylon membrane and probed with the full-length coding region of *HTM1*. *HTM1* was detected in the wild-type strain. In the Δ*htm1* strain, the probe was not able to anneal, and bands were not observed, as expected.

Figure 28. PCR-analysis to identify clones with the correctly targeted marker. (A) Cartoon representation of PCR-analysis. Oligonucleotides are designed to bind outside of the target locus (upstream and downstream primers), within the target locus (*HTM1* specific primers), and within the marker (kanMX specific primer). In diploid *S. cerevisiae*, clones with the correctly integrated marker can be detected by the appearance of two PCR products of the predicted length, one for the wild-type allele and one for the mutated allele. (B) The figure shows an agarose gel of the PCR reaction using HTM1 S9 (upstream primer), HTM1 A5 (*HTM1* specific primer), and *kanMX* primer. The HTM1 S9/HTM1 A5 pair generated a 700 bp band if *HTM1* was present, and the HTM1 S9/*kanMX* primer pair generated a 275 bp band if the kanMX2 cassette inserted at the target locus. The template for the reaction was genomic DNA from YPH499 wild-type cells (lanes 1), YPH501 wild-type cells (lane 2), Δ*htm1* YPH501 cells (lanes 3-5), and Δ*htm1* YPH499 cells (lanes6-7).

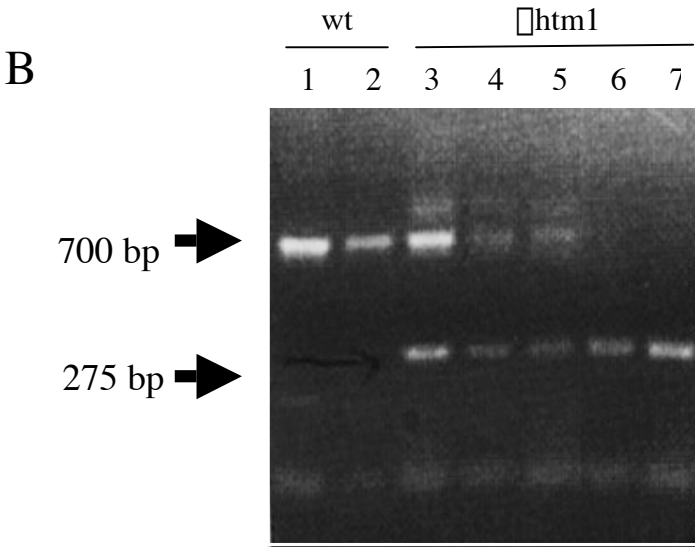
PCR Verification of Gene Deletion

A



Modified from unpublished application protocol of Achim Wach, Arndt Brachat, and Peter Philippsen

PCR Verification of Gene Deletion



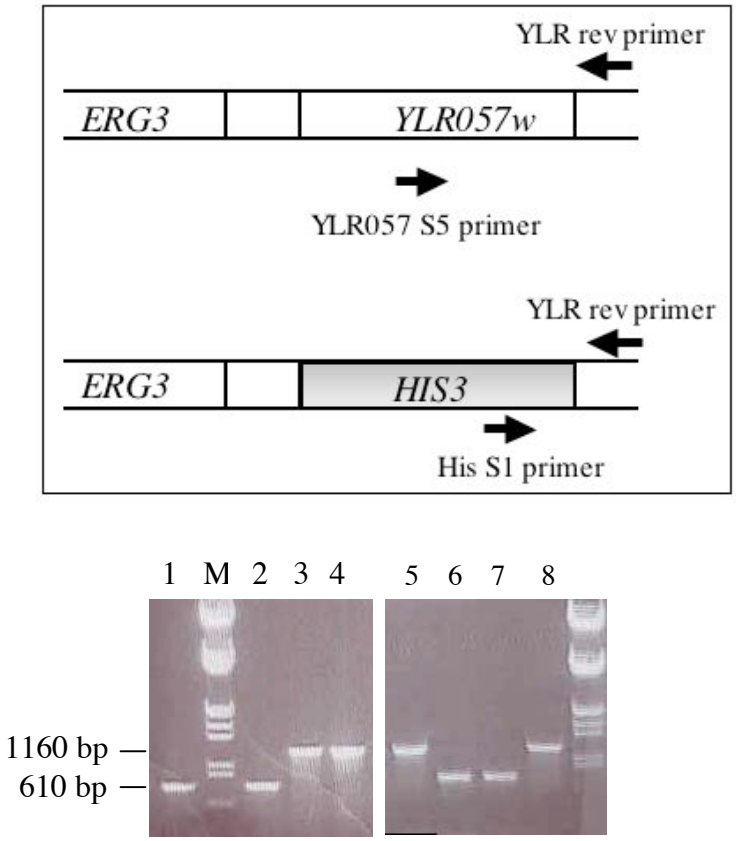


Figure 29. PCR verification of the *YLR057w* disruption in various yeast strains. His S1 primer and YLR reverse primer amplify a 610 bp band indicating the presence of the knockout cassette. YLR reverse primer and YLR057 S5 primer amplify a 1160 bp band indicating the presence of *YLR057w* in the wild-type host. Lane 1 - Δ*ylr057w* Δ*htm1* Δ*mns1* CPY*; Lane 2 - Δ*ylr057w* Δ*mns1* CPY*; Lane 3 - Δ*htm1* Δ*mns1* CPY*; Lane 4 - Δ*mns1* CPY*; Lane 5 – CPY*; Lane 6 - Δ*ylr057w* CPY*; Lane7 - Δ*ylr057w* Δ*htm1* CPY*; . Lane 8 - Δ*htm1* CPY*; M - DNA marker III (Roche).

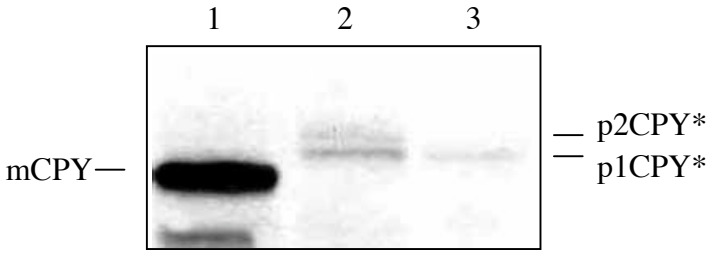


Figure 30. CPY immunoblot from cell extracts of different *S. cerevisiae* strains. Lane 1 is from Δ*htm1* yeast (wild-type CPY). Lane 2 is from Δ*htm1 prc1-1* (CPY*). Lane 3 is from *prc1-1* (CPY*). Equivalent amount of lysate (400μg) was added per lane as determined by protein assay. The blot was probed with anti-CPY antibody (gift of Markus Aebi) (1:2000) followed by anti-rabbit HRP-conjugated secondary antibody (1/20,000). The p1CPY is found in the ER and is 67 kDa. The p2CPY results from modifications in the Golgi and is 69 kDa. The mature form of CPY (mCPY) is 63 kDa.

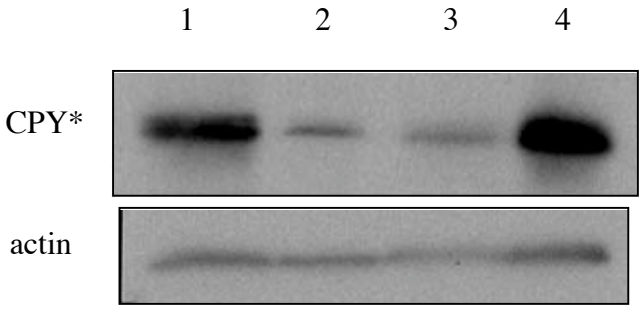
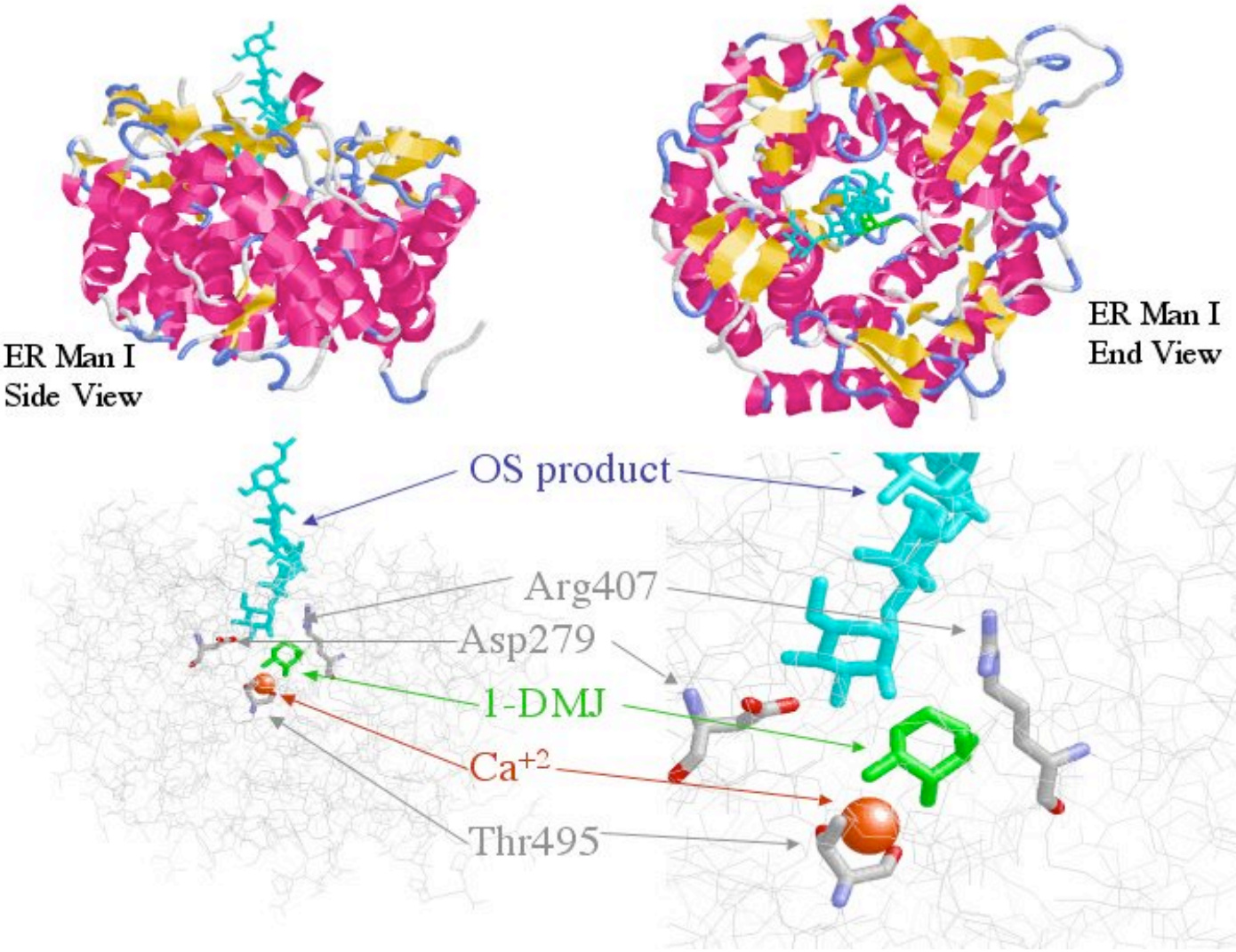


Figure 31. Immunoblot analysis of CPY* in *HTM1* disrupted *S. cerevisiae*. lane 1 - Δ*htm1 CPY**, lane 2 – CPY*, lane 3- Δ*htm1* CPY* + HTM1/p416, lane 4 - Δ*htm1* CPY* + empty p416 vector. Equivalent amount of cell lysate (from 8.4×10^7 cells) was loaded on SDS-PAGE and immunoblotted with antibody to CPY (1:2000) followed by anti-rabbit HRP-conjugated secondary antibody (1/20,000). The blot was reprobed with antibody to actin (1/2000), as a load control, followed by anti-rabbit HRP conjugated secondary antibody.

Figure 32. Location of the residues involved in the mutagenesis study. The top panels show the structure of human ER mannosidase I. The □ helices are pink and the □-sheets are yellow. The oligosaccharide (blue) was modeled into human ER mannosidase I. The bottom left panel is in the same orientation as the top left panel, with the backbone represented as a wire frame. This panel shows the position of three of the conserved residues (highlighted) for which corresponding point mutations were made in *HTM1*. From the crystal structure of ER mannosidase I, the inhibitor, deoxymannojirimycin (1-DMJ), is also shown, indicating the predicted position of the cleaved mannose residue (M10). The bottom right panel shows an enlarged view of the binding pocket.


ER Man I
Side View
ER Man I
End View
OS product
Arg407
Asp279
1-DMJ
Ca+2
Thr495

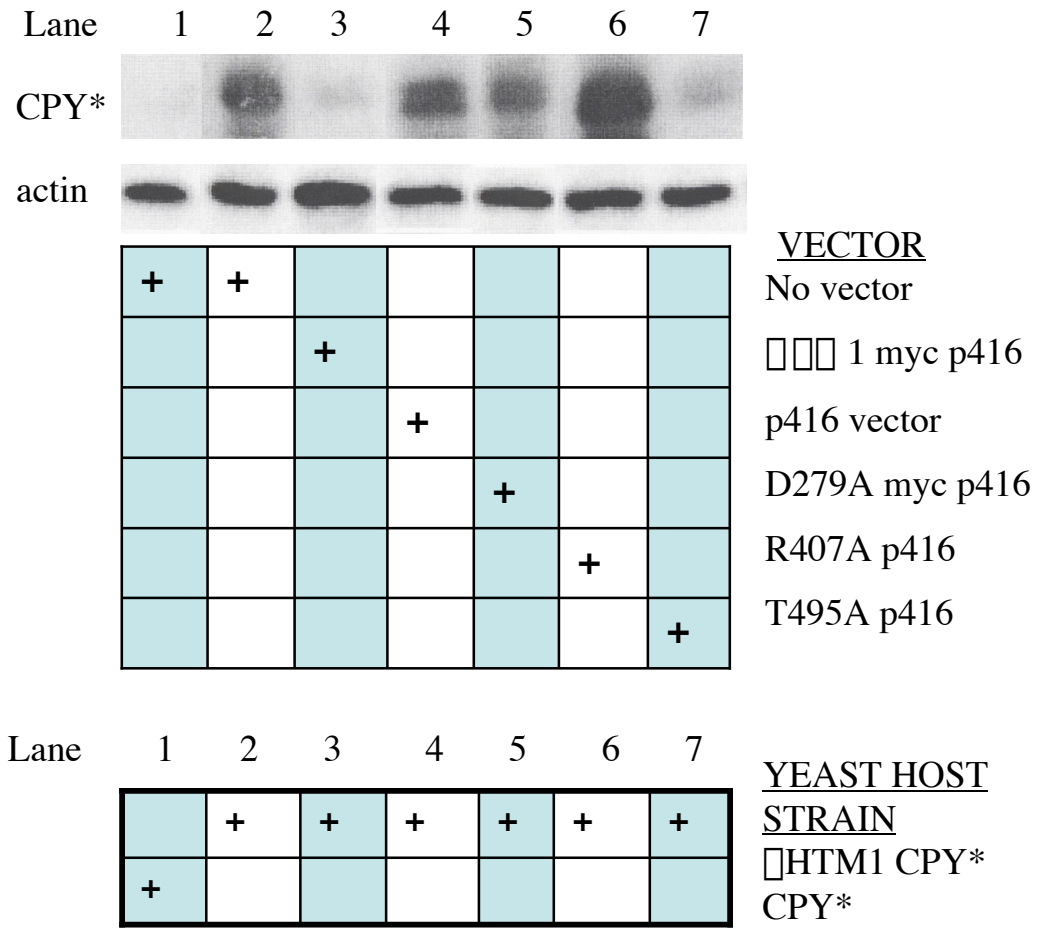


Figure 33. CPY* complementation assay for point mutants. *S. cerevisiae* strain YG618 or □*htm1* (YG618) was transformed with one of the listed vectors. Equivalent amount of cell lysate was loaded on SDS-PAGE and immunoblotted with antibody to CPY, as described in the methods section. The blot was reprobed with antibody to actin, as a load control.

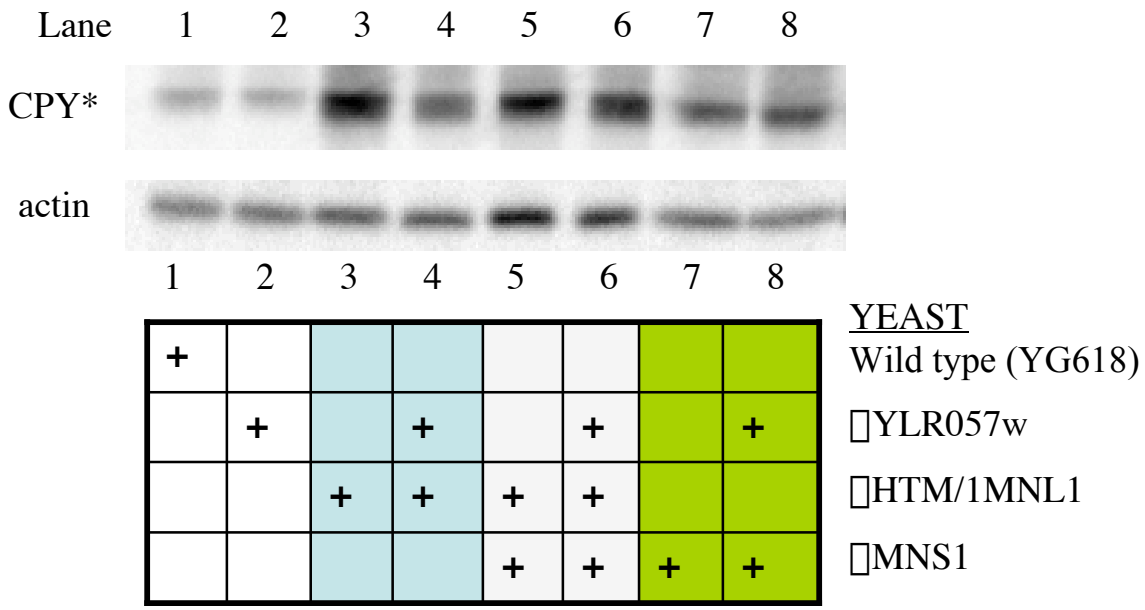


Figure 34. CPY* complementation assay comparing disruptions of *MNS1*, *HTM1*, and *YLR057w*. The knockouts were done in *S. cerevisiae* strains of the same background (YG618). Equivalent amount of cell lysate was loaded on SDS-PAGE and immunoblotted with antibody to CPY, as described in the methods section. The blot was reprobed with antibody to actin, as a load control.

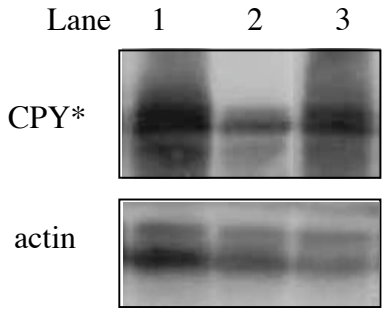


Figure 35. CPY* immunoblot of cell lysate from *DER1* disrupted *S. cerevisiae* strain. Lanes 1-3 show comparison of isogenic strains: lane 1 - Δ*htm1 CPY**, lane 2 – CPY*, lane 3 - Δ*der1 CPY**. Equivalent amount of cell lysate (from $8.4x10^7$ cells) was loaded on SDS-PAGE and immunoblotted with antibody to CPY (1/2000) followed by anti-rabbit HRP conjugated secondary antibody (1/20,000). The blot was reprobed with antibody to actin (1/2000), as a load control.

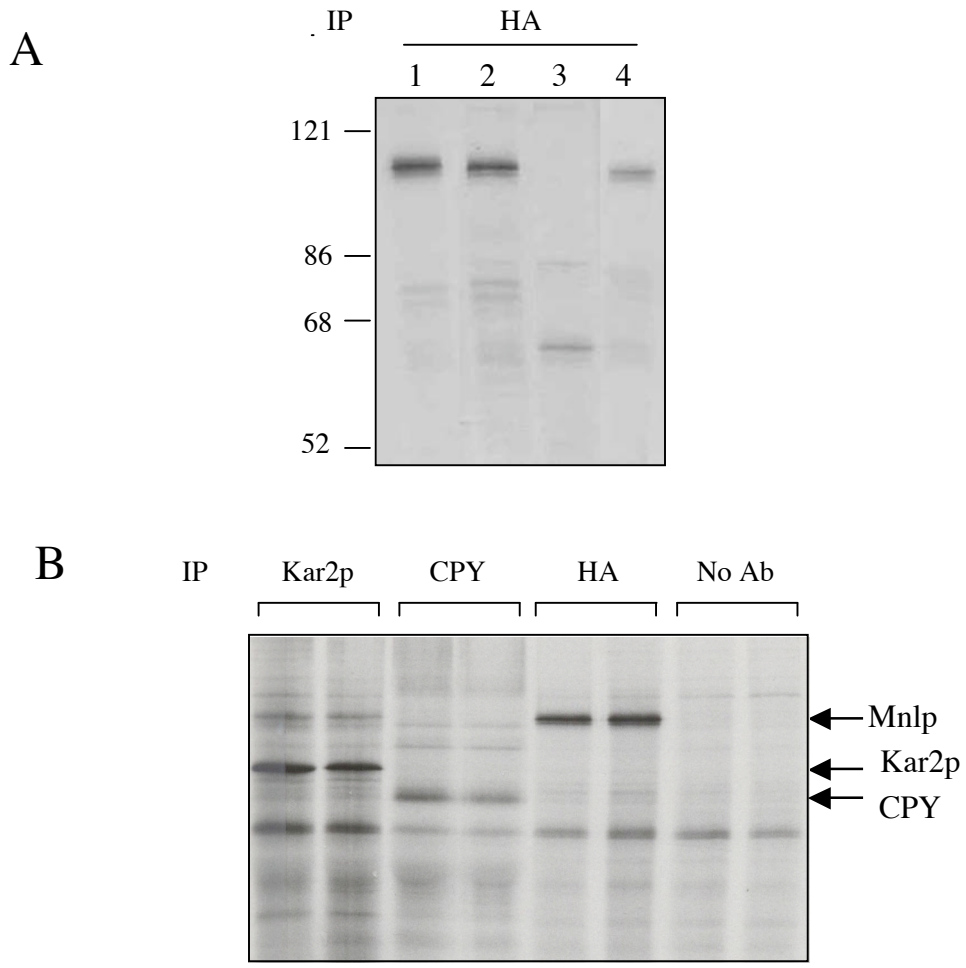


Figure 36. Radiolabel and immunoprecipitation of Htm1p from *S. cerevisiae* cell lysates. (A) As described in methods, $8x10^7$ cells were labeled with ^{35}S-Met/Cys for 30 min, lysed with YPER-S, and clarified cell lysate was immunoprecipitated with antibody to HA. The following strains (+/- vector expressed HTM1) were labeled: lane 1 – CPY* + HTM1-myc-HA-HIS/p416, lane 2 - Δ*ubc7* CPY* + HTM1-myc-HA-HIS/p416, lane 3 - Δ*ubc7* CPY*, Lane 4 - Δ*der1* CPY* + HTM1-myc-HA-HIS/p416. (B) Δ*ubc7* CPY* cells transformed with HTM1-myc-HA-His/p416 were labeled with ^{35}S-Met/Cys for 30 min, lysed with glass beads in 2% CHAPS, HBS, and clarified cell lysate was immunoprecipitated. Samples were loaded in replicates.

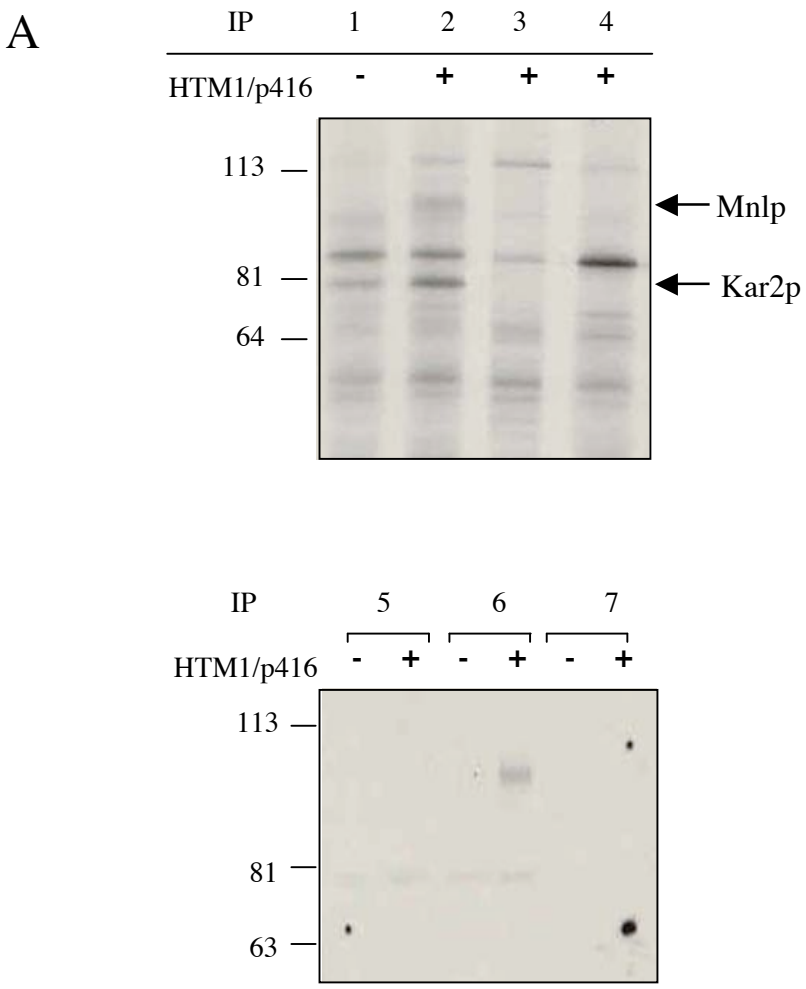


Figure 37. Radiolabel and immunoprecipitation of Kar2p from *S. cerevisiae* cell lysate. Δh*tm1* CPY* cells ($8x10^7$) were transformed with HTM1-myc-HA-His/p416, labeled with ^{35}S-Met/Cys for 30 min, and lysed with glass beads in 2% CHAPS, HBS (50 mM Hepes, 200 mM NaCl). (A) Cell lysate was immunoprecipitated with the following antibodies: IP 1 and IP 2 (Kar2 antibody), IP 3 (no antibody), and IP 4 (Sec61 antibody). (B) Sequential immunoprecipitation from cell lysate. Protein was eluted off of the resin and then immunoprecipitated with a second antibody: IP 5 (Kar2 antibody-->no antibody), IP 6 (Kar2 antibody-->HA antibody), IP 7 (No antibody-->Kar2 antibody).

Figure 38. Nucleotide and amino acid sequence of EDEM2. Highlighted on the protein sequence are the following: predicted cleavable signal sequence (underlined), mannosidase homology domain (bold), the four potential N-glycosylation sites (triangles), and four sequence polymorphisms. All four polymorphisms are different from GenBank NM_018217 (two silent and two non-conservative substitutions, asterisk), while only one silent polymorphism is found in comparison to GenBank AK023931 (diamond).

EDEM2

```
-18  TGAGAGGACACGAGCTCTATGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGTGCGCCAGGT
-6                    M  P  F  R  L  L  I  P  L  G  L  L  C  A  L  L  P  Q  H  H  G  A  P  G

73   CCCGACGGCTCCGCGCCAGATCCCGCCCACTACAGGGAGCGAGTCAAGGCCATGTTCTACCACGCCTACGACAGCTACCTGGAGAATGCC
25   P  D  G  S  A  P  D  P  A  H  Y  R  E  R  V  K  A  M  F  Y  H  A  Y  D  S  Y  L  E  N  A

                                                                  *
163  TTTCCCTTCGATGAGCTGCGACCTCTCACCTGTGACGGGCACGACACCTGGGGCAGTTTTTCTCTGACTCTAATTGATGCACTGGACACC
55   F  P  F  D  E  L  R  P  L  T  C  D  G  H  D  T  W  G  S  F  S  L  T  L  I  D  A  L  D  T

253  TTGCTGATTTTGGGGAACGTCTCAGAATTCCAAAGAGTGGTTGAAGTGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCT
85   L  L  I  L  G  N  V  S  E  F  Q  R  V  V  E  V  L  Q  D  S  V  D  F  D  I  D  V  N  A  S

343  GTGTTTGAAACAAACATTCGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGAGGCTGGATGG
115  V  F  E  T  N  I  R  V  V  G  G  L  L  S  A  H  L  L  S  K  K  A  G  V  E  V  E  A  G  W

433  CCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACTCCTCCCAGCCTTTCAGACCCCCACTGGCATGCCATATGGA
145  P  C  S  G  P  L  L  R  M  A  E  E  A  A  R  K  L  L  P  A  F  Q  T  P  T  G  M  P  Y  G

523  ACAGTGAACTTACTTCATGGCGTGAACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGCCACC
175  T  V  N  L  L  H  G  V  N  P  G  E  T  P  V  T  C  T  A  G  I  G  T  F  I  V  E  F  A  T

613  CTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGTGGCCAGAGTGGCTTTGATGCGCCTCTGGGAGAGCCGGTCAGATATCGGGCTG
205  L  S  S  L  T  G  D  P  V  F  E  D  V  A  R  V  A  L  M  R  L  W  E  S  R  S  D  I  G  L

703  GTCGGCAACCACATTGATGTGCTCACTGGCAAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTG
235  V  G  N  H  I  D  V  L  T  G  K  W  V  A  Q  D  A  G  I  G  A  G  V  D  S  Y  F  E  Y  L

793  GTGAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCCTAGAGTATAACAAAGCCATCCGGAACTACACCCGCTTCGAT
265  V  K  G  A  I  L  L  Q  D  K  K  L  M  A  M  F  L  E  Y  N  K  A  I  R  N  Y  T  R  F  D

883  GACTGGTACCTGTGGGTTCAGATGTACAAGGGGACTGTGTCCATGCCAGTCTTCCAGTCCTTGGAGGCCTACTGGCCTGGTCTTCAGAGC
295  D  W  Y  L  W  V  Q  M  Y  K  G  T  V  S  M  P  V  F  Q  S  L  E  A  Y  W  P  G  L  Q  S

973  CTCATTGGAGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGGGGCTCCCGGAATTCTACAAC
325  L  I  G  D  I  D  N  A  M  R  T  F  L  N  Y  Y  T  V  W  K  Q  F  G  G  L  P  E  F  Y  N

1063 ATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCACTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACG
355  I  P  Q  G  Y  T  V  E  K  R  E  G  Y  P  L  R  P  E  L  I  E  S  A  M  Y  L  Y  R  A  T

                                          *
1153 GGGGATCCCACCCTCCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGATTTGCAACAATCAAA
385  G  D  P  T  L  L  E  L  G  R  D  A  V  E  S  I  E  K  I  S  K  V  E  C  G  F  A  T  I  K

1243 GATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTGGCCGAGACTGTGAAATACCTCTACCTCCTGTTTGACCCAACC
415  D  L  R  D  H  K  L  D  N  R  M  E  S  F  F  L  A  E  T  V  K  Y  L  Y  L  L  F  D  P  T

1333 AACTTCATCCACAACAATGGGTCCACCTTCGACGCGGTGATCACCCCCTATGGGGAGTGCATCCTGGGGGCTGGGGGGTACATCTTCAAC
445  N  F  I  H  N  N  G  S  T  F  D  A  V  I  T  P  Y  G  E  C  I  L  G  A  G  G  Y  I  F  N

1423 ACAGAAGCTCACCCCATCGACCCTGCCGCCCTGCACTGCTGCCAGAGGCTGAAGGAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAA
475  T  E  A  H  P  I  D  P  A  A  L  H  C  C  Q  R  L  K  E  E  Q  W  E  V  E  D  L  M  R  E

                *
1513 TTCTACTCTCTCAAACGGAGCAGGTCGAAATTTCAGAAAAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACACTC
505  F  Y  S  L  K  R  S  R  S  K  F  Q  K  N  T  V  S  S  G  P  W  E  P  P  A  R  P  G  T  L

1603 TTCTCACCAGAAAACCATGACCAGGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCACTTCTCAGCTGCCCCAGTCAGCCCTTCACC
535  F  S  P  E  N  H  D  Q  A  R  E  R  K  P  A  K  Q  K  V  P  L  L  S  C  P  S  Q  P  F  T

1693 TCCAAGTTGGCATTACTGGGACAGGTTTTCCTAGACTCCTCATAA
565  S  K  L  A  L  L  G  Q  V  F  L  D  S  S  *
```

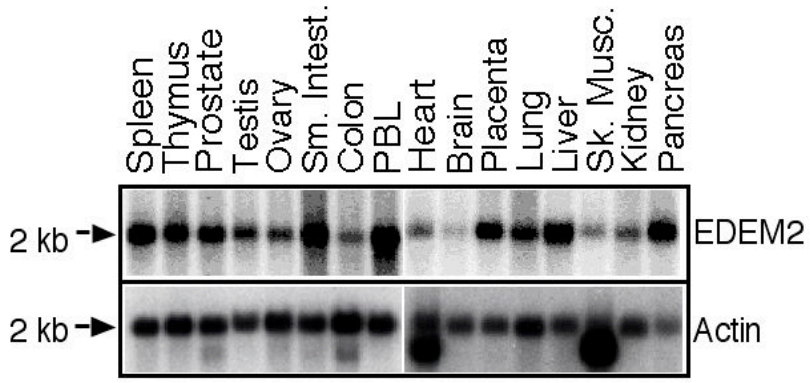


Figure 39. Northern blot of EDEM2. A Northern blot of human poly(A^+) RNAs was hybridized with radiolabeled probes generated from a 1 kb PCR amplimer from the coding region of EDEM2. The blot was rehybridized with □-actin cDNA as a control. The size of the transcripts indicated was estimated based on the electrophoretic mobility of radiolabeled RNA standards. PBL (peripheral blood leukocytes).

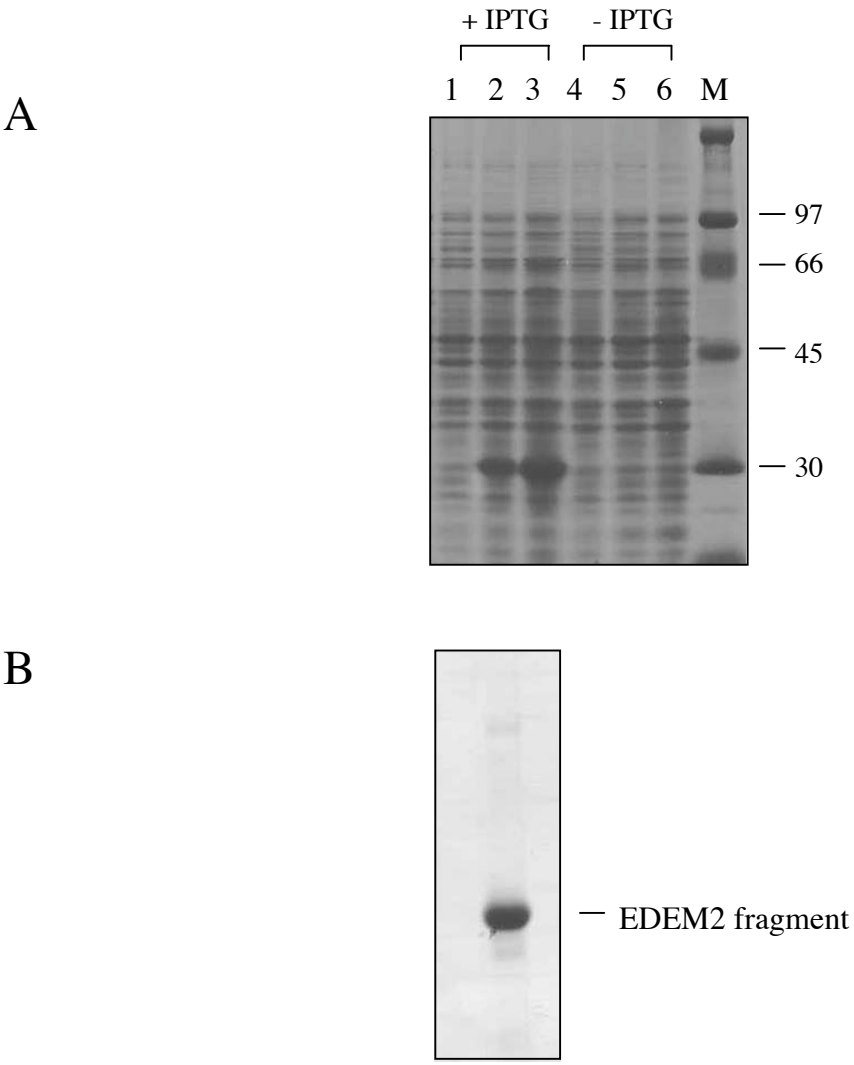


Figure 40. *E. coli* expression of a fragment of EDEM2 with a 6x His tag at the NH_2-terminal end. Expression was induced with 1 mM IPTG for the indicated times. (A) Lanes 1-3 show *E. coli* lysates after induction for 0, 1,5, and 4 h respectively. Lanes 4-6 show *E. coli* lysates without IPTG for 0, 1.5, and 4 h, respectively. Protein was visualized by Coomassie staining. (B) After induction, recombinant protein was purified by Ni^{+2}-NTA chromatography, pooled, and dialyzed. This material was used for rabbit immunization. Protein was visualized by Coomassie staining.

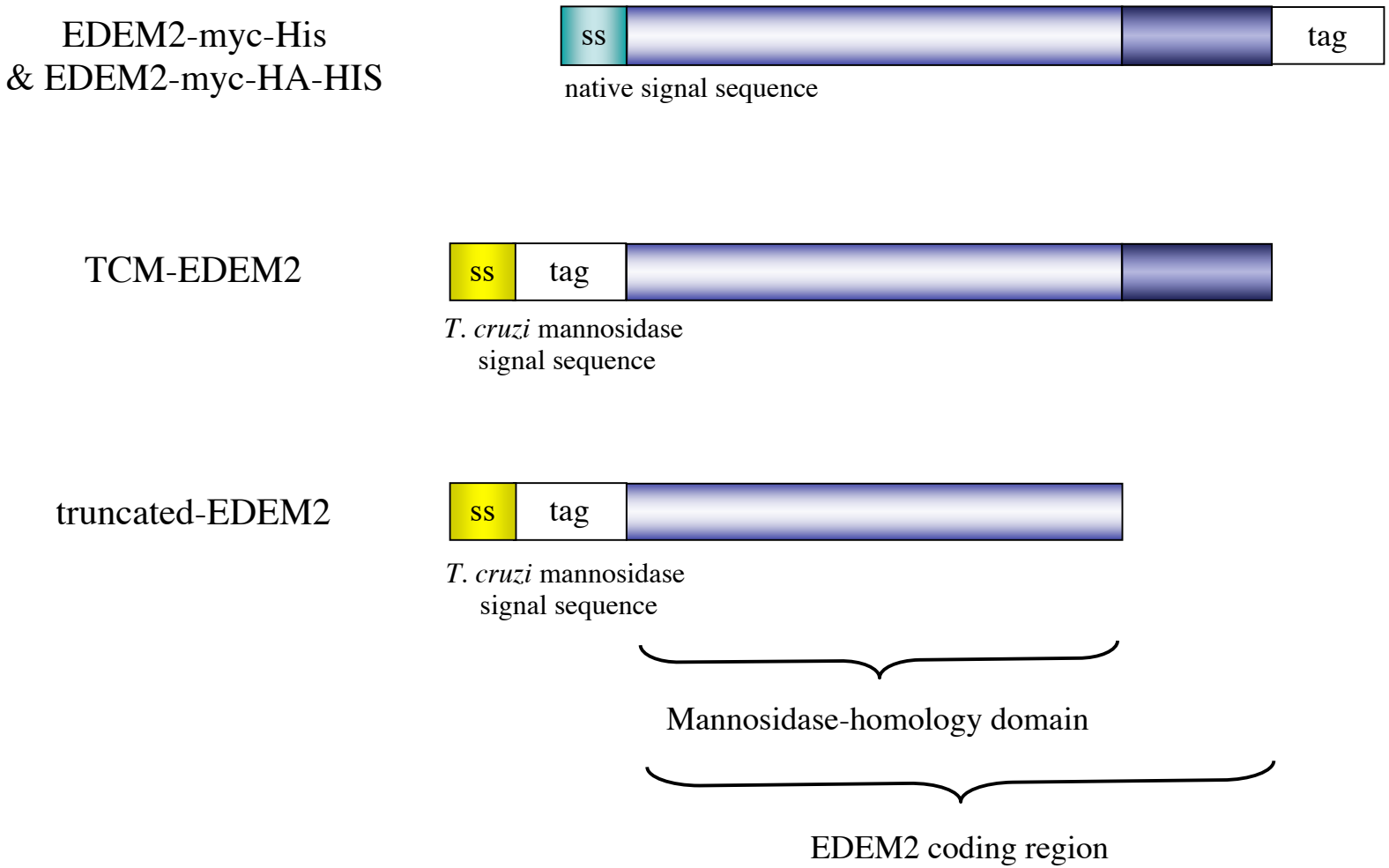
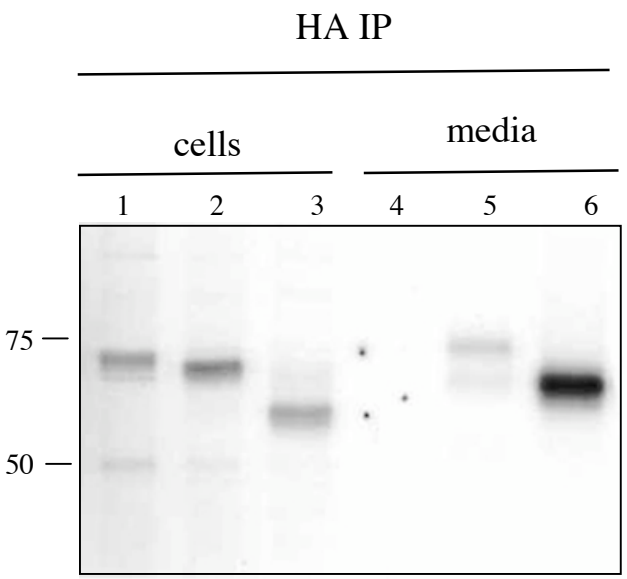


Figure 41. Diagram of the EDEM2 constructs used for characterization studies. EDEM2-myc-His and EDEM2-myc-HA-His are full length constructs with a COOH-tag. TCM-EDEM2 and truncated EDEM2 have the native signal sequence (ss) replaced with the cleavable signal sequence from *T. cruzi* mannosidase. An epitope tag (HA-His) was inserted between the signal sequence and the coding region. The truncated-EDEM2 construct only contains the mannosidase-homology domain (light blue), and not the COOH-terminal extension (dark blue).

A



B

Ratio of secreted protein/intracellular protein	
EDEM2-myc-His	0
TCM-EDEM2	0.64
Truncated-EDEM2	2.6

Figure 42. Intracellular versus secreted localizations of EDEM2.

(A) HEK293 cells stably transfected with EDEM2-myc-HA-His/pEAK (lane 1 and 4), TCM-EDEM2/pEAK (lane2 and 5), or truncated-EDEM2/pEAK (lane 3 and 6) were radiolabeled with ^{35}S-Met/Cys. Cells were lysed in 2% CHAPS, HBS. The cell lysates (equivalent to 1/2 of one 150mm x 25 mm plate) and media (1/2 media from one plate) were immunoprecipitated with antibody to the HA epitope. The samples were separated by SDS-PAGE and detected by autoradiography. (B) From panel A, the levels of secreted protein were compared to the levels of intracellular protein by densitometry.

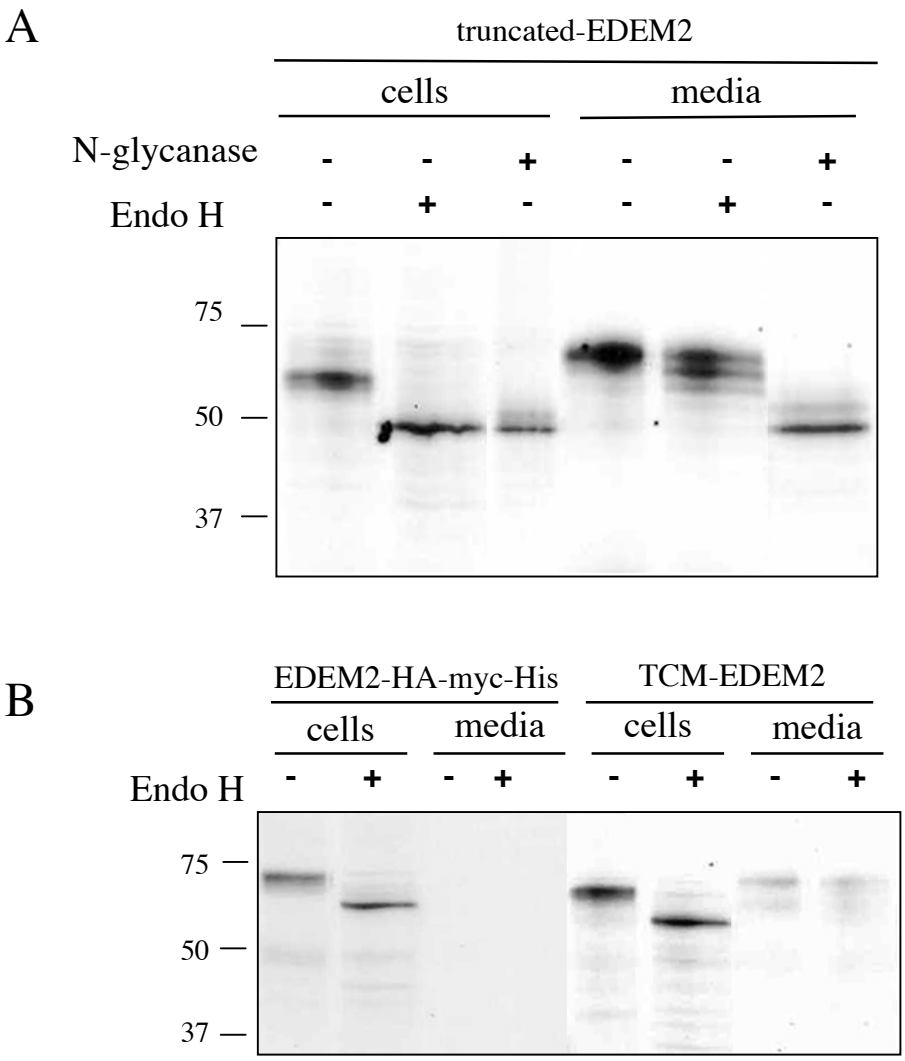


Figure 43. Glycosidase digest of EDEM2. HEK293 cells transfected with truncated-EDEM2 (A), EDEM2-myc-HA-His (B), or TCM-EDEM2 (B) were labeled for 2 h with ^{35}S-Met/Cys and chased for 3 h. Cells were lysed in 2% CHAPS, HBS. The cell lysate (equivalent to 1/3 of one 150mm x 25 mm plate) or media (1/3 of media) was immunoprecipitated with antibody to HA and digested with Endo H or N-glycanase. The samples were separated by SDS-PAGE and detected by autoradiography. When treated with N-glycanase, the bands for TCM-EDEM2 in the media and in the cells collapsed to the same size (not shown).

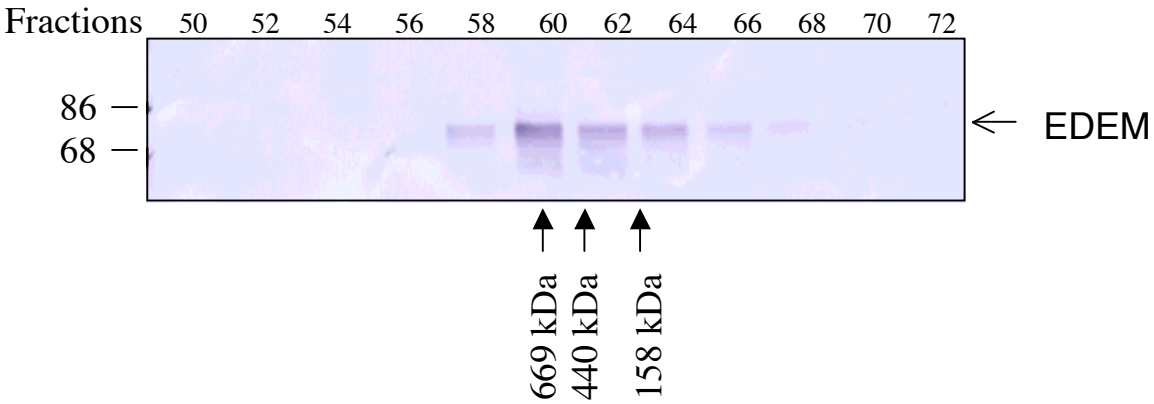
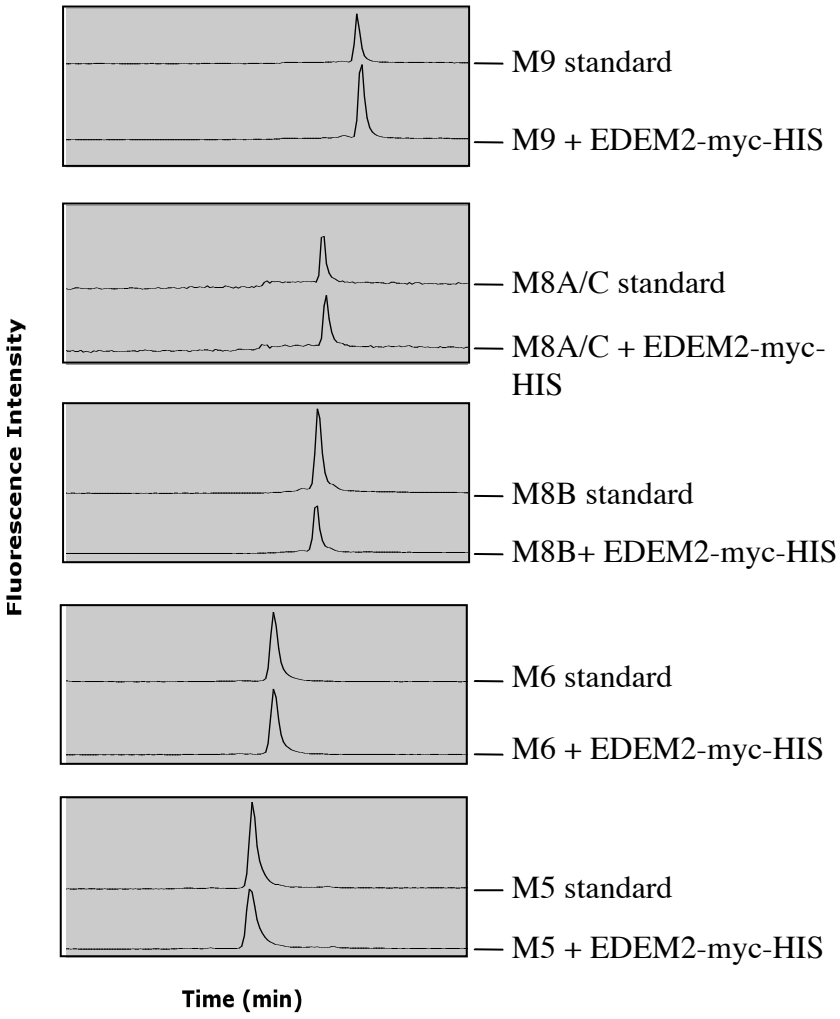


Figure 44. Gel filtration of EDEM2. Ni^{+2}-NTA purified TCM-EDEM2 was run on a Superdex 200 column. Eluted fractions were separated by SDS-PAGE and immunoblotted with antibody to HA. Arrows point to the equivalent positions where the standards were detected by UV trace (thyroglobulin (Tg), 669 kDa; ferritin, 440 kDa; and aldolase, 158 kDa).

Figure 45. α-Mannosidase assay with PA-tagged oligosaccharides. In Panel A, $Man_{9-5}GlcNAc_2$-PA tagged oligosaccharides were incubated with Ni^{+2}-NTA purified EDEM2-myc-His. The samples were resolved by NH_2-bonded HPLC. Peaks were identified by comparison to PA-tagged standards of known structures. In the samples with EDEM2, no shifts in the peaks were observed, as compared to the standards, indicating that EDEM2-myc-His did not cleave the substrate In Panel B, either ER mannosidase I (ER Man I) or Golgi mannosidase IA (Golgi Man IA) was added to the sample containing EDEM2-myc-His and PA-tagged oligosaccharide. Both ER ManI and Golgi Man I added into a sample containing EDEM2-myc-His were able to trim the substrate, indicating the buffer that EDEM2-myc-His was in did not inhibit the reaction.

A


M9 standard
M9 + EDEM2-myc-HIS
M8A/C standard
M8A/C + EDEM2-myc-HIS
M8B standard
M8B+ EDEM2-myc-HIS
M6 standard
M6 + EDEM2-myc-HIS
M5 standard
M5 + EDEM2-myc-HIS
Fluorescence Intensity
Time (min)

B

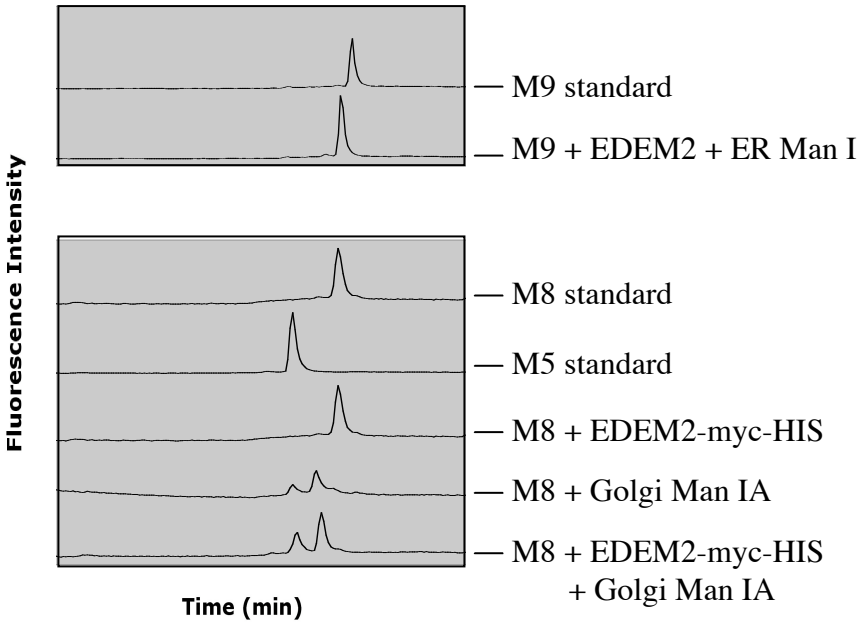
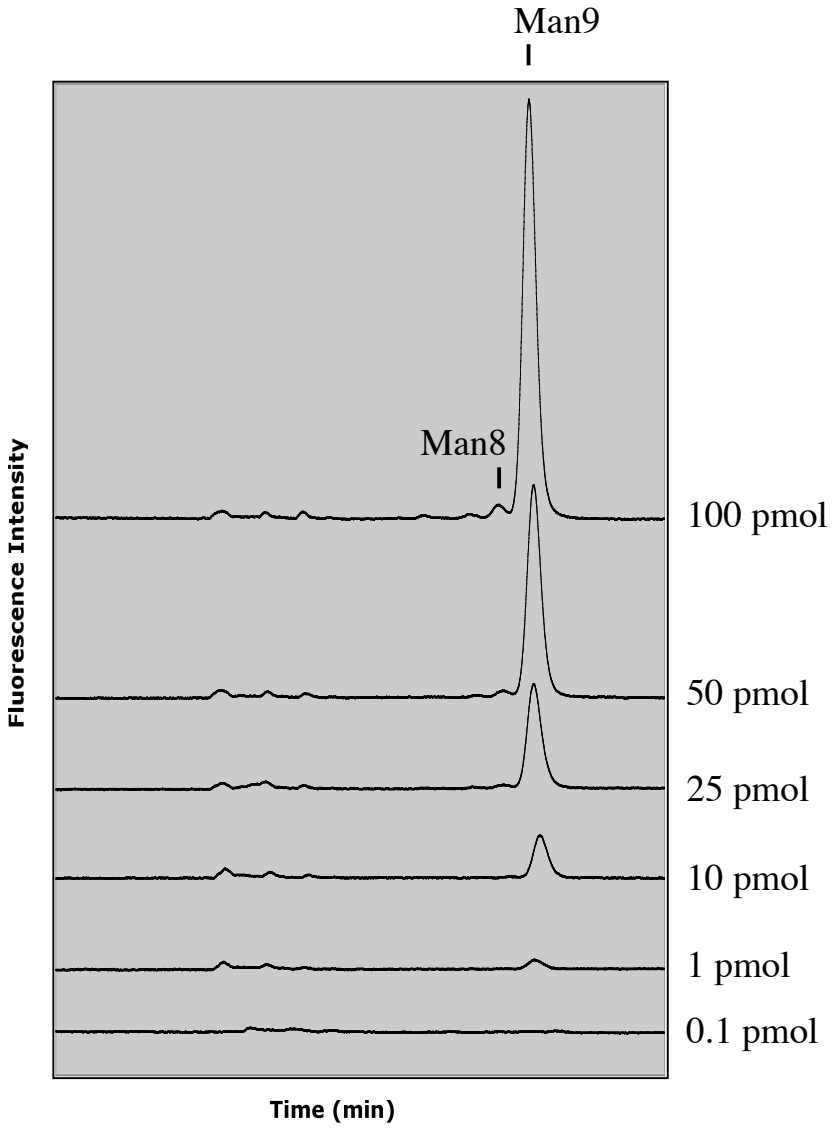

M9 standard
M9 + EDEM2 + ER Man I
M8 standard
M5 standard
M8 + EDEM2-myc-HIS
M8 + Golgi Man IA
M8 + EDEM2-myc-HIS + Golgi Man IA
Fluorescence Intensity
Time (min)

Figure 46. Dilution series of $Man_9GlcNAc_2$-PA. To demonstrate that small amounts of PA-tagged oligosaccharides can be detected in the □-mannosidase assays, 0.1-100 pmol $Man_9GlcNAc_2$-PA was resolved by NH_2-bonded HPLC. For these assays, the detection limit is 1 pmol. The table on the right hand side of the figure shows the area of the corresponding peaks. Additionally, in the 100 pmol sample, a Man_8 peak can be detected. The area of this peak is 1.5% that of the Man_9 peak.



Man9	Peak area
100 pmol	46540
50 pmol	23770
25 pmol	11822
10 pmol	4868
1 pmol	939
0.1 pmol	0

Figure 47. Immunofluorescence studies on HEK293 cells transfected with a C-terminal epitope tagged EDEM2. (A) HEK293 cells were stably transfected with EDEM2-myc-His/pEAK (Panels A-C) or co-transfected with EDEM2-myc-His/pEAK and N_{HK}/pcDNA3.1 (Panels D-P). The myc-epitope tag was detected by indirect immunofluorescence using a mouse monoclonal anti-myc antibody followed by an anti-mouse secondary antibody. Calreticulin was used as the ER marker and was detected by a polyclonal antibody against calreticulin followed by an anti-rabbit secondary antibody. Cycloheximide treatment did not alter protein localization in cells co-transfected with EDEM2-myc-His and N_{HK} (Panels H-J). Cells in Panels K-M were not treated with saponin as a negative control for accessibility of the antibody to the intracellular compartments. Panels N-P were additional negative controls where no primary antibodies were added.

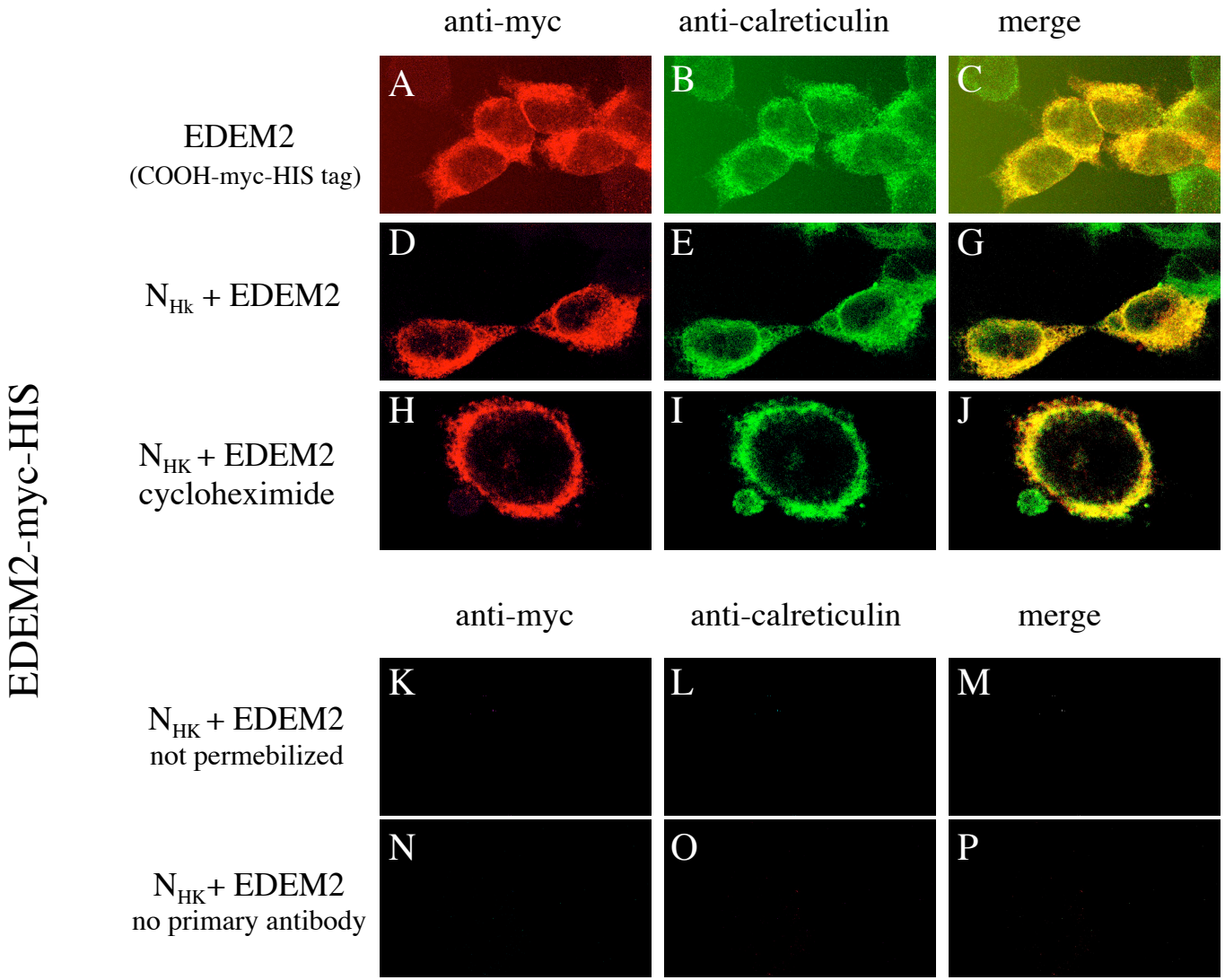

EDEM2-myc-HIS
anti-myc
anti-calreticulin
merge
A
B
C
EDEM2
(COOH-myc-HIS tag)
D
E
G
N_Hk + EDEM2
H
I
J
N_HK + EDEM2
cycloheximide
anti-myc
anti-calreticulin
merge
K
L
M
N_HK + EDEM2
not permebilized
N
O
P
N_HK+ EDEM2
no primary antibody

Figure 48. Immunofluorescence studies on HEK293 cells transfected with TCM-EDEM2/pEAK and PI Z/pcDNA3.1. Panel A, localization of the ER was detected using an anti-rabbit calreticulin antibody followed by Alexafluor 488 anti-rabbit secondary antibody. Panel B, EDEM2 was detected using an anti-mouse HA antibody followed by Alexafluor 594 anti-mouse secondary antibody. Panels C, merged picture of panels A and B. Panel D, light microscopy picture. Panel E, localization of the cytosol was detected using an anti-rabbit cytosolic mannosidase antibody followed by Alexafluor 488 anti-rabbit secondary antibody. Panels F-H follow same order as panels B-D. Panel I, localization of the Golgi was detected using an anti-rabbit Golgi mannosidase II antibody followed by Alexafluor 488 anti-rabbit secondary antibody. Panels J-L follow the same order as panels B-D.

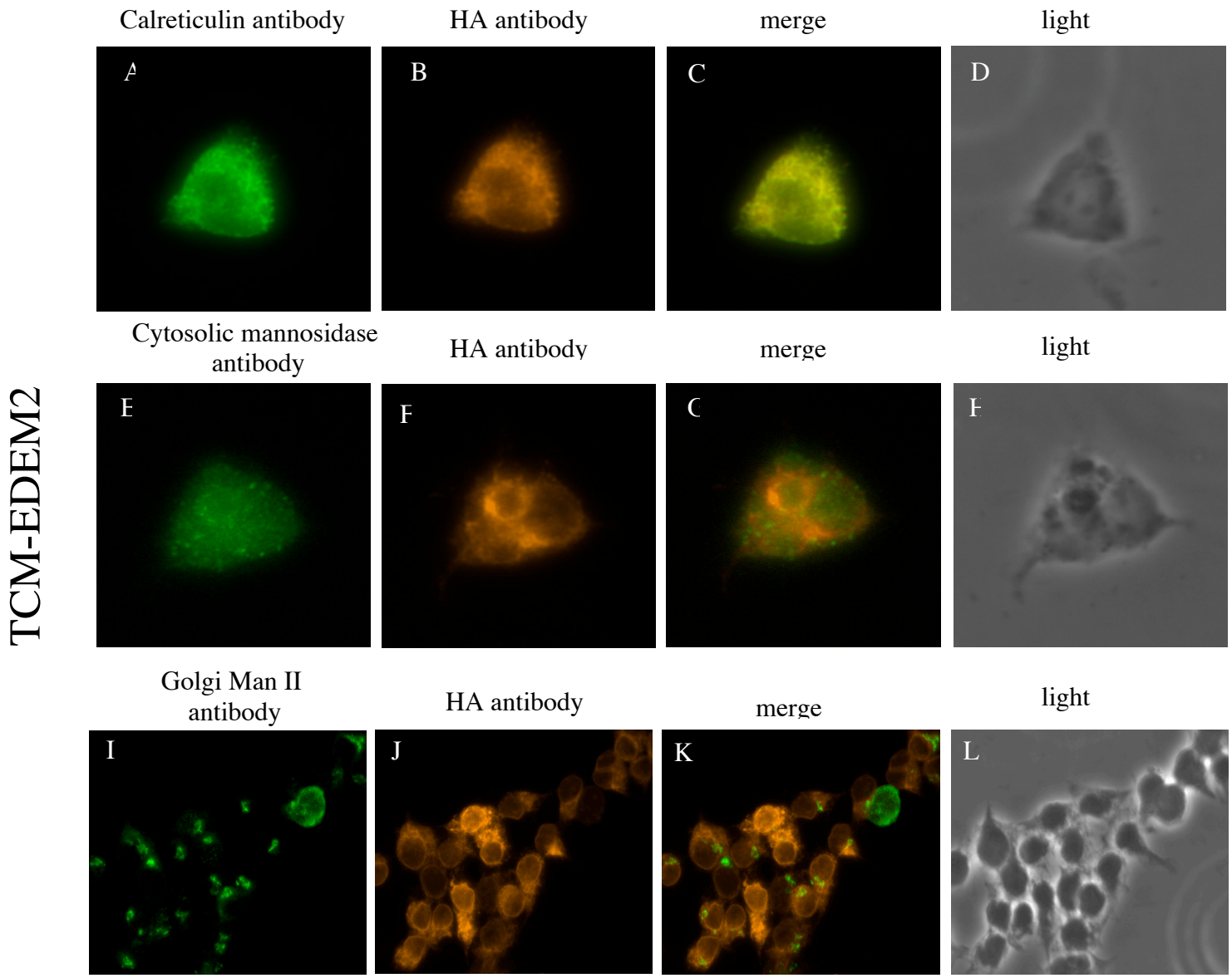
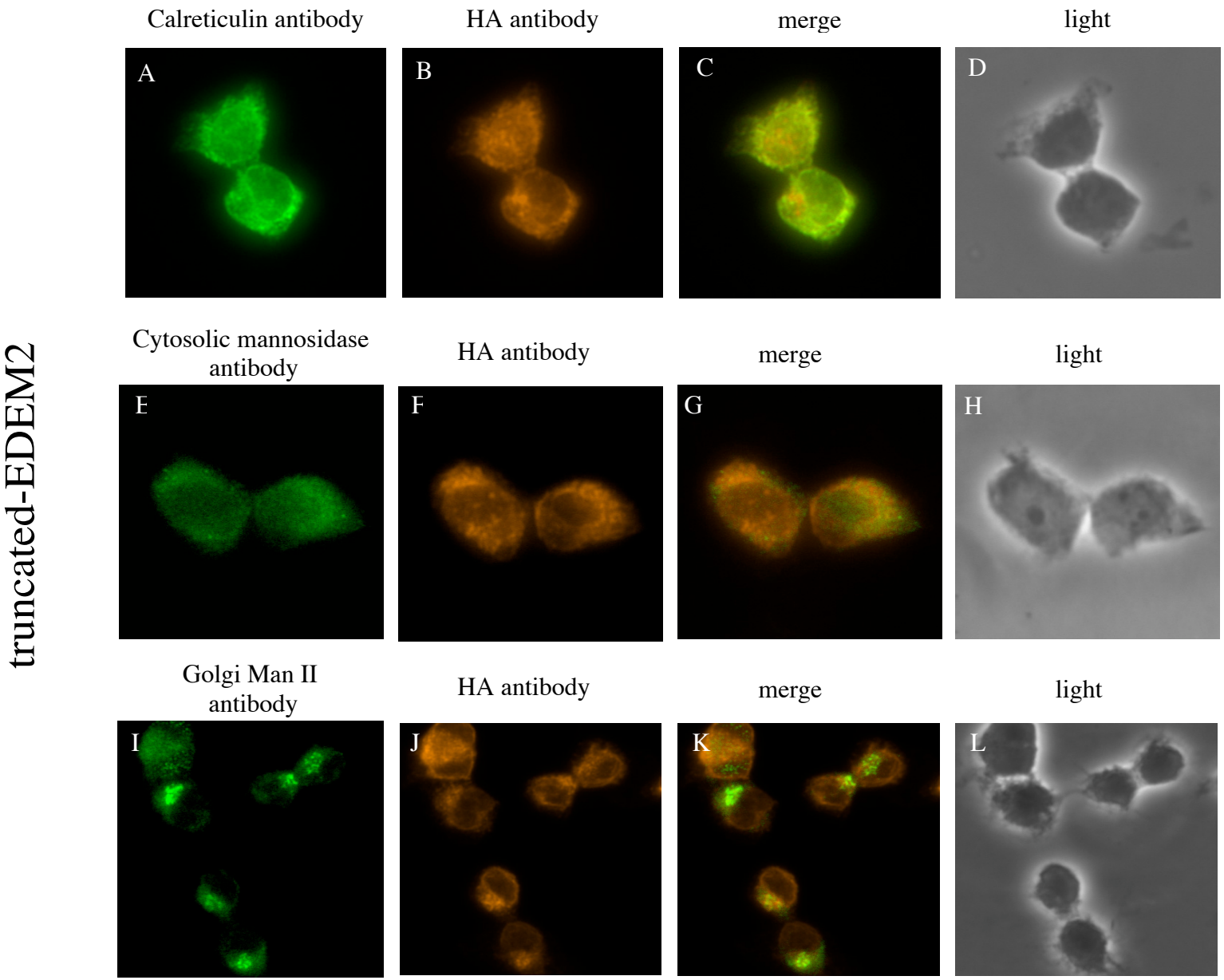

TCM-EDEM2
Calreticulin antibody
HA antibody
merge
light
A
B
C
D
Cytosolic mannosidase antibody
HA antibody
merge
light
E
F
G
H
Golgi Man II antibody
HA antibody
merge
light
I
J
K
L

Figure 49. Immunofluorescence studies on HEK293 cells transfected with truncated-EDEM2/pEAK and PI Z/pcDNA3.1. Panel A, localization of the ER was detected using an anti-rabbit calreticulin antibody followed by Alexafluor 488 anti-rabbit secondary antibody. Panel B, EDEM2 was detected using an anti-mouse HA antibody followed by Alexafluor 594 anti-mouse secondary antibody. Panel C, merged picture of panels A and B. Panel D, light microscopy picture. Panel E, localization of the cytosol was detected using an anti-rabbit cytosolic mannosidase antibody followed by Alexafluor 488 anti-rabbit secondary antibody. Panel I, localization of the Golgi was detected using an anti-rabbit Golgi mannosidase II antibody followed by Alexafluor 488 anti-rabbit secondary antibody.


truncated-EDEM2
Calreticulin antibody
HA antibody
merge
light
A
B
C
D
Cytosolic mannosidase antibody
HA antibody
merge
light
E
F
G
H
Golgi Man II antibody
HA antibody
merge
light
I
J
K
L

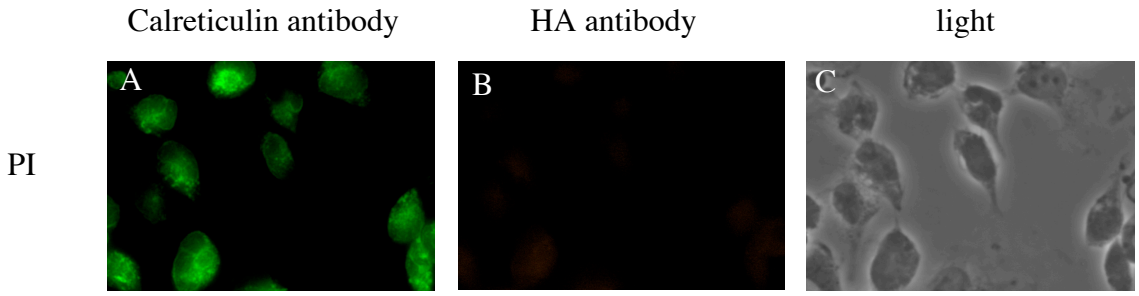


Figure 50. Immunofluorescence control of HEK293 cells transfected with the PI Z/pcDNA3.1 constructs. Panel A, localization of the ER was detected using an anti-rabbit calreticulin antibody followed by Alexafluor 488 anti-rabbit secondary antibody. Panel B, negative control using the anti-mouse HA antibody followed by Alexafluor 594 anti-mouse secondary antibody. Panel C, light microscopy picture.

Figure 51. Immunofluorescence controls. Immunolocalization using HEK293 cells stably transfected with the TCM-EDEM2 (A-C) or truncated-EDEM2 (D-F) and PI Z/pcDNA3.1 constructs. Panels A and D, negative control using only Alexafluor 488 anti-rabbit secondary antibody and no primary antibody. Panels B and E, negative control using only Alexafluor 594 anti-mouse secondary antibody and no primary antibody. Panels C, F, J, and L, light microscopy pictures. Panels H-L, as a negative control for access to intracellular compartments, cells were not treated with saponin. Panel H, anti-rabbit Golgi mannosidase II antibody followed by Alexafluor 488 anti-rabbit secondary antibody. Panel I, anti-mouse HA antibody followed by Alexafluor 594 anti-mouse secondary antibody. Panel K, anti-rabbit calreticulin antibody followed by Alexafluor 488 anti-rabbit secondary antibody.

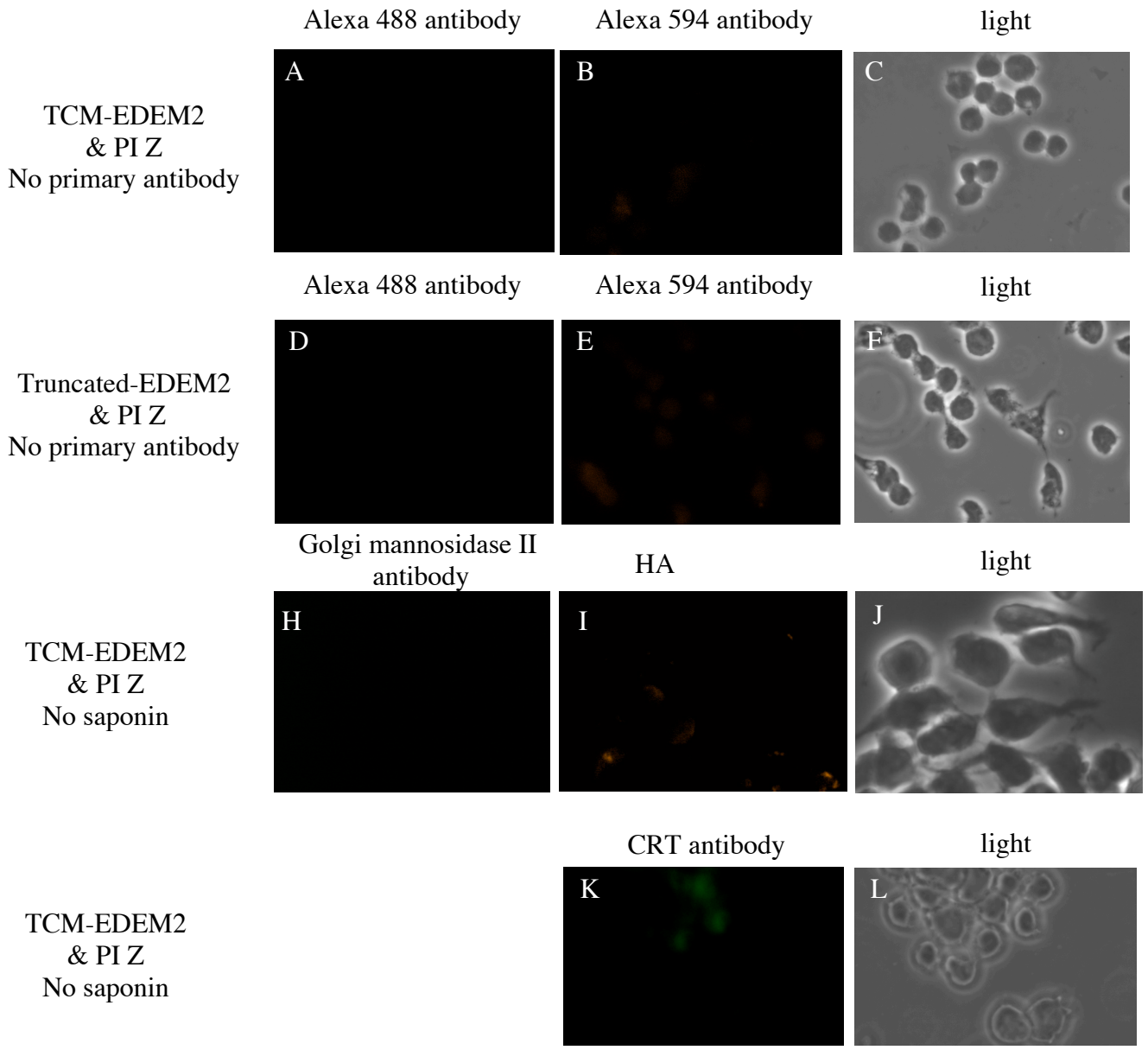
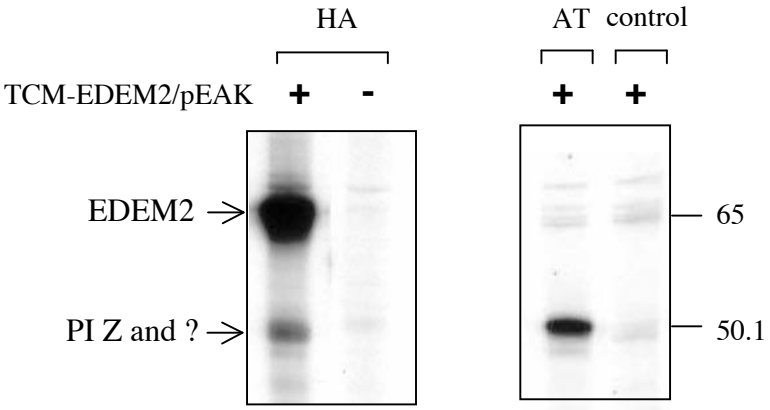
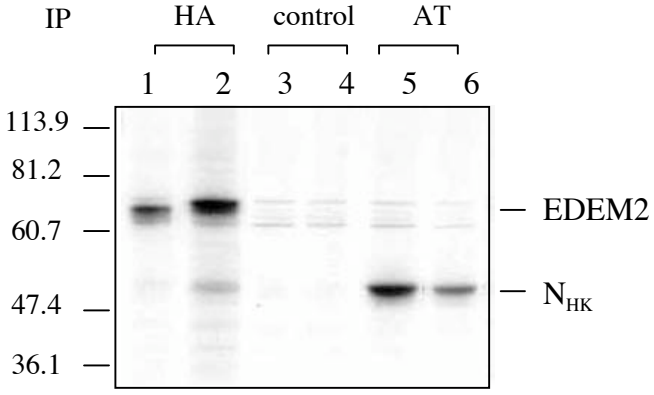

Alexa 488 antibody
Alexa 594 antibody
light
TCM-EDEM2
& PI Z
No primary antibody
A
B
C
Alexa 488 antibody
Alexa 594 antibody
light
Truncated-EDEM2
& PI Z
No primary antibody
D
E
F
Golgi mannosidase II
antibody
HA
light
TCM-EDEM2
& PI Z
No saponin
H
I
J
CRT antibody
light
TCM-EDEM2
& PI Z
No saponin
K
L

Figure 52. Radiolabel and immunoprecipitation of EDEM2 and α1-antitrypsin variants. (A) HEK293 cells stably transfected with PI Z/pcDNA3.1 and TCM-EDEM2/pEAK were radiolabeled 2 h with ^{35}S-Met/Cys and lysed in 2% CHAPS, HBS. The cell lysate was immunoprecipitated with anti-HA, anti-α1-antitrypsin antibody, or control mouse mAb (cdc2). Samples were separated by SDS-PAGE and detected by autoradiography. The band at 50 kDa (PI Z and ?) is also present in cells not expressing PI Z (not shown), indicating it is an unidentified band, partially or entirely other than PI Z. (B) HEK293 cells stably transfected with N_{HK}/pcDNA3.1 and TCM-EDEM2/pEAK (lanes 1, 3, and 5) or EDEM2-myc-HA-HIS/pEAK (lanes 2, 4, and 6) radiolabeled 2 h with ^{35}S-Met/Cys. The cell lysate was immunoprecipitated with anti-HA, anti-α1-antitrypsin antibody, or control mouse mAb. Samples were separated by SDS-PAGE and detected by autoradiography.

A


HA
AT
control
TCM-EDEM2/pEAK
+
-
+
+
EDEM2
PI Z and ?
65
50.1

B


IP
HA
control
AT
1
2
3
4
5
6
113.9
81.2
60.7
47.4
36.1
EDEM2
N_HK

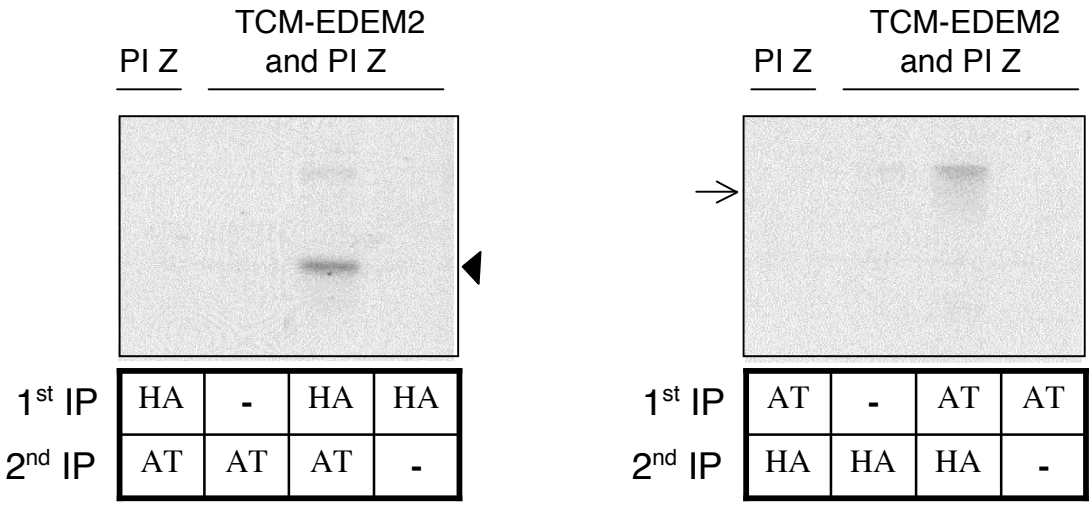
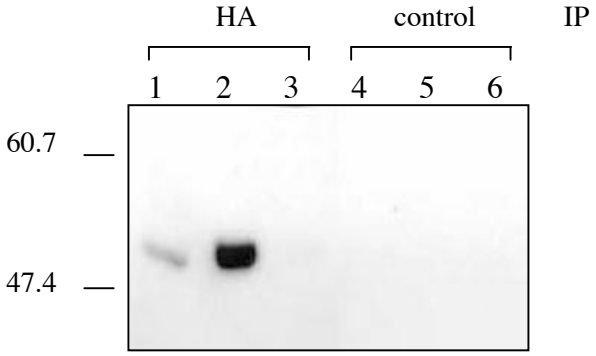


Figure 53. EDEM2-PI Z interaction detected by sequential immunoprecipitation. HEK293 cells transfected with PI Z/pcDNA3.1 ± TCM-EDEM2/pEAK label. In a sequential immunoprecipitation, lysates were immunoprecipitated first with the antibody to □1-antitrypsin; the samples were then eluted and immunoprecipitated with antibody to the HA epitope. Arrows point to EDEM2. Panel B is similar to (A), with HA immunoprecipitation followed by □1-antitrypsin immunoprecipitation. Filled triangles point to □1-antitrypsin.

Figure 54. Immunoprecipitation followed by immunoblotting to detect the EDEM2-α1-antitrypsin interaction. (A) α1-antitrypsin immunoblot after HA immunoprecipitation. HEK 293 cells were co-transfected with either EDEM2-myc-HA-HIS/pEAK and N_{HK}/pcDNA3.1 (lane1 and 6), TCM-EDEM2/pEAK and PI Z/pcDNA3.1 (lane 2 and 5), or PI Z/pcDNA3.1 (lane 3 and 4). Cell lysate was immunoprecipitated with the antibody to □□ or with nonspecific antibody (cdc2), separated by SDS-PAGE, then immunoblotted with the antibody to α1-antitrypsin (AT). (B) HA immunoblot after α1-antitrypsin immunoprecipitation. In HEK 293 cells co-transfected with TCM-EDEM2/pEAK and PI Z/pcDNA3.1, EDEM2 was purified using Ni^{+2}-NTA resin. The purified sample was immunoprecipitated with the antibody to α1-antitrypsin or nonspecific antibody, separated by SDS-PAGE, then immunoblotted with the antibody to HA. Lane 1: control, purified EDEM2, Lane2: negative control IP (goat polyclonal Sec61), Lane 3: α1-antitrypsin IP, Lane 4: α1-antitrypsin IP, no EDEM2 added.

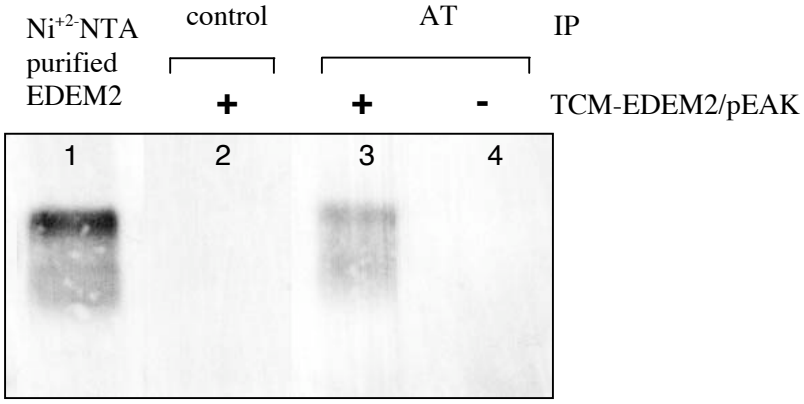
A

□1-AT immunoblot



B

□□ immunoblot



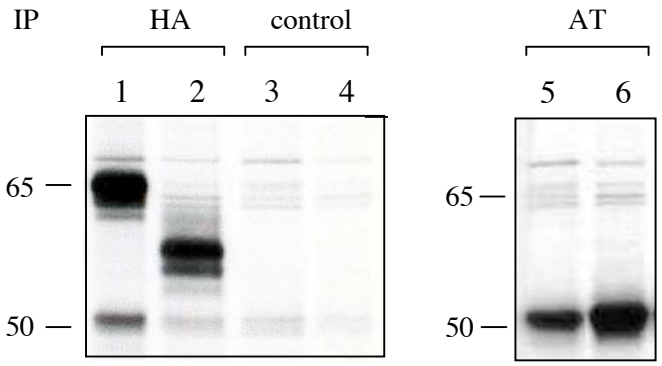
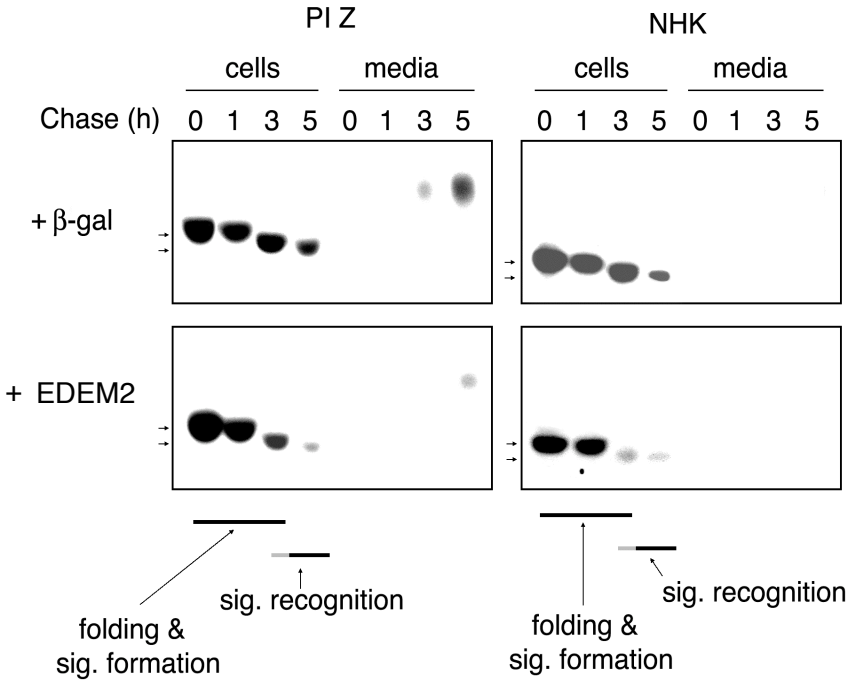


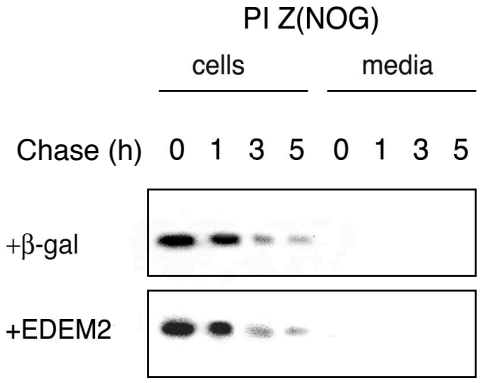
Figure 55. Comparison of TCM-EDEM2 and truncated EDEM2 in co-immunoprecipitation with PI Z. HEK293 cells stably transfected with PI Z and TCM-EDEM2 or truncated-EDEM2 were labeled with ^{35}S-Met/Cys for 2.5 h. 2 mM DTT was added during the last 30 min of the label. Lysate from cells expressing TCM-EDEM2 (lane 1, 3, and 5) or truncated-EDEM2 (lane 2, 4, and 6) was immunoprecipitated from cell lysate with HA antibody, mouse monoclonal control antibody (cdc2), or □1-antitrypsin antibody (AT).

Figure 56. EDEM2 accelerates the degradation of misfolded α1-antitrypsin. (A) Pulse-chase studies were performed in HEK293 cells that were cotransfected with TCM-EDEM and either PI Z or N_{HK} (bottom panels). Control pulse-chase studies employed a β-gal expression construct in place of TCM-EDEM2 (top panels) in combination with either PI Z or N_{HK}. Cells were labeled for 15 min with [^{35}S] Met and chased for the time indicated. Radiolabeled α1-antitrypsin was immunoprecipitated from cell lysates or media, separated by SDS-PAGE and detected by fluorography. The horizontal bars at the bottom of the figure depict the predicted folding and signal formation phase, and the signal recognition phase of ERAD. The grey portion of the bar depicts the rate-limiting step in the overall degradation process, which is affected by the availability of the proposed degradation receptor, EDEM2. (B) Similar to Panel A, left panels, except that the PI Z expression construct was replaced with a non-glycosylated PI Z mutant expression construct, PI Z(NOG). The specificity of EDEM2 for accelerating disposal of unfolded glycoproteins was indicated by the lack of enhancement of degradation of a non-glycosylated form of PI Z. (C) Similar to Panel A, except that the construct encoding truncated EDEM2 was used in place of full-length TCM-EDEM2 in the co-transfections. The full length TCM-EDEM2 protein was able to accelerate the degradation of N_{HK} and cause a partial acceleration of PI Z disposal, but COOH-terminal truncation of EDEM2 resulted in a protein that was ineffective in enhancing mutant α1-antitrypsin disposal.

A


PI Z
NHK
cells
media
cells
media
Chase (h)
0 1 3 5 0 1 3 5
0 1 3 5 0 1 3 5
+ β-gal
+ EDEM2
folding &
sig. formation
sig. recognition
folding &
sig. formation
sig. recognition

B


PI Z(NOG)
cells
media
Chase (h) 0 1 3 5 0 1 3 5
+β-gal
+EDEM2

C

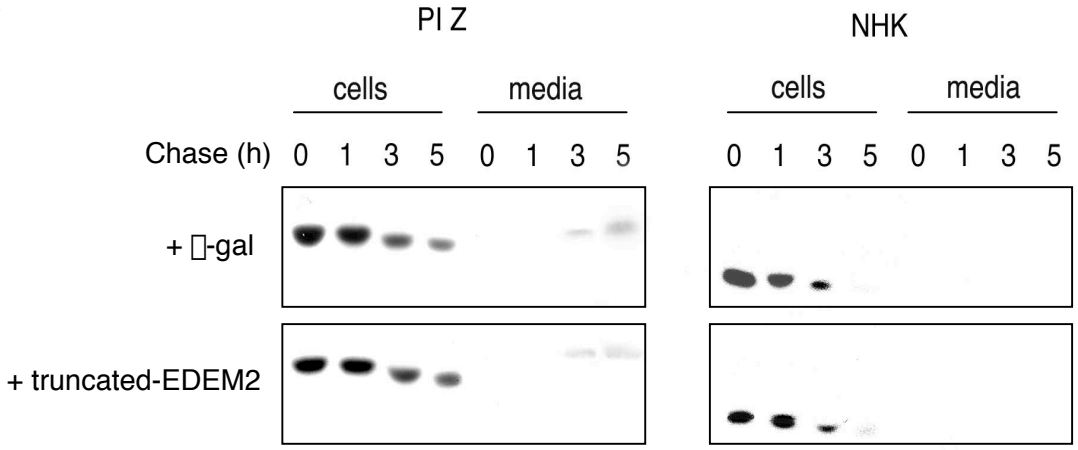

PI Z
NHK
cells
media
cells
media
Chase (h) 0 1 3 5 0 1 3 5
0 1 3 5 0 1 3 5
+ -gal
+ truncated-EDEM2

TABLE 6. *S. cerevisiae* strains

YPH499	*MAT*a *ura3-52 lys2-801amber ade2-101ocre trp1-Δ63 his3-Δ200 leu2-Δ1*
YPH501	*MAT*a/Δ *ura3-52 lys2-801amber ade2-101ocre trp1-Δ63 his3-Δ200 leu2-Δ1*
Δ*htm1*(YPH499)	*MAT*a *ura3-52 lys2-801amber ade2-101ocre trp1-Δ63 his3-Δ200 leu2-Δ1 Δhtm1::kanMX*
Δ*htm1*(YPH501)	*MAT* a/Δ *ura3-52 lys2-801amber ade2-101ocre trp1-Δ63 his3-Δ200 leu2-Δ1 Δhtm1::kanMX*/HTM1
SS328	*MAT*Δ *ade2-101 his3Δ200 ura3-52 lys2*
YG618	*MAT*Δ *ade2-101 his3Δ200 ura3-52 lys2-801 prc1-1*
YG778	*MAT*a *ade2-101 his3Δ200 ura3-52 lys2 Δmns1::kanMX prc1-1*
Δ*htm1*(SS328)	*MAT*Δ *ade2-101 his3Δ200 ura3-52 lys2 Δhtm1::kanMX*
Δ*htm1*(YG618)	*MAT*Δ *ade2-101 his3Δ200 ura3-52 lys2-801 Δhtm1::kanMX prc1-1*
Δ*der1*(YG618)	*MAT*Δ *ade2-101 his3Δ200 ura3-52 lys2-801 Δder1::kanMX prc1-1*
Δ*ylr057w*(YG618)	*MAT*Δ *ade2-101 his3Δ200 ura3-52 lys2-801 Δylr057w::HIS3 prc1-1*
Δ*ylr057w*(YG778)	*MAT*a *ade2-101 his3Δ200 ura3-52 lys2 Δmns1::kanMX Δylr057w::HIS3 prc1-1*
Δ*ylr057w* (Δ*htm1*(YG618))	*MAT*Δ *ade2-101 his3Δ200 ura3-52 lys2-801 Δhtm1::kanMX Δylr057w::HIS3 prc1-1*
Δ*ylr057w*(Δ*htm1* x Δmns1 CPY*)	*MAT? ade2-101 his3Δ200 ura3-52 lys2 Δmns1::kanMX Δhtm1::kanMX Δylr057w::*HIS3 *prc1-1*
W303-1C (YWO 0343)	*ura3-1 his3-11,15 leu2-3,112 trp1-1 ade2-1ocre can1-100 prc1-1*
W303-CQ (YWO 0350)	*ura3-1 his3-11,15 leu2-3,112 trp1-1 ade2-1ocre can1-100 ubc7:LEU2 prc1-1*

CHAPTER 5

DISCUSSION

EDEM Family

HTM1 and EDEM2 are two homologs in a group of mannosidase-like proteins, which we have termed EDEM (ER degradation enhancing □-mannosidase-like protein). EDEM proteins are found throughout the eukaryotic kingdom, and these proteins appear to target chronically misfolded glycoproteins in the ER for degradation by the cytosolic proteasome. *HTM1*, EDEM2, and the other EDEM homologs share sequence similarity to the Class I mannosidases (Family 47 glycosylhydrolases), but the sequence relationship shared between the EDEM homologs and the Class I mannosidases exists only in the mannosidase homology domain, which is the region encoding the $(\square\square)_7$ core structure (Figure 10) (156,157). Most of the Class I mannosidases have an ER lumenal catalytic domain that is linked to the transmembrane domain and cytosolic tail via a stem region (Figure 19). The fungal secreted mannosidases are also Class I mannosidases, but these proteins do not have extended NH_2- or COOH- sequences surrounding the catalytic domain.

Similar to the membrane-bound Class I mannosidases, EDEM1 and *YLR057w* each have an NH_2-extension with a predicted transmembrane domain. In contrast, EDEM2, EDEM3, and *HTM1* have only a few residues before the mannosidase homology domain and appear to have cleavable signal sequences (Figure 19 and Table 5). The major difference between the EDEM members and the □-1,2 mannosidases lies at the carboxyl-terminus. EDEM proteins contain a

carboxyl-terminal extension, which is not present in the catalytically active Class I mannosidases. The COOH-extension varies in length between the EDEM members (Table 5), and the sequence relationship shared by the EDEM homologs that exists in the mannosidase homology domain is not present in the COOH-terminus. For example, the carboxyl-extension of EDEM2 does not share sequence similarity with EDEM1 or EDEM3. Based on motif searches within the COOH-terminal extensions, the EDEM1 extension contains a putative ATP/GTP-binding site motif A (P-loop) (181), EDEM2 has no recognizable motifs, and EDEM3 has a putative protease-associated domain in the middle of the extended tail (182) as well as a COOH-terminal KDEL ER retention motif (183) . *HTM1* contains a putative GPI anchor at the carboxy-terminus (184). By computational analysis, De Groot et al. (184) suggested that the ER localization of Htm1p performed with an HA-tagged construct resulted in a mislocalization of the protein. Yet in our experiments, expression of a similar construct or of untagged Htm1p in Δ*htm1* yeast both restored CPY* to wild-type levels, supporting the data that Htm1p is ER localized. Finally, *YLR057w* has no recognizable motifs in the COOH-extension.

The EDEM orthologs have a high sequence similarity, but the paralogs (EDEM1, EDEM2, and EDEM3) are not as highly conserved. For example, the full-length protein EDEM2 shares 93% identity to its murine ortholog, but only shares 33% and 23% identity with human EDEM1 and EDEM3, respectively. Comparison of sequences using an unrooted tree dendrogram shows that EDEM1, EDEM2, and EDEM3 each form their own clusters (Figure 9). The *Saccharomyces* EDEM members, *HTM1* and *YLR057w* are further removed in sequence similarity, and they do not fall into the metazoan clusters.

Expression and characterization of HTM1

We have isolated and characterized a cDNA clone encoding *S. cerevisiae HTM1*. Full-length cDNA was obtained from genomic DNA and encoded a protein of 796 amino acids with five potential N-linked glycosylation sites (Figure 16). The hydrophobic sequence at the NH_2-terminus may be a cleavable signal sequence, or it may anchor the protein to the ER membrane (figure 17).

Characterization studies were performed using protein expressed from mammalian cells or from *S. cerevisiae*. *P. pastoris* did not provide a suitable host for expression, although, some intracellular expression in *Pichia* was observed after the native signal sequence was replaced with the *S. cerevisiae* □ factor secretion signal (Figure 21). Detection of Htm1p was accomplished either using an epitope tag or using antibodies generated against *E. coli* expressed fragment of Htm1p. The anti-HTM1 antibody worked for immunoblotting but did not have the ability to immunoprecipitate protein.

Using the HTM1-myc-His construct in *S. cerevisiae,* Htm1p had an apparent molecular mass of 97 kDa by SDS-PAGE. Following N-glycanase treatment, the size of the protein matched the predicted size of 92 kDa (Figure 25).

Gene Disruption of *HTM1*

For characterization of the function of *HTM1* in *S. cerevisiae,* the open reading frame was knocked out. PCR based gene disruption was used to generate viable, knockout strains of *HTM1*. Using the full-length coding region as a probe, Southern blots (Figure 27) and PCR-analysis (Figure 28) were used to confirm the gene disruption. In haploid *Saccharomyces* the PCR analysis indicated that the gene had been disrupted. In diploid *Saccharomyces,* PCR

analysis indicated that one copy of the coding region was present, while the second copy had been disrupted by the *kanMX* cassette.

The Δ*htm1* yeast strain displayed a normal morphology and no phenotype was detected when the cells were grown at elevated or lower temperatures or grown on a variety of carbon sources. A phenotype was detected when a Δ*htm1* strain was crossed with a strain containing a mutant of Stt3p. Stt3p is a component of the oligosaccharyl transferase complex, and the phenotype detected in Δ*htm1* x *stt3-7* was a rescue of the s*tt3-7* strain (2). This result suggested that Htm1p is involved in the degradation of the mutant Stt3-7 protein.

The accumulation of the misfolded CPY* model protein has been used in prior studies as an indicator for involvement in the ERAD pathway. In Δ*htm1* yeast, CPY* accumulated in the p1 form (ER) and p2 forms (Golgi) and was not processed to the mature form (vacuole). Wild-type CPY was processed to the mature form (Figure 30). Previous work had shown that inhibition of mannose trimming by ER mannosidase I disruption resulted in an accumulation of CPY*, and from these results it was postulated that a lectin was involved in targeting the misfolded glycoprotein for degradation (1). The result that Htm1p did not have mannosidase activity and that its disruption resulted in accumulation of CPY* suggested it might be that lectin. In addition, the Δ*htm1* Δ*mns1* double knockout did not result in any additive affect on CPY* accumulation, indicating these proteins act in the same pathway.

The gene disruption allowed for complementation assays to be performed, testing the ability of recombinant Htm1p or its point mutants to restore degradation of CPY* in Δ*htm1* cells to levels of CPY* observed in wild-type cells (Figure 31). The vector-encoded HTM1 was able to rescue the cells and restore CPY* to wild-type levels. The addition of the COOH-terminal

epitope tags (myc-His or myc-HA-His) did not interfere with the function of the protein, as they also were able to complement.

Based on the known structures of α-1,2 mannosidases (154,156-158), the point mutations in *HTM1* are all residues which reside in the center of the barrel, where the oligosaccharide interacts with the protein (Figure 32). Threonine 495 corresponds to the Thr that lies at the top of the β hairpin that plugs the barrel. A calcium ion binds to the Thr and to four water molecules that are in turn hydrogen-bonded to other conserved residues. This ion is essential for activity in the Class I mannosidases (156). This point mutant, T495A, was able to function and CPY* did not accumulate in this line. This agrees with the data where the corresponding Thr point mutant in ER mannosidase I did not abrogate the oligosaccharide interaction (personal communication, Khanita Karaveg). It is also possible that in the Thr point mutant, the Ca^{2+} retains its position by using a water molecule to replace Thr and complete the coordination. Two other point mutants in *HTM1* were not able to complement, suggesting that these residues are important for oligosaccharide binding. Aspartic acid 279 corresponds to a putative catalytic residue essential for mannosidase activity and is in a position to interact with mannose (M7) (Figure 15). In ER mannosidase I, mutation of Asp to Asn abolished oligosaccharide binding, and it is thought that this residue is not involved in catalysis (personal communication, Khanita Karaveg). In *HTM1*, the Asp residue would also be expected to be involved in binding to the oligosaccharide. Arginine 407 corresponds to the Arg residue that contacts mannose (M4) and may stabilize the α1,6-arm, and in this manner the Arg407 residue may dictate the position of the oligosaccharide required for cleavage of the glycosidic linkage between M7 and M10 (156). Arg407 corresponds to the residue that also H-bonds to a glycerol, which mimics saccharide binding and presumably occupies the location of the cleaved mannose (156). Additional work is

needed to strengthen the hypothesis that the conserved catalytic residues are not important for function in Htm1p, while the conserved residues at the oligosaccharide binding site are relevant to its function.

In order to see if there was a direct interaction between Htm1p and CPY*, immunoprecipitation from cell lysates was performed. The experiment was performed with the *S. cerevisiae* wild-type strains as well as strains with disruptions in either *UBC7*, the ubiquitin conjugating enzyme, or *DER1* (Figure 36). As CPY* is degraded in wild-type cells, disruption of *DER1* or *UBC7* would allow a pool of CPY* to accumulate. However, no complex with CPY* was observed in immunoprecipitates of HTM1-myc-HA-His from *S. cerevisiae* cell lysates, but an unidentified band at ~80 kDa was co-immunoprecipitated with Htm1p. Sequential immunoprecipitation of Kar2 followed by HA immunoprecipitation indicates that this band is Kar2p (BiP) (Figure 37). Whether this is a chaperone interaction or a functional interaction has yet to be determined.

In *S. cerevisiae*, *HTM1* and *YLR057w* are the only two EDEM homologs, but cells are still viable after deletion of both genes. Unlike *HTM1*, *YLR057w* does not appear to act on CPY*, as the gene disruption of *YLR057w* did not show any effect on the degradation of CPY* (Figure 34). The double knockouts (Δ*ylr057w* Δ*htm1* or Δ*ylr057w* Δ*mns1*) and the multiple knockout (Δ*ylr057w* Δ*htm1* Δ*mns1*) did not have any additional effect on CPY* degradation. The function of YLR057w has not been identified, and as YLR057w contains low sequence similarity compared to the other EDEM proteins, we have classified YLR057w as an EDEM-like protein.

Expression and characterization of EDEM2

We have cloned, expressed, and characterized a human cDNA encoding EDEM2, which is the gene product of chromosome 20 open reading frame 31 (C20orf31). The sequence is similar to GenBank™ AK023931, containing only one silent base pair polymorphism. EDEM2 is a ubiquitously expressed (Figure 39) member of the EDEM subfamily, which is conserved throughout eukaryotic organisms. Like the other members of the EDEM subfamily, EDEM2 is a protein that is similar in sequence to Class I mannosidases (Family 47 glycosylhydrolases) (Figure 18).

The polyclonal antibodies generated against *E. coli* expressed protein, the 506 MAP peptide, and the NH2 MAP peptide were able to detect recombinant EDEM2 by immunoblotting. The 506 MAP peptide antibody did not have the ability to immunoprecipitate protein (the NH_2 MAP peptide antibody was not tested), but the myc and HA antibodies were able to immunoprecipitate epitope-tagged protein.

Recombinant EDEM2 expressed in HEK293 cells generated ER-localized proteins of 70 kDa and 71 kDa, from NH_2- and COOH-terminal tagged constructs, respectively, as determined by SDS-PAGE. EDEM2 contains four N-linked glycosylation sites, and complete cleavage of these glycans by endoglycosidase H indicates that EDEM2 has a pre-Golgi localization (Figure 43). By gel filtration, EDEM2 exhibits an apparent molecular mass of ~600 kDa (Figure 44), suggesting that EDEM2 acts as an oligomer and/or forms a complex with other proteins, such as calnexin, α1-antitrypsin, or other components of the ERAD system.

No α-mannosidase activity was detected for NH_2- or COOH-tagged recombinant protein (Figure 45). Although it is possible that $Man_{9-5}GlcNAc_2$ are not the proper substrates, the results are consistent with a previous study in *S. cerevisiae* indicating no change in glycan processing

profiles between wild-type and Δ*htm1* strains (2). Similarly, no α-mannosidase activity was detected for mouse EDEM1 (96). The reason why the conserved αα-mannosidase homology domains of EDEM2 and other EDEM proteins do not exhibit catalytic activity has yet to be determined. Sequence similarity between EDEM2 and the Class I mannosidases lies predominantly in the α helical segments of the $(\alpha\alpha)_7$ barrel structure (156,157). Residues within the catalytic site at the core of the $(\alpha\alpha)_7$ barrel also appear to be conserved, while conservation in primary sequence is lost in the loop regions connecting the helices (Figure 18). One potential explanation for the lack of hydrolase activity is that EDEM members lack conserved cysteine residues that form a disulfide bridge between two of the helices in the $(\alpha\alpha)_7$ barrel (156,185). However, the *T. reesei* α-1,2-mannosidase also lacks these conserved cysteines residues and is still an active mannosidase, indicating that the disulfide is not essential for hydrolase activity (158).

Supporting a role for the EDEM proteins in ERAD, Htm1p was found in the ER lumen (3) and the three human homologs all appear to be localized to the ER, but each by a different mechanism. EDEM1 is predicted to be a type-II ER transmembrane protein. EDEM3 contains the COOH-terminal KDEL ER retrieval motif. EDEM2 is predicted to be a soluble protein and has no recognizable ER retention signal but was found to colocalize with the ER marker, calreticulin (Figures 47-49). EDEM2 contains a hydrophobic sequence at the NH_2-terminus predicted to be a cleavable signal sequence (Figure 17), although if this sequence acted as an uncleaved signal anchor it could account for the membrane association and retention. Exchanging the putative cleavable signal from EDEM2 sequence for the *T. cruzi* αα-mannosidase cleavable signal sequence does not alter the localization pattern of EDEM2 and argues against the idea that the hydrophobic sequence at the NH_2-terminus is responsible for ER retention. In

contrast, localization of EDEM2 is influenced by the COOH-terminal extension beyond the mannosidase homology domain. While localization is still mainly in the ER for the truncated-EDEM2 construct, a large percentage of the truncated construct is secreted, in contrast to the full-length EDEM2 that is retained in the ER (Figure 42).

Similar to murine EDEM1 (96), EDEM2 is able to interact with the misfolded α1-antitrypsin variants, N_{HK} and PI Z (Figures 52-54). The interaction with the substrate most likely includes a glycan component, as demonstrated by the ability to accelerate the degradation of α1-antitrypsin mutants but not of an unglycosylated mutant α1-antitrypsin (Figure 56). However, the interaction does not appear to be solely dependent on the oligosaccharide moiety. EDEM2 and α1-antitrypsin variants still interact after treatment with tunicamycin (not shown), indicating that a complex can be maintained and possibly formed in the absence of glycan structures on the ERAD substrate. A similar result was seen when murine EDEM could be found in association with BACE 457α after treatment with N-glycanase (100), although the co-immunoprecipitation could also be a result of a non-specific aggregation.

In contrast to the observation that both the NH_2- and the COOH-terminal tagged forms of EDEM2 were able to bind N_{HK} and PI Z, only the NH_2-tagged construct was able to accelerate the degradation of the misfolded α1-antitrypsin variants (data not shown). Similarly, COOH-terminal truncation beyond the mannosidase-homology domain eliminated the ability of EDEM2 to accelerate mutant α1-antitrypsin degradation. One possible explanation for these data is that the COOH-domain has a role in the interaction with other proteins in the ERAD pathway and that the extended tag sequence at the COOH-terminus interferes with this function.

We noted that the degradation of the α1-antitrypsin variant, N_{HK}, occurs more quickly than for PI Z. One explanation for the delayed rate of disposal for PI Z is the ability of this

mutant to spontaneously generate loop-sheet polymers that may reduce the accessibility of protein monomers to be targeted for degradation (186). Although the rate of polymerization has been reported to be low in HEK293 cells (95), aggregate formation could still influence the rate of entry into the disposal machinery.

The SDS-PAGE mobility of radiolabeled α1-antitrypsin was found to increase during the chase period and is complete by 3 h chase (Figure 56). This shift to a lower apparent molecular mass is consistent with glycan processing and is coincident with disposal of the radiolabeled protein. This is in agreement with the current model of ERAD where ER mannosidase I trimming of $Man_9GlcNAc_2$ structures to $Man_8GlcNAc_2$ generates the glycan signal for degradation. The data would also be consistent with the recent data indicating that further processing to $Man_{7-5}GlcNAc_2$ structures may be the alternate glycan signal in mammalian cells (14,187-190).

In summary, the EDEM subfamily of the Family 47 glycosylhydrolase is found throughout the eukaryotic kingdom and consists of proteins with amino acid sequence similarity to the Class I mannosidases. Deletion of *HTM1* resulted in an accumulation of the misfolded CPY* protein, and overexpression of Htm1p returned CPY* to normal levels. EDEM2 interacted with misfolded α1-antitrypsin and accelerated its disposal, while truncated-EDEM2 and COOH-tagged EDEM2 could not accelerate the degradation of mutant α1-antitrypsin. Also, initial mutagenesis studies on conserved residues in the catalytic and oligosaccharide binding site agrees with the hypothesis that the catalytic site is not important for Htm1p function, while conserved residues at the binding site are important for function. The data from Mn11p, EDEM1, and EDEM2 suggests that the EDEM group consists of catalytically inactive proteins that act to target misfolded glycoproteins to the ER-associated degradation pathway. In

Saccharomyces, if YLR057w is an EDEM protein, it appears to have a different substrate specificity than Htm1p, while the human homologs, EDEM1 and EDEM2, displayed overlapping functions with the α1-antitrypsin variant, N_{HK}.

To further investigate the biological functions of this new subfamily of mannosidase-like proteins, studies must be done to test the hypothesis that the EDEM proteins are lectins and to determine the structure, examine the mechanism of action, and differentiate the roles of the EDEM homologs. The differences in sequence between the EDEM homologs suggest overlapping, but specific roles for each member. In characterizing this subfamily, the role of the COOH-extension after the mannosidase homology domain will be a major factor in determining the mechanism of action. For EDEM2, the COOH-terminal extension appears to be involved in both ER retention and in the degradation of substrates.

For EDEM1, its role in ERAD may involve a transfer of misfolded proteins from calnexin to the putative lectin. Whether the other homologs follow this pathway is unknown. Since the present model of EDEM1 interaction with calnexin is through its transmembrane domain, the fact that EDEM2 and EDEM3 are proposed to be soluble proteins would suggest that a similar mechanism for interaction with calnexin is not likely to pertain for these latter proteins. The nature of the lectin activity of the EDEM homologs needs to be further characterized, since a direct role of a specific glycan structure in the initiation or maintenance of the interaction remains to be demonstrated. Surface plasmon resonance studies may useful for this characterization, although EDEM members may only recognize the glycan in the context of a misfolded protein. For structural studies, production of a secreted protein is necessary to generate the necessary amount of protein at a reasonable cost. Mouse knockout models may aid in determining the *in vivo* functions of the mammalian EDEM homologs.

REFERENCES

1. Jakob, C. A., Burda, P., Roth, J., and Aebi, M. (1998) *J Cell Biol* **142**, 1223-1233
2. Jakob, C. A., Bodmer, D., Spirig, U., Battig, P., Marcil, A., Dignard, D., Bergeron, J. J., Thomas, D. Y., and Aebi, M. (2001) *EMBO Rep* **2**, 423-430
3. Nakatsukasa K, N. S., Hosokawa N, Nagata K, Endo T. (2001) *J Biol Chem.* **276**, 8635-8638
4. Sitia, R., and Braakman, I. (2003) *Nature* **426**, 891-894
5. Knop M., Finger A., Braun T., Hellmuth K., and DH., W. (1996) *EMBO J.* **15**, 753-763
6. Dempski, R. E., Jr., and Imperiali, B. (2002) *Curr Opin Chem Biol* **6**, 844-850
7. Weng, S., and Spiro, R. G. (1996) *Glycobiology* **6**, 861-868
8. Karaivanova, V. K., Luan, P., and Spiro, R. G. (1998) *Glycobiology* **8**, 725-730
9. Moremen, K. W., Trimble, R. B., and Herscovics, A. (1994) *Glycobiology* **4**, 113-125
10. Dairaku, K., and Spiro, R. G. (1997) *Glycobiology* **7**, 579-586
11. Puccia, R., Grondin, B., and Herscovics, A. (1993) *Biochem J* **290 (Pt 1)**, 21-26
12. Herscovics, A. (1999) *Biochim Biophys Acta* **1473**, 96-107
13. Weng, S., and Spiro, R. G. (1996) *Arch Biochem Biophys* **325**, 113-123
14. Hosokawa, N., Tremblay, L. O., You, Z., Herscovics, A., Wada, I., and Nagata, K. (2003) *J Biol Chem* **278**, 26287-26294
15. Roth, J., Brada, D., Lackie, P. M., Schweden, J., and Bause, E. (1990) *Eur J Cell Biol* **53**, 131-141

16. de Virgilio, M., Kitzmuller, C., Schwaiger, E., Klein, M., Kreibich, G., and Ivessa, N. E. (1999) *Mol Biol Cell* **10**, 4059-4073

17. Vassilakos, A., Michalak, M., Lehrman, M. A., and Williams, D. B. (1998) *Biochemistry* **37**, 3480-3490

18. Hammond, C., Braakman, I., and Helenius, A. (1994) *Proc Natl Acad Sci U S A* **91**, 913-917

19. Sousa, M., Ferrero-Garcia, M., Parodi A.J. (1992) *Biochemistry* **28**, 97-105

20. Ellgaard, L., Molinari, M., and Helenius, A. (1999) *Science* **286**, 1882-1888

21. Sousa, M., and Parodi, A. J. (1995) *Embo J* **14**, 4196-4203

22. Trombetta, E. S., and Helenius, A. (1998) *Curr Opin Struct Biol* **8**, 587-592

23. Meunier, L., Usherwood, Y. K., Chung, K. T., and Hendershot, L. M. (2002) *Mol Biol Cell* **13**, 4456-4469

24. Hammond, C., and Helenius, A. (1994) *J Cell Biol* **126**, 41-52

25. Kamhi-Nesher, S., Shenkman, M., Tolchinsky, S., Fromm, S. V., Ehrlich, R., and Lederkremer, G. Z. (2001) *Mol Biol Cell* **12**, 1711-1723

26. McCracken, A. A., and Brodsky, J. L. (1996) *J Cell Biol* **132**, 291-298

27. Hammond, C., and Helenius, A. (1995) *Curr Opin Cell Biol* **7**, 523-529

28. Chevet, E., Wong, H. N., Gerber, D., Cochet, C., Fazel, A., Cameron, P. H., Gushue, J. N., Thomas, D. Y., and Bergeron, J. J. (1999) *Embo J* **18**, 3655-3666

29. Wong, H. N., Ward, M. A., Bell, A. W., Chevet, E., Bains, S., Blackstock, W. P., Solari, R., Thomas, D. Y., and Bergeron, J. J. (1998) *J Biol Chem* **273**, 17227-17235

30. Ou, W. J., Thomas, D. Y., Bell, A. W., and Bergeron, J. J. (1992) *J Biol Chem* **267**, 23789-23796

31. Schrag, J. D., Bergeron, J. J., Li, Y., Borisova, S., Hahn, M., Thomas, D. Y., and Cygler, M. (2001) *Mol Cell* **8**, 633-644

32. Schrag, J. D., Procopio, D. O., Cygler, M., Thomas, D. Y., and Bergeron, J. J. (2003) *Trends Biochem Sci* **28**, 49-57

33. Ellgaard, L., and Helenius, A. (2001) *Curr Opin Cell Biol* **13**, 431-437

34. Leach, M. R., Cohen-Doyle, M. F., Thomas, D. Y., and Williams, D. B. (2002) *J Biol Chem* **277**, 29686-29697

35. Frickel, E. M., Riek, R., Jelesarov, I., Helenius, A., Wuthrich, K., and Ellgaard, L. (2002) *Proc Natl Acad Sci U S A* **99**, 1954-1959

36. Saito, Y., Ihara, Y., Leach, M. R., Cohen-Doyle, M. F., and Williams, D. B. (1999) *Embo J* **18**, 6718-6729

37. Ihara, Y., Cohen-Doyle, M. F., Saito, Y., and Williams, D. B. (1999) *Mol Cell* **4**, 331-341

38. Ou, W. J., Bergeron, J. J., Li, Y., Kang, C. Y., and Thomas, D. Y. (1995) *J Biol Chem* **270**, 18051-18059

39. Stronge, V. S., Saito, Y., Ihara, Y., and Williams, D. B. (2001) *J Biol Chem* **276**, 39779-39787

40. Molinari, M., Eriksson, K. K., Calanca, V., Galli, C., Cresswell, P., Michalak, M., and Helenius, A. (2004) *Mol Cell* **13**, 125-135

41. Ellgaard, L., and Frickel, E. M. (2003) *Cell Biochem Biophys* **39**, 223-247

42. Danilczyk, U. G., Cohen-Doyle, M. F., and Williams, D. B. (2000) *J Biol Chem* **275**, 13089-13097

43. Gao, B., Adhikari, R., Howarth, M., Nakamura, K., Gold, M. C., Hill, A. B., Knee, R., Michalak, M., and Elliott, T. (2002) *Immunity* **16**, 99-109

44. Rodan, A. R., Simons, J. F., Trombetta, E. S., and Helenius, A. (1996) *Embo J* **15**, 6921-6930

45. Zapun, A., Petrescu, S. M., Rudd, P. M., Dwek, R. A., Thomas, D. Y., and Bergeron, J. J. (1997) *Cell* **88**, 29-38

46. Labriola, C., Cazzulo, J. J., and Parodi, A. J. (1999) *Mol Biol Cell* **10**, 1381-1394

47. Danilczyk, U. G., and Williams, D. B. (2001) *J Biol Chem* **276**, 25532-25540

48. Liu, Y., Choudhury, P., Cabral, C. M., and Sifers, R. N. (1999) *J Biol Chem* **274**, 5861-5867

49. Branza-Nichita, N., Petrescu, A. J., Dwek, R. A., Wormald, M. R., Platt, F. M., and Petrescu, S. M. (1999) *Biochem Biophys Res Commun* **261**, 720-725

50. Zhang, J. X., Braakman, I., Matlack, K. E., and Helenius, A. (1997) *Mol Biol Cell* **8**, 1943-1954

51. Fernandez, F. S., Trombetta, S. E., Hellman, U., and Parodi, A. J. (1994) *J Biol Chem* **269**, 30701-30706

52. Trombetta, S. E., and Parodi, A. J. (1992) *J Biol Chem* **267**, 9236-9240

53. Parodi, A. J. (2000) *Annu Rev Biochem* **69**, 69-93

54. Ritter, C., and Helenius, A. (2000) *Nat Struct Biol* **7**, 278-280

55. Fujimoto, K., and Kornfeld, R. (1991) *J Biol Chem* **266**, 3571-3578

56. Esmon, B., Esmon, P. C., and Schekman, R. (1984) *J Biol Chem* **259**, 10322-10327

57. Fewell, S. W., Travers, K. J., Weissman, J. S., and Brodsky, J. L. (2001) *Annu Rev Genet* **35**, 149-191

58. Gething, M. J. (1999) *Semin Cell Dev Biol* **10**, 465-472

59. Simons, J. F., Ferro-Novick, S., Rose, M. D., and Helenius, A. (1995) *J Cell Biol* **130**, 41-49

60. Johnson, A. E., and Haigh, N. G. (2000) *Cell* **102**, 709-712

61. Flaherty, K. M., DeLuca-Flaherty, C., and McKay, D. B. (1990) *Nature* **346**, 623-628

62. Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996) *Science* **272**, 1606-1614

63. Knarr, G., Kies, U., Bell, S., Mayer, M., and Buchner, J. (2002) *J Mol Biol* **318**, 611-620

64. Cunnea, P. M., Miranda-Vizuete, A., Bertoli, G., Simmen, T., Damdimopoulos, A. E., Hermann, S., Leinonen, S., Huikko, M. P., Gustafsson, J. A., Sitia, R., and Spyrou, G. (2003) *J Biol Chem* **278**, 1059-1066

65. Kurisu, J., Honma, A., Miyajima, H., Kondo, S., Okumura, M., and Imaizumi, K. (2003) *Genes Cells* **8**, 189-202

66. Soldano, K. L., Jivan, A., Nicchitta, C. V., and Gewirth, D. T. (2003) *J Biol Chem* **278**, 48330-48338

67. Tu, B. P., and Weissman, J. S. (2002) *Mol Cell* **10**, 983-994

68. Tu, B. a. W., J. (2004) *Journal of Cell Biology* **164**, 341-346

69. Frickel, E. M., Frei, P., Bouvier, M., Stafford, W. F., Helenius, A., Glockshuber, R., and Ellgaard, L. (2004) *J Biol Chem*

70. Benedetti, C., Fabbri, M., Sitia, R., and Cabibbo, A. (2000) *Biochem Biophys Res Commun* **278**, 530-536

71. Pagani, M., Fabbri, M., Benedetti, C., Fassio, A., Pilati, S., Bulleid, N. J., Cabibbo, A., and Sitia, R. (2000) *J Biol Chem* **275**, 23685-23692

72. Tsai, B., Rodighiero, C., Lencer, W. I., and Rapoport, T. A. (2001) *Cell* **104**, 937-948

73. Lappi, A. K., Lensink, M. F., Alanen, H. I., Salo, K. E., Lobell, M., Juffer, A. H., and Ruddock, L. W. (2004) *J Mol Biol* **335**, 283-295

74. Lumb, R. A., and Bulleid, N. J. (2002) *Embo J* **21**, 6763-6770

75. Russell S, R. L., Salo K, Oliver J, Roebuck Q, Llewellyn, H. Llewelyn Roderick D, Koivunen P, Myllyharju J, and High S. (2004) *J Biol Chem* **in press Feb 2004**

76. Antoniou, A. N., Ford, S., Alphey, M., Osborne, A., Elliott, T., and Powis, S. J. (2002) *Embo J* **21**, 2655-2663

77. Paulsson, K. M., and Wang, P. (2004) *Faseb J* **18**, 31-38

78. Teckman, J. H., Gilmore, R., and Perlmutter, D. H. (2000) *Am J Physiol Gastrointest Liver Physiol* **278**, G39-48

79. McGee, T. P., Cheng, H. H., Kumagai, H., Omura, S., and Simoni, R. D. (1996) *J Biol Chem* **271**, 25630-25638

80. Cabral, C. M., Choudhury, P., Liu, Y., and Sifers, R. N. (2000) *J Biol Chem* **275**, 25015-25022

81. Yang, M., Omura, S., Bonifacino, J. S., and Weissman, A. M. (1998) *J Exp Med* **187**, 835-846

82. Chung, D. H., Ohashi, K., Watanabe, M., Miyasaka, N., and Hirosawa, S. (2000) *J Biol Chem* **275**, 4981-4987

83. Marcus, N. Y., and Perlmutter, D. H. (2000) *J Biol Chem* **275**, 1987-1992

84. Liu, Y., Choudhury, P., Cabral, C. M., and Sifers, R. N. (1997) *J Biol Chem* **272**, 7946-7951

85. Knop M., H. N., Wolf D.H. (1996) *Yeast* **12**, 1229-1238

86. Wu, Y., Swulius, M. T., Moremen, K. W., and Sifers, R. N. (2003) *Proc Natl Acad Sci U S A* **100**, 8229-8234

87. Pilon, M., Schekman, R., and Romisch, K. (1997) *Embo J* **16**, 4540-4548

88. Wiertz, E. J., Tortorella, D., Bogyo, M., Yu, J., Mothes, W., Jones, T. R., Rapoport, T. A., and Ploegh, H. L. (1996) *Nature* **384**, 432-438

89. Bebok, Z., Mazzochi, C., King, S. A., Hong, J. S., and Sorscher, E. J. (1998) *J Biol Chem* **273**, 29873-29878

90. Schmitz, A., Herrgen, H., Winkeler, A., and Herzog, V. (2000) *J Cell Biol* **148**, 1203-1212

91. Tsai, B., Ye, Y., and Rapoport, T. A. (2002) *Nat Rev Mol Cell Biol* **3**, 246-255

92. Plemper, R. K., Bordallo, J., Deak, P. M., Taxis, C., Hitt, R., and Wolf, D. H. (1999) *J Cell Sci* **112 (Pt 22)**, 4123-4134

93. Fagioli, C., Mezghrani, A., and Sitia, R. (2001) *J Biol Chem* **276**, 40962-40967

94. Bonifacino, J. S., and Weissman, A. M. (1998) *Annu Rev Cell Dev Biol* **14**, 19-57

95. Cabral, C. M., Liu, Y., Moremen, K. W., and Sifers, R. N. (2002) *Mol Biol Cell* **13**, 2639-2650

96. Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L. O., Herscovics, A., and Nagata, K. (2001) *EMBO Rep* **2**, 415-422

97. Nagase, T., Seki, N., Ishikawa, K., Ohira, M., Kawarabayasi, Y., Ohara, O., Tanaka, A., Kotani, H., Miyajima, N., and Nomura, N. (1996) *DNA Res.* **3**, 321-329

98. Yoshida, H., Matsui, T., Hosokawa, N., Kaufman, R. J., Nagata, K., and Mori, K. (2003) *Dev Cell* **4**, 265-271

99. Lee, A. H., Iwakoshi, N. N., and Glimcher, L. H. (2003) *Mol Cell Biol* **23**, 7448-7459

100. Molinari, M., Calanca, V., Galli, C., Lucca, P., and Paganetti, P. (2003) *Science* **299**, 1397-1400

101. Oda, Y., Hosokawa, N., Wada, I., and Nagata, K. (2003) *Science* **299**, 1394-1397

102. Vashist, S., Frank, C. G., Jakob, C. A., and Ng, D. T. (2002) *Mol Biol Cell* **13**, 3955-3966

103. Suzuki, T., and Lennarz, W. J. (2002) *Glycobiology* **12**, 803-811

104. Johnson, A. E., and van Waes, M. A. (1999) *Annu Rev Cell Dev Biol* **15**, 799-842

105. Dudek, J., Volkmer, J., Bies, C., Guth, S., Muller, A., Lerner, M., Feick, P., Schafer, K. H., Morgenstern, E., Hennessy, F., Blatch, G. L., Janoscheck, K., Heim, N., Scholtes, P., Frien, M., Nastainczyk, W., and Zimmermann, R. (2002) *Embo J* **21**, 2958-2967

106. Wilkinson, B. M., Tyson, J. R., Reid, P. J., and Stirling, C. J. (2000) *J Biol Chem* **275**, 521-529

107. Brodsky, J. L., and Schekman, R. (1993) *J Cell Biol* **123**, 1355-1363

108. Benach, J., and Hunt, J. F. (2004) *Nature* **427**, 24-26

109. Van den Berg, B., Clemons, W. M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004) *Nature* **427**, 36-44

110. Wesche, J., Rapak, A., and Olsnes, S. (1999) *J Biol Chem* **274**, 34443-34449

111. Brunger, A. T., and DeLaBarre, B. (2003) *FEBS Lett* **555**, 126-133

112. Lord, J. M., Ceriotti, A., and Roberts, L. M. (2002) *Curr Biol* **12**, R182-184

113. Elkabetz, Y., Shapira, I., Rabinovich, E., and Bar-Nun, S. (2004) *J Biol Chem* **279**, 3980-3989

114. Flierman, D., Ye, Y., Dai, M., Chau, V., and Rapoport, T. A. (2003) *J Biol Chem* **278**, 34774-34782

115. Ye, Y., Meyer, H. H., and Rapoport, T. A. (2003) *J Cell Biol* **162**, 71-84

116. Bays, N. W., Gardner, R. G., Seelig, L. P., Joazeiro, C. A., and Hampton, R. Y. (2001) *Nat Cell Biol* **3**, 24-29

117. Kaneko, M., Ishiguro, M., Niinuma, Y., Uesugi, M., and Nomura, Y. (2002) *FEBS Lett* **532**, 147-152

118. Knop M, F. A., Braun T, Hellmuth K, Wolf DH. (1996) *EMBO J* **15**, 753-763

119. Cyr, D. M., Hohfeld, J., and Patterson, C. (2002) *Trends Biochem Sci* **27**, 368-375

120. Gardner, R. G., Shearer, A. G., and Hampton, R. Y. (2001) *Mol Cell Biol* **21**, 4276-4291

121. Caldwell, S. R., Hill, K. J., and Cooper, A. A. (2001) *J Biol Chem* **276**, 23296-23303

122. Haynes, C. M., Caldwell, S., and Cooper, A. A. (2002) *J Cell Biol* **158**, 91-101

123. Pickart, C. M. (2001) *Annu Rev Biochem* **70**, 503-533

124. Yoshida, Y., Chiba, T., Tokunaga, F., Kawasaki, H., Iwai, K., Suzuki, T., Ito, Y., Matsuoka, K., Yoshida, M., Tanaka, K., and Tai, T. (2002) *Nature* **418**, 438-442

125. Hampton, R. Y. (2002) *Curr Opin Cell Biol* **14**, 476-482

126. Fang, S., Ferrone, M., Yang, C., Jensen, J. P., Tiwari, S., and Weissman, A. M. (2001) *Proc Natl Acad Sci U S A* **98**, 14422-14427

127. Hirsch, C., Misaghi, S., Blom, D., Pacold, M. E., and Ploegh, H. L. (2004) *EMBO Rep* **5**, 201-206

128. Hirsch, C., Blom, D., and Ploegh, H. L. (2003) *Embo J* **22**, 1036-1046

129. Bochtler, M., Ditzel, L., Groll, M., Hartmann, C., and Huber, R. (1999) *Annu Rev Biophys Biomol Struct* **28**, 295-317

130. Voges, D., Zwickl, P., and Baumeister, W. (1999) *Annu Rev Biochem* **68**, 1015-1068

131. Shen, J., Chen, X., Hendershot, L., and Prywes, R. (2002) *Dev Cell* **3**, 99-111

132. Travers, K. J., Patil, C. K., Wodicka, L., Lockhart, D. J., Weissman, J. S., and Walter, P. (2000) *Cell* **101**, 249-258

133. Harding, H. P., Calfon, M., Urano, F., Novoa, I., and Ron, D. (2002) *Annu Rev Cell Dev Biol* **18**, 575-599

134. Ma, Y., and Hendershot, L. M. (2001) *Cell* **107**, 827-830

135. Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P., and Ron, D. (2000) *Nat Cell Biol* **2**, 326-332

136. Rutkowski, D. T., and Kaufman, R. J. (2004) *Trends Cell Biol* **14**, 20-28

137. Kopito, R. R. (1999) *Physiol Rev* **79**, S167-173

138. Plemper, R. K., and Wolf, D. H. (1999) *Trends Biochem Sci* **24**, 266-270

139. Kudo, T., Katayama, T., Imaizumi, K., Yasuda, Y., Yatera, M., Okochi, M., Tohyama, M., and Takeda, M. (2002) *Ann N Y Acad Sci* **977**, 349-355

140. Ryu, E. J., Harding, H. P., Angelastro, J. M., Vitolo, O. V., Ron, D., and Greene, L. A. (2002) *J Neurosci* **22**, 10690-10698

141. Story, C. M., Furman, M. H., and Ploegh, H. L. (1999) *Proc Natl Acad Sci U S A* **96**, 8516-8521

142. Schubert, U., Anton, L. C., Bacik, I., Cox, J. H., Bour, S., Bennink, J. R., Orlowski, M., Strebel, K., and Yewdell, J. W. (1998) *J Virol* **72**, 2280-2288

143. Tyers, M., and Jorgensen, P. (2000) *Curr Opin Genet Dev* **10**, 54-64

144. Grondin, B., and Herscovics, A. (1992) *Glycobiology* **2**, 369-372

145. Gonzalez, D. S., Karaveg, K., Vandersall-Nairn, A. S., Lal, A., and Moremen, K. W. (1999) *J Biol Chem* **274**, 21375-21386

146. Jelinek-Kelly, S., Akiyama, T., Saunier, B., Tkacz, J. S., and Herscovics, A. (1985) *J Biol Chem* **260**, 2253-2257

147. Jelinek-Kelly, S., and Herscovics, A. (1988) *J Biol Chem* **263**, 14757-14763

148. Moremen, K. W. (2000) in *Carbohydrates in chemistry and biology* (Beat Ernst, G. W. H., Pierre Sanaÿ, ed), 4 vols., Weinheim, New York

149. Lal, A., Pang, P., Kalelkar, S., Romero, P. A., Herscovics, A., and Moremen, K. W. (1998) *Glycobiology* **8**, 981-995

150. Bieberich, E., Treml, K., Volker, C., Rolfs, A., Kalz-Fuller, B., and Bause, E. (1997) *Eur J Biochem* **246**, 681-689

151. Lal, A., Schutzbach, J. S., Forsee, W. T., Neame, P. J., and Moremen, K. W. (1994) *J Biol Chem* **269**, 9872-9881

152. Herscovics, A., Schneikert, J., Athanassiadis, A., and Moremen, K. W. (1994) *J Biol Chem* **269**, 9864-9871

153. Tremblay, L. O., and Herscovics, A. (2000) *J Biol Chem* **275**, 31655-31660

154. Lobsanov, Y. D., Vallee, F., Imberty, A., Yoshida, T., Yip, P., Herscovics, A., and Howell, P. L. (2002) *J Biol Chem* **277**, 5620-5630

155. Ichishima, E., Taya, N., Ikeguchi, M., Chiba, Y., Nakamura, M., Kawabata, C., Inoue, T., Takahashi, K., Minetoki, T., Ozeki, K., Kumagai, C., Gomi, K., Yoshida, T., and Nakajima, T. (1999) *Biochem J* **339 (Pt 3)**, 589-597

156. Vallee, F., Lipari, F., Yip, P., Sleno, B., Herscovics, A., and Howell, P. L. (2000) *Embo J* **19**, 581-588

157. Vallee, F., Karaveg, K., Herscovics, A., Moremen, K. W., and Howell, P. L. (2000) *J Biol Chem* **275**, 41287-41298

158. Van Petegem, F., Contreras, H., Contreras, R., and Van Beeumen, J. (2001) *J Mol Biol* **312**, 157-165

159. Lipari, F., Gour-Salin, B. J., and Herscovics, A. (1995) *Biochem Biophys Res Commun* **209**, 322-326

160. Mulakala, C., and Reilly, P. J. (2002) *Proteins* **49**, 125-134

161. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) *J Bacteriol* **153**, 163-168

162. Wach, A., Brachat, A., Pohlmann, R., and Philippsen, P. (1994) *Yeast* **10**, 1793-1808

163. Gietz, R. D., and Woods, R. A. (1994) in *Molecular Genetics of Yeast, A practical approach* (Johnston, J. R., ed, ed), pp. 121-134, IRL Press, Oxford, UK

164. Moremen, K. W., and Robbins, P. W. (1991) *J Cell Biol* **115**, 1521-1534

165. Sikorski, R. S., and Hieter, P. (1989) *Genetics* **122**, 19-27

166. Sambrook, J., Fritsch, E.F. and Maniatis , T. (1989) *Molecular Cloning : a Laboratory Manual*, 2nd edn Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

167. RN Sifers, S. B.-M., VJ Kidd, H Muensch, and SL Woo. (1988) *J. Biol. Chem.* **263**, 7330-7335

168. Le, A., Ferrell, G. A., Dishon, D. S., Le, Q. Q., and Sifers, R. N. (1992) *J Biol Chem* **267**, 1072-1080

169. Nakai, K., and Kanehisa, M. (1992) *Genomics* **14**, 897-911

170. Persson, B., and Argos, P. (1994) *J Mol Biol* **237**, 182-192

171. Mumberg, D., Muller, R., and Funk, M. (1995) *Gene* **156**, 119-122

172. Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F., and Cullin, C. (1993) *Nucleic Acids Res* **21**, 3329-3330

173. Spirig, U., Glavas, M., Bodmer, D., Reiss, G., Burda, P., Lippuner, V., te Heesen, S., and Aebi, M. (1997) *Mol Gen Genet* **256**, 628-637

174. Finger, A., Knop, M., and Wolf, D. H. (1993) *Eur. J. Biochem.* **218**, 565-573

175. Hiller, M. M., Finger, A., Schweiger, M., and Wolf, D. H. (1996) *Science* **273**, 1725-1728

176. Hitt, R., and Wolf, D. H. (2004) *FEMS Yeast Res* **4**, 721-729

177. Kyte, J., and Doolittle, R. F. (1982) *J Mol Biol* **157**, 105-132

178. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998) *Bioinformatics* **14**, 378-379

179. Briand, J. P., Barin, C., Van Regenmortel, M. H., and Muller, S. (1992) *J Immunol Methods* **156**, 255-265

180. Vandersall-Nairn, A. S., Merkle, R. K., O'Brien, K., Oeltmann, T. N., and Moremen, K. W. (1998) *Glycobiology* **8**, 1183-1194

181. Falquet, L., Pagni, M., Bucher, P., Hulo, N., Sigrist, C. J., Hofmann, K., and Bairoch, A. (2002) *Nucleic Acids Res* **30**, 235-238

182. Mahon, P., and Bateman, A. (2000) *Protein Sci* **9**, 1930-1934

183. Munro, S., and Pelham, H. R. (1987) *Cell* **48**, 899-907

184. De Groot, P. W., Hellingwerf, K. J., and Klis, F. M. (2003) *Yeast* **20**, 781-796

185. Lipari, F., and Herscovics, A. (1996) *J Biol Chem* **271**, 27615-27622

186. Lomas, D. A., Evans, D. L., Finch, J. T., and Carrell, R. W. (1992) *Nature* **357**, 605-607

187. Frenkel, Z., Gregory, W., Kornfeld, S., and Lederkremer, G. Z. (2003) *J Biol Chem* **278**, 34119-34124

188. Kitzmuller, C., Caprini, A., Moore, S. E., Frenoy, J. P., Schwaiger, E., Kellermann, O., Ivessa, N. E., and Ermonval, M. (2003) *Biochem J* **Pt**

189. Ermonval, M., Kitzmuller, C., Mir, A. M., Cacan, R., and Ivessa, N. E. (2001) *Glycobiology* **11**, 565-576

190. Foulquier, F., Duvet, S., Klein, A., Mir, A. M., Chirat, F., and Cacan, R. (2004) *Eur J Biochem* **271**, 398-404