Doctoral Thesis

Htm1p function in ER-associated protein degradation

Author[s]: Clerc, Simone Christine

Publication Date: 2009

Permanent Link: https://doi.org/10.3929/ethz-a-005911904

Rights / License: In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Htm1p function in ER-associated protein degradation

A dissertation submitted to
ETH ZÜRICH

for the degree of
DOCTOR OF SCIENCES

presented by
SIMONE CHRISTINE CLERC
Eidg. dipl. Apothekerin ETHZ

born on September 10, 1977
citizen of Hauteville FR, Switzerland

accepted on the recommendation of
Prof. Dr. Markus Aebi, examiner
Prof. Dr. Ari Helenius, co-examiner

2009
Acknowledgements

I am deeply indebted to Prof. Dr. Markus Aebi who accepted to be my Doktorvater and gave me the opportunity to do my thesis in his laboratory. I cordially thank him for his supervision and support. I very much appreciate his scientific acumen as well as his caring personality.

I express my gratitude to Dr. Claude Jakob for supervision and support during the first half of the thesis. I thank him for introducing me to basic research, work with yeast and to the fascinating topic of ER-associated protein degradation.

I thank Prof. Dr. Ari Helenius for kindly agreeing to co-referee my thesis.

I cordially thank Dr. Paola Deprez for her scientific and personal advice, encouragement, support and all the funny jokes she made.

Special thanks go to my student Daniela Oggier. With her diploma thesis she contributed essentially to the advance of this project.

I thank all former and present members of the Institute of Microbiology, the Aebi group and the yeast subgroup for the very pleasant ambiance. In particular, I thank Alex, Christine, Farnoush, Martin, Paola, Reto and S. cerevisiae. They were all important for my scientific progress and I very much enjoyed the time we spent together, be it at work or in spare time.

I thank the management and the staff of the Institute of Microbiology, particularly Alain, Daniel, Fabian, Jaques, Palmira, Paul, Silvia and Stefan.

I express my heartfelt gratitude to my boyfriend Daniel and to my family for their love, patience and support, enabling me to complete the thesis.
# Table of Contents

Summary .................................................................................................................. ii

Zusammenfassung ................................................................................................... iv

Chapter 1 ............................................................................................................... 1
    Protein folding and ER-associated degradation

Chapter 2 .............................................................................................................. 57
    Screening for Htm1p interacting proteins

Chapter 3 ............................................................................................................. 85
    Htm1 protein generates the N-glycan signal for glycoprotein degradation
    in the endoplasmic reticulum

Chapter 4 ............................................................................................................. 127
    Analysis of the substrate specificity of Htm1p

Curriculum vitae .................................................................................................. 143
Summary

To maintain protein homeostasis in secretory compartments, eukaryotic cells harbor a quality control system that monitors protein folding and protein complex assembly in the endoplasmic reticulum (ER). Proteins that do not fold properly or integrate into cognate complexes are degraded by ER-associated degradation (ERAD), involving retrotranslocation to the cytoplasm and proteasomal peptide hydrolysis. N-linked glycans and processing products thereof are key players not only in protein folding but also in the degradation of misfolded glycoproteins: The covalent structure of the oligosaccharide represents the folding status of the protein. It was proposed that the Man₈GlcNAc₂ glycan structure serves as the N-glycan signal for protein misfolding, and that this signal is decoded by the apparently inactive member of the GH47 family of glycosylhydrolases, the Htm1 (Homologous to mannosidase I) protein in S. cerevisiae. An overview on protein folding, quality control and ER-associated degradation of N-glycosylated proteins is given in chapter one.

Goal of this thesis was to elucidate the function of Htm1p in ERAD.

As described in chapter two, in a first approach the protein environment of Htm1p was assessed in vivo in a yeast two-hybrid screen. Interacting proteins of Htm1p were identified that may reveal novel and relevant aspects in protein folding, quality control and ERAD in the future.

In chapter 3 the question whether Htm1p may function as an active enzyme in ERAD, was addressed. It was found that Htm1p removes the terminal α1,2-linked mannose of the C branch from N-glycans on misfolded glycoproteins that have been processed to Man₈GlcNAc₂. By this, it generates a free α1,6-linked mannose which is the N-glycan ER-degradation signal, that is recognized by the ER degradation receptor Yos9p. These findings suggest a novel working model for ERAD in S. cerevisiae, in which Htm1p is the key enzyme that makes the distinction between unfolded and terminally misfolded proteins.

Analysis and comparison of the N-glycan substrate specificity of Htm1p with other members of the GH47 family revealed a unique specificity. While Htm1p removes
one mannose (similar ER mannosidase), it processes the C branch of the glycan (similar to Golgi mannosidases). Moreover, ER mannosidase I and Golgi mannosidase members act complementary and independently of each other whereas Htm1p is functionally dependent on Mns1p. It is possible that the residue S277 is a key residue in the regulation of this specificity.
Zusammenfassung


Ziel dieser Doktorarbeit war herauszufinden, welche Funktion das Htm1 Protein in ERAD hat.

Wie in Kapitel zwei beschrieben ist, wurde in einem ersten Ansatz die Proteinumgebung von Htm1p ermittelt. Dies erfolgte in vivo mittels eines, ‘yeast two- hybrid screens’. Es wurden verschiedene Proteine gefunden die mit Htm1p interagieren. Daraus könnten sich in der Zukunft neue und relevante Aspekte für die Proteinfaltung, Qualitätskontrolle und ERAD ergeben.

In Kapitel 3 wurde die Frage gestellt, ob Htm1p als aktives Enzym in ERAD funktionieren könnte. Es wurde herausgefunden, dass Htm1p die endständige α1,2- gebundene Mannose vom C Arm von missgefaulten Glykoproteinen schneidet, die bereits zu Man8GlcNAc2 getrimmt worden sind. Dadurch erzeugt Htm1p eine freie α1,6-gebundene Mannose, die das N-Glykan ERAD Signal ist. Dieses wird durch den ERAD Rezeptor Yos9p erkannt. Diese Erkenntnisse ergeben ein neues ERAD
Zusammenfassung

Modell für *S. cerevisiae*. In diesem Modell ist Htm1p das Schlüsselenzym, das die Unterscheidung zwischen ungefalteten und missgefalteten Proteinen macht.

Analyse und Vergleich der N-Glykan Substratspezifität von Htm1p mit anderen Mitgliedern der GH47 Famile ergab, dass Htm1p einzigartig ist. Während Htm1p eine einzelne Mannose vom Glykan abschneidet (ähnlich wie die ER Mannosidase I), schneidet es den C Arm des Glykans (ähnlich der Golgi Mannosidasen). Zudem ist Htm1p funktionell von der ER Mannosidase I abhängig, während die ER Mannosidase I und Golgi Mannosidasen komplementär und unabhängig voneinander arbeiten. Es könnte sein, dass die Aminosäure S277 eine Schlüsselrolle in der Regulation dieser Spezifität hat.
Chapter 1

Protein folding and ER-associated degradation
Chapter 1

1 Protein folding in the ER .........................................................4

1.1 Introduction ................................. ........................................4

1.2 Protein translocation ..........................................................5
  1.2.1 Co-translational translocation ........................................ 5
  1.2.2 Post-translational translocation ..................................... 5
  1.2.3 Translocon ................................................................. 6
  1.2.4 Molecular mechanism of translocation ............................. 6

1.3 ER resident folding factors ...................................................7
  1.3.1 Classical chaperones ................................................... 7
    1.3.1.1 BiP and its cofactors ............................................. 7
    1.3.1.2 GRP94 .............................................................. 8
  1.3.2 Peptide bond isomerases .............................................. 9
  1.3.3 Protein disulfide isomerases ...................................... 10
    1.3.3.1 PDI protein ....................................................... 10
    1.3.3.2 ERp57 ............................................................ 11
    1.3.3.3 ERdj5 ............................................................ 11
  1.3.4 N-linked glycans ....................................................... 11
    1.3.4.1 Transfer of the N-glycan ..................................... 13
    1.3.4.2 Mammalian co- and post-translational glycosylation .... 13

1.4 N-glycan assisted protein folding .........................................14
  1.4.1 N-glycan processing .................................................. 15
    1.4.1.1 Glucosidase I ..................................................... 15
    1.4.1.2 Glucosidase II ................................................... 15
    1.4.1.3 UDP-glucose:glycoprotein glucosyltransferase ............ 16
    1.4.1.4 ER mannosidase I .............................................. 17
  1.4.2 Lectin chaperones calnexin and calreticulin ...................... 17
  1.4.3 Calnexin/calreticulin cycle ........................................ 18

1.5 Quality control .................................................................19

1.6 Export of proteins from the ER ............................................19
  1.6.1 N-glycans in ER export ............................................. 20

1.7 Quality control beyond ER ..................................................21
  1.7.1 N-glycans in post-ER quality control .............................. 21

2 ER-associated degradation ....................................................23

2.1 Introduction .................................................................23

2.2 Distinction between unfolded and terminally misfolded ............24
  2.2.1 Importance of N-glycans ............................................ 24
    2.2.1.1 Importance of mannose trimming – Mannose timer .... 24
2.2.1.1 Processing of Man₈GlcNAc₂ in S. cerevisiae .........25
2.2.1.1.1 Processing of Man₈GlcNAc₂ in higher eukaryotes .....25
2.2.1.1.3 Mannose processing in S. pombe? .........................26
2.2.1.2 Enzymes implicated in the processing of Man₈GlcNAc₂ ...28
2.2.1.3 Htm1p and EDEM proteins ....................................28
2.2.1.3.1 Comparison and description .................................28
2.2.1.3.2 Importance for ERAD ..................................29
2.2.1.3.3 Mannosidase function ..................................31

2.3 ERAD substrate recognition ...........................................32
2.3.1 Yos9p, OS-9 and XTP3-B proteins .................................32
2.3.1.1 Yos9p ..............................................................33
2.3.1.1.1 Description ..................................................33
2.3.1.1.2 Importance for ERAD .................................33
2.3.1.1.3 Proof-reading function ................................33
2.3.1.2 OS-9 and XTP3-B ..............................................34

2.4 ERAD substrate retrotranslocation ............................35
2.4.1 Classical chaperones ............................................35
2.4.2 Protein disulfide isomerases .....................................36
2.4.3 Retrotranslocon ....................................................36
2.4.3.1 Translocon .....................................................37
2.4.3.2 Derlin-1 .........................................................37
2.4.3.3 E3 ubiquitin ligase ...........................................37
2.4.3.4 Alternative concepts ......................................38

2.5 ERAD substrate ubiquitylation .......................................38
2.5.1 Ubiquitylation reaction ...........................................38
2.5.2 Ubiquitylation enzymes ......................................39

2.6 Cytoplasmic extraction ..............................................40
2.6.1 Cdc48 complex ..................................................40

2.7 Proteasomal degradation .............................................40
2.7.1 Proteasomal targeting ...........................................40
2.7.2 Deglycosylation and deubiquitylation .....................41

List of references ..............................................................42
1 Protein folding in the ER

1.1 Introduction

The information about the native conformation of a polypeptide chain is encoded in its amino acid sequence due to the specific physico-chemical properties of the residues (Anfinsen et al., 1961). However, the driving force to build the three-dimensional structure is to reach an energetically favorable state. According to the Levinthal paradox (Levinthal, 1968; Levinthal, 1969; Zwanzig et al., 1992) a sequential sampling of all possible conformations does not occur but protein folding is rather a directed process. The currently hold concept of a funnel-like energy landscape states that although different routes may be taken, in general they drive downhill in the energy landscape (Leopold et al., 1992). By this the number of taken routes can be restricted to some extent. As the native conformation is an intrinsic property of a polypeptide chain, the protein may be able to reach this state without assistance. Though, the exposure of hydrophobic residues (which are not yet buried) may lead to formation of undesirable interactions within one or between different polypeptide chains in the cellular context. This can lead to misfolding and aggregation of the polypeptide chains. Aggregation constitutes a major off-pathway that can potentially be taken (Bukau et al., 2006). Protein folding is assisted by molecular chaperones and folding enzymes. These folding factors on the one hand smooth the energy landscape as they catalyze rate-limiting reactions in the folding process and on the other hand they guide the folding proteins on their route downwards to the energy minimum as they prevent non-productive interactions with other proteins or prevent proteins from getting trapped in off-pathway intermediates (Christis et al., 2008). By this, folding factors increase the probability for polypeptide chains to reach the native conformation in a biologically acceptable time span. In vivo protein synthesis and with this protein folding takes place in the cytosol, in mitochondria, in chloroplasts, and in the endoplasmic reticulum (ER) where correspondingly different folding factors are found (Hebert and Molinari, 2007). This implies that the set of folding factors that assists a specific protein depends on the final destination of the protein or on the route it takes to get there. In the following, the focus will be on the folding of proteins that are synthesized in the ER.
1.2 Protein translocation

To access the ER, nascent polypeptide chains need to dislocate from the cytosol to the ER compartment. This can occur either during or after translation.

1.2.1 Co-translational translocation

In higher eukaryotes most proteins are believed to be translocated into the ER during their synthesis (co-translational). For this, targeting to the ER membrane is mostly mediated by a signal peptide which is a hydrophobic sequence of about 20 residues at the N-terminus of the nascent protein (Blobel and Dobberstein, 1975). This sequence is bound by the signal recognition particle (SRP) which is an RNA/protein complex (Walter and Blobel, 1982). The translation is temporarily stalled and the ribosome is targeted to a translocon pore in the ER membrane (Walter and Blobel, 1981) where translation of the mRNA is resumed and the SRP is released. However, recent data suggests that mRNA itself may target the translating ribosome to the ER (Pyhtila et al., 2008). In the co-translational translocation mode, the folding of (luminal segments of) the protein will directly be assisted by ER localized folding factors.

1.2.2 Post-translational translocation

The translocation of proteins after their synthesis is termed post-translational translocation. It is unknown how often this mode is used in higher eukaryotes, but it seems to be quite active in fast growing cells such as yeast (Kalies and Hartmann, 1998; Rapoport, 2007). Prior to the post-translational translocation, nascent polypeptide chains that emerge from the ribosome channel into the cytosol, contact the cytosolic folding machinery (Wegrzyn and Deuerling, 2005; Young et al., 2004). These folding factors prevent their client proteins from misfolding and aggregation and thus ensure that the polypeptide chain remains in a folding-competent state until it reaches the ER. Translocation starts with binding of the substrate to the translocation pore in the ER membrane and the release of the cytosolic chaperones (Plath and Rapoport, 2000).
1.2.3 Translocon

In both, co- and post-translational translocation, a heterotrimeric Sec61 protein complex functions as translocation pore. It is a proteinaceous channel that comprises the Secα, -β, and -γ subunit in mammals of which the α subunit forms the actual channel (Van den Berg et al., 2004). In S. cerevisiae two Sec61 complexes exist: One comprises Sec61p, Sbh1p, and Sss1p, of which the α and the γ subunit proteins Sec61p and Sss1p are essential for viability. The second Sec61 complex (termed Ssh1 complex) is composed of Ssh1p, Sbh2p, and Sss1p. However, this complex is not essential and seems to function exclusively in co-translational translocation (Finke et al., 1996; Osborne et al., 2005).

1.2.4 Molecular mechanism of translocation

The translocation pore is a passive channel that in principle allows a polypeptide chain to slide back and forth. In dependence of the translocation mode, different additional proteins are engaged in the translocation process to ensure the forward transport of the peptide into the ER (Osborne et al., 2005).

In co-translational translocation, the ribosome is bound to the translocation complex and the driving force for the uni-directional translocation is given by the GTP-hydrolysis-driven ribosomal translation process (Osborne et al., 2005). In contrast, in post-translational translocation, the Sec61 translocation complex is associated with another membrane protein complex, the Sec62/63 complex, and with the luminal ER chaperone BiP. The Sec62/63 complex is a tetramer in yeast comprising the essential subunits Sec62p, and Sec63p, as well as the non-essential Sec71p, and Sec72p. In mammalian cells, the latter two are absent (Meyer et al., 2000; Tyedmers et al., 2000). The stability of the Sec62/63 complex requires the phosphorylated C-terminal domain of Sec63p (Wang and Johnsson, 2005). Substrate recognition is mediated by Sec62p. Upon association with the Sec61 complex a seven-component Sec complex results (Deshaies et al., 1991; Panzner et al., 1995). In this translocation mode, the driving force is provided by a ratcheting mechanism (Matlack et al., 1999): The J domain of Sec63p stimulates BiP to hydrolyse ATP (Misselwitz et al., 1998). In consequence, BiP binds to parts of the polypeptide chain that emerged into the ER. This association blocks the chain from moving backward while
subsequent polypeptide segments that emerge into the ER lumen (due to Brownian motion) are again bound by further BiP molecules. Net forward movement results and iteration of this process finally completes the protein translocation.

1.3 ER resident folding factors

The ER is a special folding compartment. On one hand it harbors folding factors that have homology to folding factors present in other folding compartments. On the other hand it introduces special features into the proteins and their folding process: It provides conditions and folding factors for the formation of disulfide bonds and it builds and transfers N-linked glycans, that are effective in protein folding themselves, and function as targets for ER-specific folding factors. In the ER, further post-translational modifications such as O-linked glycosylation or GPI anchorage occur. Discussion thereof would go beyond the scope of this introduction.

1.3.1 Classical chaperones

Classical molecular chaperones prevent aggregation of unfolded chains, facilitate protein maturation and retain folding proteins in appropriate micro- and macro-environments enriched with folding enzymes (Hebert and Molinari, 2007). While the cytosolic members of the evolutionarily conserved heat shock protein (Hsp) family are transcriptionally induced upon temperature stress; ER resident members are up-regulated under glucose deprivation, overload of the ER with misfolded proteins, as well as with unbalanced calcium and redox conditions (Lee, 2001; Shiu et al., 1977). Based on their respective molecular weight, Hsp’s are grouped in sub-families. Members of the Hsp70, Hsp40, Hsp90, Hsp100, and Hsp110 family have been reported (Hebert and Molinari, 2007).

1.3.1.1 BiP and its cofactors

The ER resident Hsp70 chaperone is called BiP/glucose-regulated protein (GRP)78 in higher eukaryotes (Munro and Pelham, 1986), or Kar2p in yeast (Brodsky and Schekman, 1993). It was shown to preferentially bind peptides with alternating aromatic and hydrophobic residues (Blond-Elgundi et al., 1993; Flynn et al., 1991).
BiP has an N-terminal ATPase domain that regulates its C-terminal substrate binding domain. When ATP is bound, BiP has a low peptide affinity open conformation, while ADP converts BiP into a high-affinity closed state conformation. Thus, alternation of ATP-hydrolysis and ADP to ATP exchange results in cycles of substrate binding and release, which is also termed chaperoning cycle (Mayer et al., 2003). BiP has several co-chaperones that belong to the Hsp40 family and are called ERdj1/2/3/4/5 (Bies et al., 1999; Brightman et al., 1995; Cunnea et al., 2003; Shen et al., 2002; Skowronek et al., 1999). ERdj proteins broaden the peptide-sequence specificity of BiP (Misselwitz et al., 1998) and they enhance the ATPase activity of BiP through their J domain (Awad et al., 2008; Cheetham and Caplan, 1998). ERdj2/Sec63p is acting as co-chaperone with BiP in the ratcheting mechanism during post-translational translocation. Characteristic for ERdj5/JPDI is the presence of an oxidoreductase activity in addition to the J domain (Cunnea et al., 2003; Hosoda et al., 2003).

Two factors have been identified as nucleotide exchange factors for BiP: BAP/Sil1 (Chung et al., 2002), and GRP170 (Steel et al., 2004). These factors assist in the ADP to ATP swapping, thus in the generation of the low-affinity conformation required for the substrate release. GRP170/Lsh1p, a member of the Hsp110 family, is itself a chaperone (Lin et al., 1993). In turn, its ATPase activity is stimulated by BiP (Steel et al., 2004). Contrary to BiP, Lsh1p is not required for the folding of nascent polypeptide chains but for refolding of proteins after heat shock induced misfolding (Saris and Makarow, 1998).

1.3.1.2 GRP94

The ER localized Hsp90 chaperone is called GRP94. Whereas BiP is evolutionarily conserved from yeast to human, GRP94 is only found in vertebrates (Lee, 2001). There, it constitutes one of the most abundant ER-resident chaperones (Koch et al., 1986). GRP94 has high calcium binding affinity and it functions as an important calcium buffer (Van et al., 1989). GRP94, like other Hsp90 proteins, is organized in three domains: The N-terminal ATP-binding domain, a central substrate-binding domain, and a C-terminal dimerization domain. Regulating mechanisms for GRP94 are largely unknown (Dollins et al., 2005; Immormino et al., 2004; Soldano et al., 2003). Recently, comparable ATPase activities for GRP94 and yeast Hsp90 were
assessed. Yet, GRP94 does not show the same conformational changes during the chaperoning cycle: the open and the closed conformations exist each in both, the ATP- as well as the ADP-bound state (Dollins et al., 2007; Frey et al., 2007). GRP94 appears to bind more advanced folding intermediates than BiP does, as GRP94 does not interact with nascent polypeptide chains but associates with substrates only after their release from BiP (Melnick et al., 1994). In addition, GRP94 seems to have narrower substrate specificity compared with BiP since it acts on a more specific set of proteins (Argon and Simen, 1999). While GRP94 interacts with the major histocompatibility complex (MHC) class II, it does not act on the structurally related MHC class I chains (Schaiff et al., 1992). Moreover, GRP94 was shown to interact with a variety of receptors such as Toll-like receptors, insulin-like growth factor receptors, and integrins (Randow and Seed, 2001).

1.3.2 Peptide bond isomerases

During mRNA translation the amino acids of the nascent polypeptide chain are coupled in trans-configuration (Lim and Spirin, 1986). Due to the lower energy content most peptide bonds are also trans in the finally folded protein (Scherer et al., 1998), though, some bonds are finally cis. The transition from trans to cis may be rate-limiting for folding (Odefey et al., 1995). The situation is different for peptide-bonds between an amino acid and a proline (X-P). Here, the cis- and trans-isomers are similarly stable (Stein, 1993). Depending on the side chain, 6 to 38% of X-P peptides in the folded protein are cis-isomers (Reimer et al., 1998). Yet, refolding experiments showed that isomerization of peptidyl-prolyl bonds is a rate-limiting step in the polypeptide folding process (Kiefhaber et al., 1990). These reactions are catalyzed by peptidyl-prolyl cis-trans isomerases (PPIase) (Fischer et al., 1984; Fischer et al., 1989). Depending on their binding specificity for immunosuppressive drugs the PPIases are classified in three families, namely parvulins, cyclophilins (Cyps), and FK506-binding proteins (FKBPs) (Gothel and Marahiel, 1999). Members of the latter two families have been identified in the ER. CypB, a cyclophilin member, was found in functional complexes with the Hsp47 chaperone required for the export of procollagen from the ER (Smith et al., 1995). It was also found in multiprotein complexes with other ER chaperones (Meunier et al., 2002; Zhang and Herscovitz, 2003). FKBP10/FKBP65 associates with BiP (Davis et al., 1998) and
FKBP7/FKBP23 binds and modulates BiP’s ATPase activity by means of its PPIase activity (Wang et al., 2007; Zhang et al., 2004). Expression of the PPIase FKBP2/FKBP13 is induced in response to accumulation of unfolded proteins in the ER (Bush et al., 1994; Partaledis and Berlin, 1993).

1.3.3 Protein disulfide isomerasers

For a large fraction of proteins that fold in the ER, oxidation of cysteins into disulfide bonds is essential to reach the native conformation (Tu and Weissman, 2004). The ER favors the formation of disulfide bonds as it is relatively more oxidizing than the cytosol and it harbors enzymes that catalyze the formation of native disulfides called protein disulfide isomerasers. General feature of this protein family is the presence of one or more thioredoxin-like domains. These domains can be catalytically active, then termed a domain or catalytically inactive, termed b domain. A catalytically active domain comprises two cysteins in a CXXC motif, being the redox-active site. Most commonly XX represents GH (Appenzeller-Herzog and Ellgaard, 2008). PDIs may oxidize or reduce and isomerize their client protein. The preference for oxidation or reduction depends on the surrounding residues within and outside of the active motif (Grauschopf et al., 1995; Huber-Wunderlich and Glockshuber, 1998). When the CXXC motif in PDIs is oxidized, the disulfide bond can be transferred via mixed disulfides to the client protein (Frand and Kaiser, 1999). In turn, reduced PDIs can reduce or isomerize their clients. Isomerization may occur by direct rearrangement of intramolecular disulfide bonds (Kersteen et al., 2005) or by cycles of substrate reduction and subsequent oxidation (Schwaller et al., 2003). The mammalian protein disulfide isomerase family has nineteen published members (Appenzeller-Herzog and Ellgaard, 2008), while the protein family in yeast counts five members that are functionally not interchangeable (Norgaard et al., 2001).

1.3.3.1 PDI protein

The PDI protein (Pdi1p) comprises four thioredoxin-like domains following the order $a$, $b$, $b'$, and $a'$. Additionally, it has a C-terminal domain which is highly acidic and shows low calcium-binding affinity but does not seem to have a function in the catalysis of native disulfide bond formation (Hatahet and Rudock, 2007; Macer and
Koch, 1988). The crystal structure of Pdi1p demonstrated that the four thioredoxin-like domains form the shape of a ‘twisted U’. The active sites are at the two topping ends of the U and face each other, suggesting cooperativity (Tian et al., 2006). The non-catalytic sites localize to the surface on the inside of the U, where they are in an area rich in hydrophobic residues. The non-catalytic sites may be crucial for interaction with the substrate and its presentation to the catalytic domains. The b' domain provides the principal peptide-binding site (Klappa et al., 1998). Besides the oxidoreductase activity, PDI also has chaperone activity. It can inhibit the aggregation of non-native proteins that do not contain any disulfide bonds (Cai et al., 1994). For this process the active-site cysteins are not required (Quan et al., 1995).

1.3.3.2 ERp57

ERp57, a mammalian protein disulfide isomerase, has a domain organization like PDI (a, b, b', a') (Hatahet and Ruddock, 2007). It is associated with calnexin/calreticulin and specifically interacts with N-glycosylated proteins (Maattanen et al., 2006; Oliver et al., 1999; Oliver et al., 1997). In ERp57, a positively charged region in the b' domain electrostatically interacts with a negatively charged region localizing to the tip of the P-domains of calnexin and calreticulin (Frickel et al., 2002; Kozlov et al., 2006; Leach et al., 2002; Pollock et al., 2004).

1.3.3.3 ERdj5

Mammalian ERdj5 contains both, thioredoxin-like domains and a J-domain (Cunnea et al., 2003; Hosoda et al., 2003). The four thioredoxin-like domains are of type a and have the sequences: CSHC, CPPC, CHPC and CGPC. Thus, three of them resemble the CXPC active-site motifs of cytosolic and mitochondrial thioredoxins (Powis and Montfort, 2001).

1.3.4 N-linked glycans

The majority of proteins that fold in the ER receive N-glycans (Helenius and Aebl, 2004). For this, a precursor tetradecasaccharide is assembled on a dolichyl-pyrophosphate carrier in the ER membrane. The assembly occurs by sequential
action of glycosyltransferases connecting two N-acetylglucosamines, nine mannoses, and three glucose residues to build the \( \text{Glc}_3\text{Man}_9\text{GlcNAc}_2 \) oligosaccharide (Burda and Aebi, 1999). In eukaryotes the pre-built oligosaccharides are transferred \textit{en bloc} to asparagine residues in the specific amino acid sequon N-X-S/T, X being any amino acid but proline (Figure 1) (Bause, 1983; Bause and Legler, 1981; Gavel and von Heijne, 1990).

Figure 1. Synthesis and transfer of the N-linked core oligosaccharide to a polypeptide chain.

Synthesis of the N-linked core oligosaccharide starts on the cytoplasmic side of the ER. The linear and stepwise biosynthetic pathway starts with a GlcNAc-1-phosphate that is transferred from UDP-GlcNAc to dolichylpyrophosphate. A further GlcNAc-1-phosphate and five mannose residues are added. Then the lipid-linked Man\(_5\)GlcNAc\(_2\) is flipped into the ER. On the luminal side, four mannose and three glucose residues are added. For the luminal sugar extension dolichol-P-Man or dolichol-P-Glc, respectively, is used. The complete Glc\(_3\)Man\(_9\)GlcNAc\(_2\) is transferred by the oligosaccharyltransferase (OST) onto an N-X-S/T sequon in a peptide chain. Depicted is a co-translocational N-glycosylation reaction. ALG stands for 'Asparagine Linked Glycosylation'. The yeast loci performing the specific steps are indicated. Modified from (Helenius and Aebi, 2004).
1.3.4.1 Transfer of the N-glycan

The transfer reaction is catalyzed by the multisubunit enzyme oligosaccharyltransferase (OTase). In mammalian cells and *S. cerevisiae* the OTase is an oligomeric membrane protein, comprising seven to eight non-identical subunits. In *S. cerevisiae* they are Ost1p, Ost2p, Wbp1p, Stt3p, Swp1p, Ost4p, Ost5p, and Ost3/6p (Karaoglu et al., 1995; Kelleher and Gilmore, 2006; Spirig et al., 1997). STT3 proteins are the active subunits of the eukaryotic OTases (Nilsson et al., 2003; Yan and Lennarz, 2002). *S. cerevisiae* encodes one Stt3 protein, which uses preferentially the completely synthesized Glc$_3$Man$_9$GlcNAc$_2$ oligosaccharide as donor substrate. In contrast, in vertebrates, insects and plants, two STT3 proteins (STT3A and STT3B) are present. While STT3A also preferably transfers the complete Glc$_3$Man$_9$GlcNAc$_2$ oligosaccharide, STT3B can also use incomplete core glycans (Kelleher et al., 2003).

1.3.4.2 Mammalian co- and post-translational glycosylation

In co-translocational glycosylation, the N-glycans are transferred to the nascent polypeptide chain as it enters the ER. The transfer reaction occurs after the N-X-S/T motif has emerged into the ER lumen for 65-75 residues of protein between the glycosylation site and the peptidyl-transferase site of the large ribosomal subunit (Deprez et al., 2005; Nilsson et al., 2003; Whitley et al., 1996). But it was found that for some proteins glycosylation can occur after the entire protein has been translocated into the ER lumen (post-translational glycosylation) (Bolt et al., 2005). Recent findings from the Gilmore group suggest that the STT3A and STT3B OTase isoforms cooperate in a sequential manner to glycosylate the full range of acceptor sites in nascent glycoproteins. While the co-translocational glycosylation reaction is primarily accomplished by the STT3A OTase isoform, STT3B OTase can glycosylate sites that have been skipped by STT3A. Some glycosylation sites may only be accessible to STT3B. Post-translocational glycosylation is mediated by STT3B OTase (Ruiz-Canada et al., 2009). Since STT3A has the more exquisite selection for the Glc$_3$Man$_9$GlcNAc$_2$ oligosaccharide than STT3B, co-translocational glycosylation ensures that the
transferred glycans have the correct structure for the N-glycan dependent ER protein quality control pathway (Ruiz-Canada et al., 2009).

1.4 N-glycan assisted protein folding

Transferred oligosaccharides extend for about 30Å from the polypeptide backbone of their host protein. This relatively bulky and hydrophilic appendix substantially changes the biophysical properties of the nascent yet unstructured polypeptide chain. It is likely that some glycans have a role in promoting and stabilizing local structures (Petrescu et al., 2004). But, N-glycans may also have more global effects, such as increased solubility (Helenius and Aebi, 2004).

Importantly, N-glycans also determine the affinity for the folding factors which assist the protein in the folding process: On one hand, the hydrophilicity may (locally) repel the classical chaperones such as BiP, on the other hand, upon specific processing of
the oligosaccharide N-glycan structures are generated that recruit specific lectin chaperones assisting the glycoprotein in the protein folding process (Helenius and Aebi, 2004; Molinari and Helenius, 2000).

1.4.1 N-glycan processing

Immediately after the addition of the core oligosaccharide, the N-glycan is processed by a set of hydrolases and glycosyltransferases. They act in a sequential order and some of these enzymes modify only N-glycans of unfolded proteins. This orchestrated processing results in a set of defined glycan signals that display the folding status of the host protein and are interpreted by a general folding and degradation machinery (Table 1) (Helenius and Aebi, 2004).

1.4.1.1 Glucosidase I

N-glycan trimming is initiated by glucosidase I. It is an α1,2 exo-glucosidase of the glycosylhydrolase (GH) family 63 and a type II membrane glycoprotein. It trims glucose residue n (Figure 2) of the Glc₃Man₉GlcNAc₂ core glycan. Glucosidase I activity was shown to occur as soon as the N-glycan has a distance of 72 residues from the ribosome P-site. Since OST activity requires the same distance, this indicates that the processing can occur immediately after the N-glycan has been transferred to the nascent polypeptide chain (Deprez et al., 2005; Jakob et al., 1998b). In combination with the subsequent processing by glucosidase II, glucosidase I prevents binding of the protein-bound glycan to the OTase and makes the entry of the substrate into the calnexin/calreticulin cycle possible (Helenius and Aebi, 2004).

1.4.1.2 Glucosidase II

Glucose m and l (Figure 2) are removed by the α1,3 exo-glucosidase glucosidase II which is a member of the GH family 31. It is a soluble heterodimer of a catalytic α-, and a regulatory β-subunit (Trombetta et al., 1996). While the β-subunit is dispensable for enzyme activity in vitro (Trombetta et al., 2001), the β-subunit was
shown to be required for the enzyme’s second action, the removal of glucose \( l \), in yeast (Wilkinson et al., 2006) and seems to confer N-glycan substrate specificity for the \( \alpha \)-subunit (Watanabe et al., 2009). The \( \beta \)-subunit has homology to the mannose-6-phosphate receptor. According to the proposed transactivation model, the \( \beta \)-subunit associates with the 6'-tetramannosyl branch of an N-glycan on the polypeptide chain. This might induce a conformational change in the enzyme and allows the correct positioning of the \( \alpha \)-subunit’s active site on another glycan for processing of the glucose \( m \). Therefore, more than one glycan is required for glucosidase II function (Deprez et al., 2005). However, \textit{in vitro} and \textit{in vivo} experiments showed that this transactivation does not seem to be a strict requirement for glucosidase II since methotrexate carrying solely one N-glycan was efficiently processed (Totani et al., 2006). Similar observations were made with BACE (Vanoni et al., 2008).

The function of glucosidase II processing is dual: The first processing step generates the signal for entry into the calnexin/calreticulin cycle, while the second action allows exit from the cycle. The first processing by glucosidase II, removal of glucose \( m \), is relatively rapid. However, glucose \( l \) is differently oriented in space than glucose \( m \) (Petrescu et al., 1997) and trimming of glucose \( l \) requires transient separation and repositioning of glucosidase II. The thereby generated time window may be exploited by the lectin chaperones calnexin and calreticulin to associate with the monoglucosylated N-glycan processing intermediate (Deprez et al., 2005).

1.4.1.3 UDP-glucose:glycoprotein glucosyltransferase

Glycoproteins may fold properly after one single association with a lectin chaperone (Hebert and Molinari, 2007). However, proper maturation of the protein may require more than a unique association. The mammalian ER contains a folding sensor, the UDP-glucose:glycoprotein glucosyltransferase (GT). This protein consists of a large N-terminal domain that binds non-native structures and it has a C-terminal carbohydrate transferase domain (Guerin and Parodi, 2003). It is not fully understood how GT can distinguish between native and non-native structures. However, it may well be that this depends on the presence of exposed hydrophobic residues, mobile loops, non-native determinants or the lack of compactness (Olivari and Molinari,
Yet, GT has affinity for glycoproteins that are in a molten globule-like state, or nearly native folding intermediates (Caramelo et al., 2003; Caramelo et al., 2004; Ritter and Helenius, 2000). Upon detection of such structures, the C-terminal domain of GT transfers a glucose residue onto the terminal α1,2-linked mannose of the A-branch (mannose g, Figure 2) of a high mannose N-glycan. The efficiency decreases with decreasing number of mannoses in branches B and C (Sousa et al., 1992). The reglucosylated N-glycan promotes the reassociation of the maturing protein with the lectin chaperones (Parodi, 2000).

With regard to the spatial requirements, two, apparently opposing, models have been proposed. One suggests that the reglucosylated glycan is in close proximity to the unfolded part (Ritter and Helenius, 2000; Ritter et al., 2005) while the second model proposes independence of their proximity (Taylor et al., 2004). No functional GT homolog has been identified in S. cerevisiae (Fernandez et al., 1994).

1.4.1.4 ER mannosidase I

The product of glucosidase II action, the Man$_9$GlcNAc$_2$ N-glycan, is subjected to processing by the ER α1,2 mannosidase I to yield the final ER N-glycan structure Man$_8$GlcNAc$_2$ (Byrd et al., 1982; Jakob et al., 1998b). Mns1p is a type II membrane glycoprotein. It belongs to the GH47 family and has specificity for the terminal B branch mannose (mannose i, Figure 2), which is conserved in ER mannosidase I (Herscovics et al., 2002). Processing by mannosidase I is relatively slow, and terminates the maturation phase of the protein (mannose timer) (Helenius, 1994). Proteins that have completed folding by the time Man$_8$GlcNAc$_2$ is generated pass the quality control and can leave the ER to their final destination. However, proteins that are still unfolded by then are recognized by members of the GH47 family, which further process the N-glycan to subject the protein to ER-associated degradation.

1.4.2 Lectin chaperones calnexin and calreticulin

The homologous proteins calnexin and calreticulin were initially named after their ability to bind calcium (Fliegel et al., 1989; Wada et al., 1991). Calnexin is composed of a globular carbohydrate-binding domain, folding into a leguminous (L)-type lectin-
like β-sandwich structure (Schrag et al., 2001) and a proline-rich P-domain. This domain adopts a hairpin structure and recruits the accessory oxidoreductase ERp57, which is involved in disulfide bond formation and isomerization (Frickel et al., 2002; Kozlov et al., 2006). While calnexin is a type I transmembrane protein, its paralog calreticulin is soluble (Hebert and Molinari, 2007).

Calnexin and calreticulin bind to monoglucosylated N-glycans (Kapoor et al., 2003). These may have been generated by trimming of glucose m (Figure 2) by glucosidase II, or alternatively by addition of glucose I by GT. Calnexin/calreticulin binding exposes the bound and maturing glycoprotein to ERp57 and generally slows the folding reaction helping to increase the overall efficiency (Helenius and Aebi, 2004). Thus, lectin-chaperone binding helps to control the rate of the folding process and minimizes the formation of non-productive interactions that may lead the protein into off-pathway intermediates or final off-products. While calnexin and calreticulin seem to have the same molecular specificities, they have a distinct set of client substrates. Calnexin is thought to interact with glycans that are closer to the membrane, while calreticulin binds rather peripheral glycans (Molinari et al., 2004; Pieren et al., 2005).

In S. cerevisiae Cne1p has been identified as the functional ortholog of calnexin (Xu et al., 2004). Cne1p has a short cytosolic domain, a lectin domain that mediates association with monoglucosylated substrates, as well as a conserved P-domain that mediates the interaction with the oxidoreductase Mpd1p. Similarly to the calnexin-ERp57 complex, the yeast Mpd1p interacts with Cne1p to form a functional complex that enhances the catalytic activity of the oxidoreductase (Kimura et al., 2005; Olivari and Molinari, 2007). However, no calreticulin ortholog has been identified in S. cerevisiae.

1.4.3 Calnexin/calreticulin cycle

Repeated action of glucosidase II and GT results in alternation between the un- and the mono-glucosylated oligosaccharide state. This directs the maturing host protein into association with the lectin chaperones or dissociation, respectively, which is termed calnexin/calreticulin cycle (Hammond and Helenius, 1993; Hammond and Helenius, 1994; Nauseef et al., 1995; Peterson et al., 1995). For two recombinant proteins, the T cell receptor α-subunit (Van Leeuwen and Kearse, 1997) and transferrin (Wada et al., 1997), it was shown that they undergo repeated de- and
reglucosylation cycles. Yet, it is largely unknown how many rounds mammalian N-glycoproteins require in the calnexin/calreticulin cycle (Ruddock and Molinari, 2006). However, in *S. pombe*, although the GT is an active enzyme, most polypeptides reach the native conformation upon single binding to calnexin (Fanchiotti et al., 1998). In contrast, in *S. cerevisiae* no reglucosylation activity could be detected, implying that the calnexin cycle does not exist in this system (Fernandez et al., 1994).

1.5 Quality control

Newly synthesized proteins that have acquired a properly folded and assembled structure are transported from the ER to their final destination. In contrast, incompletely folded and assembled proteins are retained and eventually degraded. The molecular mechanism of this sorting phenomenon has been called ‘quality control’ (Hammond and Helenius, 1995; Hurtley and Helenius, 1989). For the retention of non-native proteins in the ER, an active mechanism was proposed: While a protein is in the unfolded state, it remains in association with molecular chaperones and folding enzymes. In contrast, when the protein reaches its native conformation, the structures and signals to promote association with the chaperones and folding enzymes are no longer exposed (e.g. hydrophobic patches) or present (e.g. monoglucosylated N-glycans). Thus, the protein is released from the folding machinery and may leave the ER (Ellgaard and Helenius, 2003).

1.6 Export of proteins from the ER

The release of proteins that have passed quality control occurs in transport vesicles that are coated with coat protein complex-II (COPII) proteins, budding at ER exit sites (Gurkan et al., 2006). While in yeast the COPII-coated cargo vesicles are directly delivered to the Golgi compartment (Bonifacino and Glick, 2004), in mammalian cells, the COPII-coated vesicles undergo homotypic fusion to generate a stationary ER-Golgi intermediate compartment (ERGIC). From there, the cargo proteins are transported in COPI-coated vesicles to the cis-Golgi (Appenzeller-Herzog and Hauri, 2006; Aridor et al., 1995; Pepperkok et al., 1993). Concentration of proteins in COPI vesicles for ER exit is mediated by ER export signals within the cargo proteins.
interacting either directly with COPII components or with ER export cargo receptors (Barlowe, 2003; Lee et al., 2004a). Erv29p is such a yeast ER export cargo receptor responsible for the packaging of soluble cargo proteins such as correctly folded carboxypeptidase Y (CPY), proteinase A (PrA), and glycosylated pro-α-factor (gpαf) (Belden and Barlowe, 2001; Caldwell et al., 2001). Indeed, a specific sequence within gpαf was identified to be responsible for binding to Erv29p (Otte and Barlowe, 2004). Membrane-spanning proteins interact with components of the COPII coat involved in cargo selection (Miller et al., 2003; Votsmeier and Gallwitz, 2001) due to specific cytosolic exit signals, including the di-acidic D/E-X-D/E motif (Nishimura and Balch, 1997).

1.6.1 N-glycans in ER export

N-glycans are also involved in the transport of proteins from the ER to the Golgi. In mammals, N-glycosylated proteins are recognized by a group of cargo receptors that belong to the family of L-type lectins. L-type lectins possess a luminal carbohydrate recognition domain, composed of an antiparallel β-sheet that binds to high-mannose-type oligosaccharides in a calcium-dependent manner (Appenzeller-Herzog et al., 2004; Kamiya et al., 2005; Yamaguchi et al., 2007). The most completely characterized is ERGIC-53 (Kamiya et al., 2008). It cycles between the ER and ERGIC and functions as cargo receptor of several soluble glycoproteins like pro-cathepsin C and pro-cathepsin Z (Appenzeller et al., 1999; Vollenweider et al., 1998), coagulation factors V and VIII (Zhang and Herscovitz, 2003), and α1-antitrypsin (Nyfeler et al., 2008). ERGIC-53 binds high-mannose-type oligosaccharides, with a moderate tendency to bind better the larger oligosaccharides (Kamiya et al., 2008), in accordance with the N-glycan structures of folded (and exported) glycoproteins, being Man₈GlcNAc₂. Vesicular integral protein 36 (VIP36), is concentrated in the cis-Golgi and might facilitate the transport of glycoproteins from ERGIC to the cis-Golgi. Though, an alternative model has been postulated that VIP36 rather retrieves glycoproteins from the Golgi back to the ER (Kamiya et al., 2008).
Chapter 1

1.7 Quality control beyond ER

Despite quality control in the ER, there is a number of proteins that exit the ER and traffic to the Golgi when they are non-native (Caldwell et al., 2001; Coughlan et al., 2004; Hong et al., 1996; Taxis et al., 2002; Vashist et al., 2001; Vashist and Ng, 2004), given that functional ER-exit signals are present (Kincaid and Cooper, 2007). It has been proposed that quality control checkpoints may exist beyond the ER, in order to detect non-native proteins that have escaped from the ER or that may misfold after their exit. L-type lectins may be involved in such post-ER quality control processes (Freeze et al., 2009).

1.7.1 N-glycans in post-ER quality control

Recently, it was shown that VIP36 binds high-mannose-type oligosaccharides with an intact trimannosyl A branch. But processing of the α1,2-linked mannoses of either the B or the C branch (mannose i or k, Figure 2), or both, results solely in minor reduction of VIP36 binding. Since VIP36 has a slightly acidic pH optimum, corresponding to the luminal pH of the Golgi, it was postulated that VIP36 acts in post-ER quality control, and that it may retrieve non-native glycoproteins back to the ER (Kamiya et al., 2008). This notion is interesting, as the proposed N-glycan signal for ER-associated degradation is the Man7GlcNAc2 structure, lacking the terminal B and C branch mannose (Clerc et al., 2009; Quan et al., 2008). However, it has also been suggested that misfolded proteins can be degraded in a post-Golgi compartment without necessarily recycling to the ER (Arvan et al., 2002).
Table 1. Overview on N-glycan structures and their function, generated in the ER of *S. cerevisiae*.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Generated by</th>
<th>Signal for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glc$_3$Man$_9$</td>
<td>Stt3p</td>
<td>N-glycan dependent ER protein quality control pathway</td>
</tr>
<tr>
<td>Glc$_3$Man$_9$</td>
<td>Glucosidase I</td>
<td>Release from OTase Preparation for lectin chaperone</td>
</tr>
<tr>
<td>Glc$_1$Man$_9$</td>
<td>Glucosidase II</td>
<td>Association with lectin chaperone</td>
</tr>
<tr>
<td>Man$_9$</td>
<td>Glucosidase II</td>
<td>Dissociation from lectin chaperone</td>
</tr>
<tr>
<td>Man$_9$</td>
<td>ER mannosidase I</td>
<td>Folding phase is terminated</td>
</tr>
<tr>
<td>Man$_7$C</td>
<td>Htm1p</td>
<td>Degradation</td>
</tr>
</tbody>
</table>
2 ER-associated degradation

2.1 Introduction

Even though the ER is conducive for protein folding, a polypeptide chain may still enter off-pathways and therefore not, or not within biologically acceptable time, reach the native tertiary and quarternary structure. Indeed, it seems that some proteins have an intrinsically low folding efficiency, as the cystic fibrosis transmembrane conductance regulator (CFTR) has an estimated yield of folded proteins of around 25% (Kopito, 1999). Moreover, polypeptide chains with faulty sequence may be produced during transcription and translation, having estimated error rates of \(1 \times 10^{-3}\) to \(1 \times 10^{-4}\) (Hebert and Molinari, 2007). The resulting defects like mutations, deletions or truncations may directly disturb folding. Independent of the cause, folding-incompetent polypeptides need to be disposed in order to maintain ER homeostasis (Meusser et al., 2005) and to keep the capacity to assist further, newly synthesized polypeptide chains entering the ER in their maturation (Eriksson et al., 2004). The overall extent of degradation in a cell may vary: A first study determined a fraction of more than 30% that was subjected to rapid degradation (Schubert et al., 2000); however, a later study reported much lower turnover rates (Vabulas and Hartl, 2005). Generally, proteins which do not achieve the proper fold in due time, as well as excess subunits of protein complexes, are retained in the ER and get ultimately degraded by the cytoplasmically located 26S proteasome. The process of recognition, retrotranslocation, ubiquitylation and degradation of terminally misfolded proteins is referred to as ER-associated degradation (ERAD) (Brodsky and McCracken, 1999; Plemper and Wolf, 1999a; Sommer and Wolf, 1997). The location of the misfolded lesion dictates which factors are engaged for ERAD of a substrate. In *S. cerevisiae*, the terminology ERAD-L, ERAD-C and ERAD-M (L, C and M stand for Lumen, Cytosol and Membrane, respectively) has been introduced for description of the different pathways taken by the ERAD substrates with lesions in luminal, cytosolic, or membrane domains (Bonifacino et al., 1990; Taxis et al., 2003; Vashist and Ng, 2004). In the following, the focus will be on the ERAD-L pathway of N-glycosylated proteins.
2.2 Distinction between unfolded and terminally misfolded

Native proteins do not expose motifs like hydrophobic patches or unpaired cysteins. Therefore, they no longer elicit binding of chaperones or folding enzymes and it appears evident that the cell can distinguish between folded and non-folded proteins. In contrast, how an unfolded protein can be differentiated from a terminally misfolded protein, is less obvious: both conformations are aggregation prone and attract molecular chaperones. How this process is regulated, is better known for N-glycosylated proteins. It is widely accepted that the oligosaccharide appendices and their processing by glucosidases and glucosyltransferases play a crucial regulatory function in this process (Hebert et al., 2005).

2.2.1 Importance of N-glycans

Analysis in *S. cerevisiae* revealed that the ERAD model substrate CPY* (the G255R mutant of the N-glycosylated vacuolar protease carboxypeptidase Y) does not correctly fold. Instead, it is retained in the ER and rapidly degraded by ERAD (Finger et al., 1993). In contrast, a non-glycosylated version of CPY* is retained but it does not get degraded (Knop et al., 1996).

2.2.1.1 Importance of mannose trimming - Mannose timer

Both, in *S. cerevisiae* and in higher eukaryotes, the importance of mannose trimming for ERAD of misfolded glycoproteins has been reported. In *S. cerevisiae* mannosidase I deletion mutants showed a delay in glycoprotein degradation (Jakob et al., 1998a; Knop et al., 1996), while mammalian cells revealed impaired degradation after treatment with the \( \alpha \)-mannosidase inhibitors kifunensine and 1-deoxymannojirimycin (Su et al., 1993). In turn, overexpression of mammalian ER mannosidase I accelerates glycan trimming and degradation (Hosokawa et al., 2003). The finding that inhibition of the ER mannosidase I protected defective folding-peptides from disposal led to the concept of a mannose timer (Helenius, 1994). Since ER \( \alpha \)-mannosidase is acting slowly, compared to the glucosidases, it was proposed that this enzyme might act as a molecular timer and that the Man\(_9\)GlcNAc\(_2\) glycan structure may function as the N-glycan degradation signal (Jakob et al., 1998a).
However, processing to the Man\textsubscript{8}GlcNAc\textsubscript{2} structure alone would not be sufficient because only misfolded proteins with this structure are degraded (Jakob et al., 1998a).

### 2.2.1.1.1 Processing of Man\textsubscript{8}GlcNAc\textsubscript{2} in S. cerevisiae

Recent analysis in S. cerevisiae revealed that the removal of one further \(\alpha\)1,2-linked mannose from the Man\textsubscript{8}GlcNAc\textsubscript{2} structure is required for degradation of misfolded glycoproteins. It was shown that the N-glycan degradation signal, which is a free \(\alpha\)1,6-linked mannose, is generated by processing of the C-branch (removal of mannose k, Figure 2) of the Man\textsubscript{8}GlcNAc\textsubscript{2} glycan (Clerc et al., 2009; Quan et al., 2008). This structure is generated by Htm1p which functionally depends on prior processing of the N-glycan by Gls1p, Gls2p and Mns1p (Clerc et al., 2009).

### 2.2.1.1.2 Processing of Man\textsubscript{8}GlcNAc\textsubscript{2} in higher eukaryotes

In higher eukaryotes further processing of the Man\textsubscript{8}GlcNAc\textsubscript{2} was observed, and it turned out that the generation of Man\textsubscript{7}GlcNAc\textsubscript{2}, Man\textsubscript{6}GlcNAc\textsubscript{2} and Man\textsubscript{5}GlcNAc\textsubscript{2} glycan structures precedes or elicits degradation (Ermonval et al., 2001; Foulquier et al., 2004; Foulquier et al., 2002; Frenkel et al., 2003; Hosokawa et al., 2003; Kitzmuller et al., 2003). Interestingly, processing of \(\alpha\)1,2 mannoses was as well required in the cell lines B3F7 and MadIA214, which do not have the cleavable mannose on branch B, but a complete A branch (Ermonval et al., 2001; Olivari et al., 2006). This led to a modification of the hypothesis that Man\textsubscript{8}GlcNAc\textsubscript{2} may be the N-glycan degradation signal: While the Man\textsubscript{8}GlcNAc\textsubscript{2} glycan still exposes the mannose on branch A, which is the sole residue of the protein-bound oligosaccharide that can be reglucosylated by GT, cycling in the calnexin chaperone system and with this protection from degradation may still occur. For irreversible extraction of folding-defective glycopolypeptides removal of the mannose on branch A is required (Hebert and Molinari, 2007).

Based on the data from the ERAD competent S. cerevisiae model system in which the calnexin/calreticulin cycle is absent, a further extension of the mammalian working model has been proposed: While the removal of the terminal \(\alpha\)1,2-linked
mannose of the A branch ensures the exit from the calnexin/calreticulin cycle, additional trimming of the C branch generates the terminal misfolding signal, the α1,6-linked mannose which targets the N-glycoprotein for ERAD (Clerc et al., 2009). Recent investigations support this notion, as indeed a terminal α1,6-linked mannose is the specific functionality, recognized by the human ER-degradation lectin OS-9 (Hosokawa et al., 2009).

2.2.1.1.3 Mannose processing in *S. pombe*?

ER mannosidase I from *S. pombe* is a functional enzyme, though it seems to have poor activity *in vivo* and the primary N-glycan structure detected on misfolded glycoproteins is Man₉GlcNAc₂. Solely minor processing was observed. Nevertheless, CPY* degrades efficiently in wild type cells, while it depends on Spmns1p and Sphtm1p (Movsichoff et al., 2005). Movsichoff et al. challenged the mannosidase timer hypothesis and proposed that Spmns1 (and Sphtm1) might rather act as a lectin-like factor in ERAD (Movsichoff et al., 2005). However, it is interesting to note, that the smallest detected glycan species was Man₇GlcNAc₂ exposing an α1,6-linked mannose (Movsichoff et al., 2005). Moreover, *S. pombe* also has a Yos9p homolog (Banerjee et al., 2007). One may speculate that the ER-degradation signal in *S. pombe* might also be the free α1,6-linked C branch mannose, potentially generated by the concerted action of Spmns1p and Sphtm1p on specific N-glycans.
Figure 3. Quality control and ERAD-L in *S. cerevisiae*.

Top left: Immediately after the addition of the core N-glycan to an unfolded protein (black line with yellow stars), the outermost of the three glucoses, glucose n, is trimmed by glucosidase I. Then, glucosidase II trims the second glucose (m). The glycoprotein carrying the monoglucosylated N-glycan, binds to the lectin chaperone Cne1p/calnexin. Cne1p exposes the polypeptide chain to the protein disulfidemersase Mpd1p/ERp57 (grey), supporting the formation of disulfide bonds in unfolded protein structures. Processing of the remaining glucose (l) by glucosidase II, results in the dissociation of the protein from the lectin chaperone. The N-glycan is further processed by Mns1p that slowly removes the terminal mannose of the B branch (i). Then, the protein encounters the quality control checkpoint Htm1p/Pdi1p: If it is correctly folded (black line without yellow stars), it is not recognized by Htm1p and ready for export. In contrast, if it is still unfolded (black line with a yellow star), it is sequestered by Htm1p, which processes the C branch of the N-glycan (mannose k), in order to mark the protein 'misfolded'. The hereby generated free α1,6 mannose, is recognized by the proof-reading factor Yos9p in Yos9p/Hrd3p/Kar2p subcomplex, associated with the Hrd complex (Figure 4). The misfolded glycoprotein is targeted to retrotranslocation and proteasomal degradation. Fast N-glycan processing steps are depicted with dotted arrows; slow processing is represented by a continuous arrow. Yellow stars represent unfolded structures. Blue dots represent chaperones and folding enzymes that support protein folding in a N-glycan independent manner.
2.2.1.2 Enzymes implicated in the processing of Man₈GlcNAc₂

Enzymes that have been implicated in the processing of the Man₈GlcNAc₂ glycan belong to the family GH47 of the glycoside hydrolase classification glycoproteins (Molinari, 2007; Moremen and Molinari, 2006). The first subgroup of the GH47 family comprises the ER mannosidase I; the second subgroup includes the Golgi α1,2 mannosidases IA, IB, and IC; and the third subgroup comprises the ER degradation enhancing α-mannosidase-like (EDEM) and the Homologous to mannosidase I (Htm1p) proteins.

2.2.1.3 Htm1p and EDEM proteins

2.2.1.3.1 Comparison and description

*HTM1* encodes a protein of 796 amino acids, EDEM1 a protein of 652 amino acids, EDEM2 and EDEM3 one of 578 or 931 amino acids, respectively. Sequence comparison of ER mannosidase I, EDEM1/2/3 and Htm1p reveals a mannosidase homology region with significant sequence identity among the homologs. EDEM proteins conserve all catalytic residues required for glycolytic activity and structural modeling indicates no difference in their spatial location (Moremen and Molinari, 2006). However, the proteins differ in their remaining domains. One particular characteristic is the presence of a C-terminal extension which varies among the EDEM/Htm1p homolog (Kanehara et al., 2007; Mast et al., 2005). In Htm1p, this domain constitutes almost one third, in EDEM3 almost half of the whole protein. The only member of the EDEM family that reveals an obvious ER retrieval signal is EDEM3 being a KDEL motif, suggesting that Htm1p, EDEM1 and EDEM2 require association with other ER resident proteins (Kanehara et al., 2007).

EDEM1 was found as a type II ER transmembrane protein, when determined in COS cells, but it seems to be a luminal protein in HEK293 cells (Olivari et al., 2005). Alkaline extraction from canine pancreas RER microsomes, and membranes from HepG2 cells indicated that EDEM1 exists as both, a soluble and membrane-associated glycoprotein (Zuber et al., 2007). The contrasting results were explained by variations in signal peptide cleavage in different cell lines (Olivari et al., 2005).
Chapter 1

However, EDEM2 and EDEM3 are soluble proteins (Mast et al., 2005).

Based on the UniGene expression profile, the three mammalian proteins are frequently expressed in the same organs but at different levels (Olivari and Molinari, 2007).

While the mammalian EDEM proteins are major targets of the ER-stress-induced Ire1/Xbp1 pathway, S. cerevisiae HTM1 is not UPR up-regulated (Travers et al., 2000). The Ire1/Xbp1 pathway is used to enhance the capacity for ERAD in response to an increase in cargo load and/or accumulation of misfolded polypeptides (Hebert and Molinari, 2007).

EDEM1 was detected in small (about 150nm-sized) ER-derived vesicles that lack conventional ER markers, as well as a COPII coat (Zuber et al., 2007). Moreover, EDEM1 was found to have an unusual short half-life, compared to other ER-resident folding factors, and an ERAD tuning mechanism was proposed to contribute to the maintenance of ER homeostasis: The selective segregation and rapid turnover of ERAD factors reduces the competition between the folding and the degradation machinery, at the advantage of the folding process (Cali et al., 2008). While Cali et al. proposed a novel mechanism that is distinctive from macroautophagy, a recent report shows that endogenous EDEM1 is degraded by basal autophagy (Le Fourn et al., 2009).

2.2.1.3.2 Importance for ERAD

EDEM1

EDEM1 function was first analyzed in COS cells where EDEM1 overexpression was found to accelerate the turnover of the NHK variant of α1-antitrypsin and it was found in association with NHK α1-antitrypsin in co-immunoprecipitation experiments (Hosokawa et al., 2001; Molinari et al., 2003).

In mammalian quality control, the time that the synthesized polypeptide is allocated to reach the native structure, is for ERAD substrates seen as a lag phase before degradation onset (Molinari, 2007). In this, the N-glycosylated ERAD candidates are trapped in the calnexin/calreticulin chaperone system and undergo formation and/or
isomerization and/or reduction of intra- and intermolecular disulfide bonds, attempting to reach a conformation that passes the ER quality control (Molinari et al., 2005). Overexpression of EDEM1 leads to the extraction of the misfolded proteins from the calnexin/calreticulin cycle, as it reduces the lag phase and promotes an earlier onset of protein degradation (Molinari et al., 2003), what seems to be facilitated by the direct physical interaction of calnexin and EDEM via their transmembrane domains (Molinari et al., 2003; Oda et al., 2003).

Moreover, EDEM1 was found to prevent the formation of disulfide-bonded dimers (Hosokawa et al., 2006). In line with this, EDEM1 associates with ERdj5, which regulates the degradation of the NHK α1-antitrypsin variant, as it accelerates the formation of degradation-competent NHK monomers from disulfide-linked dimers (Ushioda et al., 2008).

EDEM1 has also been shown to participate in the retro-translocation of the active subunit of the ricin toxin (Slominska-Wojewodzka et al., 2006). And, EDEM1 was co-immunoprecipitated with Derlin-2 and Derlin-3, which are homologs of the yeast Der1p (Oda et al., 2006). The Derlin family of integral membrane proteins has been proposed to form all or part of the membrane channel for substrate retrotranslocation (Lilley and Ploegh, 2004; Ye et al., 2001).

**EDEM2**

EDEM2 overexpression accelerates the degradation of two variants of the glycosylated α1-antitrypsin, NHK and PI Z, but not of a non-glycosylated variant. Neither in vivo nor in vitro analysis revealed α1,2-mannosidase activity for EDEM2 (Mast et al., 2005; Olivari et al., 2005). The similarity of EDEM1 and EDEM2 function and substrate specificity suggested that these proteins might perform semi-redundant functions. Indeed, BACE457Δ degradation was only modestly affected, when solely EDEM1 expression had been knocked-down (Molinari et al., 2003).

**EDEM3**

EDEM3 overexpression also accelerates the degradation of glycosylated ERAD substrates, as was determined with NHK and the unassembled α subunit of the T-cell receptor (Hirao et al., 2006).
**Htm1p**

Initial reports describing the function of *S. cerevisiae* Htm1p provided evidence that yeast deleted in the HTM1 locus does not exhibit defects in glycan trimming in vivo. However, they reveal strong defects in the degradation of misfolded glycoproteins that may be soluble or membrane proteins like CPY* or mutated variant of Stt3p, the Stt3-7. In contrast, the degradation rates of non-glycosylated ERAD substrates, a misfolded variant of proα-factor (ΔGpαF) and the unstable Sec61-2p translocon subunit, were unaffected by the deletion (Jakob et al., 2001; Nakatsukasa et al., 2001). Further genetic analysis suggested that Htm1p functions in the ERAD-L pathway (Vashist and Ng, 2004), upstream of the Hrd1p complex (Carvalho et al., 2006; Denic et al., 2006; Gauss et al., 2006a). Moreover, Htm1p was found to physically interact with Pdi1p via its C-terminal domain (Clerc et al., 2009; Sakoh-Nakatogawa et al., 2009). This interaction includes an intermolecular disulfide bond, and is essential for subsequent introduction of a disulfide bond in the mannosidase homology domain of Htm1p by Pdi1p (Sakoh-Nakatogawa et al., 2009).

2.2.1.3.3 Mannosidase function

EDEM3 was the first member of this GH 47 protein subfamily that was shown to have mannosidase activity. Its overexpression leads to mannose trimming in vivo to Man₇GlcNAc₂, Man₆GlcNAc₂ and to lesser extent to Man₅GlcNAc₂ (Hirao et al., 2006). Contrary to the initial reports, EDEM1 has also been proposed to function as processing mannosidase, for overexpression of EDEM1 enhances ERAD substrate demannosylation. This was found both, in wild type as well as in B3F7 cells, lacking mannoses on both, the B and the C branch (Olivari et al., 2006). On this basis, a novel model for ERAD function of mammalian EDEM proteins was formulated by the Molinari group: EDEM1 (possibly also EDEM2 and EDEM3) would be activated under ER stress conditions, leading to rapid and extensive demannosylation of N-glycans, and in particular their A-branch. This reduces the reglucosylation of the substrates by GT and accelerates the clearance of aberrant proteins (Olivari et al., 2006).

Recently, it was shown that *S. cerevisiae* Htm1p is also functioning as an active α1,2-mannosidase in ERAD. It was found, that it removes the terminal α1,2-mannose of the C branch from N-glycans on misfolded glycoproteins that have been processed.
to Man₈GlcNAc₂. By this, it generates a free α1,6-linked mannose which is the N-glycan ER-degradation signal, that is recognized by the ER degradation receptor Yos9p (Clerc et al., 2009; Quan et al., 2008). The functional dependence on prior processing by glucosidase I, glucosidase II and mannosidase I ensures, that the proteins can interact with the folding machinery, and that the proteins get a fair time window for folding (dependence on the mannose timer), before the N-glycan signal for protein destruction is generated. The current working model suggests, that Htm1p is the key enzyme that makes the distinction between unfolded and terminally misfolded proteins: If the protein signal is ‘non-folded’ and the glycan signal is Man₈GlcNAc₂, Htm1p generates the free α1,6-linked mannose, the signal for ‘terminal misfolding’. How Htm1p selects non-folded proteins remains elusive. However, it may be that the protein substrate specificity is mediated by Pdi1p (Figure 3).

Analysis of the N-glycan substrate specificity of Htm1p did not reveal a clear resemblance to one of the other two defined GH47 family subgroups. While Htm1p removes one mannose (similar to subgroup 1), it processes the C-branch of the glycan (similar to subgroup 2). Moreover, subgroup 1 and subgroup 2 members act complementary and independently of each other whereas Htm1p is functionally dependent on Mns1p (Chapter 4).

2.3 ERAD substrate recognition

2.3.1 Yos9p, OS-9 and XTP3-B proteins

The finding that N-glycans and specifically their structure is important in the degradation of misfolded glycoproteins, led early to the proposal that a lectin receptor is involved in the recognition of ERAD substrates (Jakob et al., 1998a; Liu et al., 1997; Yang et al., 1998). Yos9p and mammalian OS-9 and XTP3-B constitute group of proteins that have lectin-like domains with homology to the mannose-6-phosphate receptor (MPR) family. They are implicated in the recognition of misfolded glycoproteins for degradation.
2.3.1.1 Yos9p

2.3.1.1.1 Description

YOS9 (yeast osteosarcoma 9) gene was named after its similarity to OS-9, a gene found amplified in chromosomal DNA from human tumors (Su et al., 1996). YOS9 encodes a 542 amino acid protein with one MRH domain (Munro, 2001); and a C-terminal HDEL sequence. It has an N-terminal signal sequence which is cleaved. Early, biochemical analysis suggested that Yos9p is peripherally associated with the ER membrane (Friedmann et al., 2002). Studies with the Hrd1 E3 ubiquitin ligase revealed that Yos9p is part of the Hrd1 complex (Carvalho et al., 2006; Denic et al., 2006; Gauss et al., 2006a).

2.3.1.1.2 Importance for ERAD

Yos9p was found to participate in glycoprotein turnover in a genetic screen (Buschhorn et al., 2004). It is required for the efficient degradation of N-glycosylated proteins but not non-glycoproteins and it is involved in the degradation of soluble and membrane proteins. Yos9p was found to bind specifically to misfolded, but not to folded proteins, as was determined in co-immunoprecipitation and cross-linking experiments (Bhamidipati et al., 2005; Kim et al., 2005; Szathmary et al., 2005). The finding that mutation of the MRH-domain abolished ERAD function provided indirect evidence that Yos9p functions as lectin in ERAD (Bhamidipati et al., 2005; Kim et al., 2005; Szathmary et al., 2005).

2.3.1.1.3 Proof-reading function

Yos9p is part of the Hrd1p complex to which it is bound via Hrd3p, a transmembrane protein with a large luminal domain (Figure 4). The N-glycan-MRH-Yos9p interaction does not seem to be essential for N-glycoprotein substrate recruitment to the Hrd1p complex, but Yos9p was proposed to perform a proof-reading or gating function in the complex (Denic et al., 2006; Gauss et al., 2006b). In this, Hrd3p recruits the misfolded proteins while Yos9p scans the substrate for the correct N-glycan structure, signaling terminal misfolding. Recently, the N-glycan substrate specificity of
Yos9p was determined and revealed that Yos9p is binding N-glycans that have a free α1,6-linked mannose (Quan et al., 2008), corroborating the working model that this N-glycan structure, which is generated by Htm1p (Clerc et al., 2009), is the N-glycan signal for glycoprotein degradation.

2.3.1.2 OS-9 and XTP3-B

Recent studies revealed that the two MRH-domain-containing lectins, OS-9 and XTP3-B, are ER luminal proteins that are involved in ER quality control. Moreover, they were found in large complexes containing the HRD1-SEL1L (homologs of Hrd1p and Hrd3p) ubiquitin-ligase. (Alcock and Swanton, 2009; Bernasconi et al., 2008; Christianson et al., 2008; Hosokawa et al., 2008; Mueller et al., 2008). The components of this mammalian complex are similar to the one in S. cerevisiae, suggesting evolutionary conservation. Studies using lectin mutants suggested that the MRH domains may not be required for binding to ERAD substrates, but rather for interaction with SEL1L (Christianson et al., 2008), which carries several N-glycans (Lilley BN and Ploegh HL, 2005, PNAS; Mueller B, Lilley BN et al., JCB, 2006). Also similar to Yos9p, the lectin activity appears to be dispensable for OS-9 binding to misfolded glycoproteins (Alcock and Swanton, 2009; Bernasconi et al., 2008; Christianson et al., 2008). Additionally, OS-9 facilitates the ubiquitylation of misfolded glycoproteins (Alcock and Swanton, 2009). Finally it was shown that the lectin activity of OS-9 is required for enhancement of glycoprotein ERAD, and that it specifically binds N-glycans lacking the terminal mannose from the C branch (Hosokawa et al., 2009). OS-9 is upregulated in response to ER stress upon activation of the Ire1/Xbp1 pathway (Alcock and Swanton, 2009; Bernasconi et al., 2008). The precise function of XTP3-B remains elusive.
The Hrd1p complex includes the E2 ubiquitin conjugating enzyme Ubc7p which is tethered to the membrane via Cue1p. Through Ubx2p the AAA⁺ ATPase Cdc48p with its cofactors Npl4p and Ufd1p is attached (Carvalho et al., 2006). Furthermore, Hrd1p binds in a 1:1 interaction the transmembrane protein Hrd3p, which has a large luminal domain (Gardner et al., 2000), and in turn attaches Yos9p to the complex. Additionally, Der1p is bound to Hrd1p through Usa1p. The Yos9p/Hrd3p/Kar2p subcomplex is believed to specifically recognize and target misfolded proteins to degradation. While ERAD-M substrates are also believed to use the Hrd1p complex, the ERAD-C pathway engages a different complex, termed Doa10p complex.

2.4 ERAD substrate retrotranslocation

After the recognition step, the ERAD substrates need to be dislocated from the ER back to the cytosol. ERAD substrates are kept in a retrotranslocation-competent state by the action of classical chaperones and proteins engaged in disulfide bond formation during productive protein folding, like PDI.

2.4.1 Classical chaperones

The ER-luminal Hsp70-family member BIP/Kar2p, as well as ER resident Hsp40-family members, like ERdj4 and ERdj5, were found in association with several ERAD substrates, maintaining their solubility for retrotranslocation (Dong et al., 2008;
Knittler et al., 1995; Nishikawa et al., 2001; Schmitz et al., 1995). Whether for all ERAD substrates binding by classical chaperones is required, is unclear to date (Vembar and Brodsky, 2008). However, Kar2p was also found to be a member in the Hrd1p complex (Denic et al., 2006); and, cytoplasmic Hsp70 protein members were shown to recruit ubiquitin-ligases in the process of CFTR degradation (Meacham et al., 2001). This indicates that Hsp70s may have a general and important role during ERAD, as prolonged interaction between an ERAD substrate and a Hsp70 might be sufficient to recruit a ubiquitin ligase (Vembar and Brodsky, 2008).

2.4.2 Protein disulfide isomerases

Besides classical chaperones, proteins engaged in disulfide bond formation during productive protein folding, like PDI are also involved in ERAD of several substrates (Molinari et al., 2002; Svedine et al., 2004). PDI was found to facilitate retrotranslocation of cholera toxin and simian virus 40 (SV40) polyoma virus. Both of them exploit the endocytic pathway to invade host cells (Schelhaas et al., 2007; Tsai et al., 2001). Retrotranslocation of the SV40 virion was also shown to be supported by ERdj5 (Schelhaas et al., 2007). And, ERdj5 was found to regulate the degradation of the disease-causing null Hong Kong (NHK) α1-antitrypsin variant, as it accelerates the formation of degradation-competent NHK monomers from disulfide-linked dimers (Ushioda et al., 2008). This suggests that the function of ERdj5 in ERAD depends on its oxidoreductase activity and is in line with the observation that degradation efficiencies can be enhanced when cells are treated with reducing agents (Valetti and Sitia, 1994). However, yeast Pdi1p is apparently able to support degradation of cysteine-free ERAD substrates (Gillece et al., 1999). In general, it is not known to which extent ERAD substrates need to be reduced, but it seems that complete reduction is not a strict requirement since at least for some substrates completely reduced forms cannot be identified prior to degradation (Molinari et al., 1997; Okuda-Shimizu and Hendershot, 2007).

2.4.3 Retrotranslocon

The retrotranslocon has not yet been identified, though several candidates have been proposed to function in the ERAD substrate dislocation.
2.4.3.1 Translocon

The Sec61 translocon has early been suggested to function as the retrotranslocation channel for various reasons. For instance, human major histocompatibility complex class I (MHC class I) and a non-glycosylated yeast mating pheromone (pαF) were found in co-precipitates or in cross-links with Sec61 (Wiertz et al., 1996). Consistently, mutations in SEC61 slowed the degradation of pαF and CPY* (Pilon et al., 1997; Plemper et al., 1997). Moreover, the binding of ribosomes to ER-derived vesicles abrogated the retrotranslocation of cholera toxin (Schmitz et al., 2000). Also, interactions between proteasome subunits and Sec61 were reported (Kalies et al., 2005). However, Sec61 was not found to be a member of the Hrd1 or the Doa10 complex (Carvalho et al., 2006; Denic et al., 2006; Gauss et al., 2006a). Nevertheless, the formation of disulfide bonds between an ERAD substrate, which had been trapped prior to degradation and Sec61 was found (Scott and Schekman, 2008). Recently, genetic dissection of the essential Sec61 translocation- and ERAD-function revealed that Sec61 is required for ERAD-L (Willer et al., 2008).

2.4.3.2 Derlin-1

Alternatively, the derlin proteins may function as retrotranslocation channel. Der1p is a component of the Hrd1p complex (Carvalho et al., 2006), what would provide ideal spatial conditions to function as retrotranslocation channel. Derlin-1, one of the three derlin homologues in human cells, was shown to be involved in the turnover of MHC class I and interacts with components of the targeting and ubiquitylation machinery (Lilley and Ploegh, 2004; Ye et al., 2004). An in vitro assay with semi-reconstituted ER-derived vesicles also revealed Derlin-1 to be required for retrotranslocation (Wahlman et al., 2007) and the dislocation of the SV40 virion also depends on Derlin-1 (Schelhaas et al., 2007).

2.4.3.3 E3 ubiquitin ligases

Moreover, the Hrd1p and Doa10p E3 ubiquitin ligases, being multi-spanning membrane proteins, have themselves been proposed to function as, or to contribute integral components for the retrotranslocation channel (Kostova et al., 2007; Kreft et
Similarly, also the human gp78 E3 ubiquitin ligase, which is a polytopic, integral membrane protein was proposed to function as retrotranslocation channel (Zhong et al., 2004).

2.4.3.4 Alternative concepts

It was suggested that the retrotranslocon might be formed only transiently and/or be composed of several components (Nakatsukasa and Brodsky, 2008). Alternatively, a retrotranslocation channel might not always be required, as membrane proteins might be extracted directly from the membrane (Nakatsukasa and Brodsky, 2008); or they might leave the ER through lipid droplets (Ploegh, 2007). In the case of membrane proteins a combination is also thinkable: The proteasome might clip cytoplasmic domains, while the resulting intra-membrane and luminal domains might be transported to different cellular compartments for degradation, or become substrates for other, yet ill-defined, proteases (Vembar and Brodsky, 2008).

2.5 ERAD substrate ubiquitylation

Most ERAD substrates are ubiquitylated prior to proteasome targeting.

2.5.1 Ubiquitylation reaction

In the ubiquitylation reaction, the 76 amino acid peptide ubiquitin (Ub) is covalently attached to ε-amino groups of lysine residues in the substrates. The transfer of ubiquitin to the substrate requires a cascade of actions involving E1, E2, E3. Initially, the C-terminal glycine of the ubiquitin is adenylated by the E1 ubiquitin-activating enzyme and then displaced to a cystein residue in the E1, forming a thioester linkage. The ubiquitin is then transferred from E1 to the E2 ubiquitin-conjugating enzyme through formation of another thioester linkage. Finally, the formation of the covalent linkage between ubiquitin and the target proteins is established by the E3 ubiquitin ligase. Iteration of this process leads to ubiquitin chain elongation, and the formation of tetra-ubiquitin chains, which is the minimal-chain length required for proteasomal degradation (Vembar and Brodsky, 2008). However, in some cases, ERAD is promoted by the action of E4 ubiquitin-chain-extension enzymes (Kohlmann...
et al., 2008; Nakatsukasa et al., 2008; Richly et al., 2005). The importance of chain elongation during ERAD might derive from the fact that the polyubiquitin appendage must reach a crucial length before a substrate can be retrotranslocated (Jarosch et al., 2002).

2.5.2 Ubiquitylation enzymes

In *S. cerevisiae*, two E3 ligases have been identified in the ER membrane: Hrd1p and Doa10p. Both are components of large multisubunit transmembrane complexes (Figure 4). Hrd1p and Doa10p cooperate with different E2 ubiquitin-conjugating enzymes. Hrd1p uses Ubc7p (Bays et al., 2001), or alternatively also Ubc1p (Hampton, 2002). Doa10p uses Ubc7p or Ubc6p (Ravid et al., 2006). Ubc6p is an integral ER-membrane protein, whereas Ubc7p is membrane associated through Cue1p (Biederer et al., 1997). Other E3s that participate in ERAD are cytoplasmic but are recruited to the ER membrane for ubiquitylation of particular proteins, like for example Rsp5p in yeast and parkin in mammals (Haynes et al., 2002; Kostova and Wolf, 2003; Tanaka et al., 2004).

Mammalian cells use at least four E3 ubiquitin ligases CHIP, RMA1, GP78, and HRD1/synoviolin for ERAD. HRD1/Hrd1p is implicated in the turnover of several substrates including the Pae1 receptor, misfolded insulin, and the unassembled T-cell receptor subunits TCR-α and CD3δ (Allen et al., 2004; Kikkert et al., 2004; Omura et al., 2006). The function of TEB4, the Doa10p homolog, however remains to be established in mammalian ERAD (Kreft et al., 2006). Moreover, the mammalian E3 ubiquitin ligases Fbs1 and Fbs2, belonging to the SCF ubiquitin ligase family, have a sugar-binding domain and it has been shown that Fbs1 binds strongly to the chitobiose moiety of denatured glycoproteins, while Fbs2 also recognizes the mannosyl-residues of the oligosaccharide (Yoshida et al., 2003). Though, the substrates of the SCF\textsuperscript{Fbs1} complex have not yet been characterized, it may clear the cytosol of glycoproteins that have escaped the classical ERAD pathways or leaked from the ER through damage in the lipid layer (Hirsch et al., 2009).
2.6 Cytoplasmic extraction

After ubiquitylation, the proteins are extracted from the membrane, shuttled to the proteasome, deglycosylated and degraded.

2.6.1 Cdc48 complex

Although some cases have been reported where the proteasome was sufficient to extract substrates (Lee et al., 2004b; Mayer et al., 1998), in general, the proteins are retrotranslocated by means of the Cdc48 complex, which provides the mechanical energy for the extraction of the substrates from the ER to the cytosol (Jentsch and Rumpf, 2007). In *S. cerevisiae* it consists of Cdc48p, which is a hexameric AAA⁺ ATPase, and two associated factors, Ufd1p and Npl4p. Correspondingly, in mammals the Cdc48 homolog p97 is associated with the conserved UFD1 and NPL4 homologs.

It is unknown how Cdc48p gains access to the ERAD substrates, which appears especially challenging in the case of ER luminal proteins. However, the polyubiquitin moiety has been postulated to function as a handle for the complex to initiate the extraction (Flierman et al., 2003). In any case, the Cdc48 complex might be recruited to the ER membrane (at least in part) through its interaction with Ubx2p in *S. cerevisiae* (Neuber et al., 2005; Schuberth and Buchberger, 2005), or VIMP in mammals (Ye et al., 2004). Recent evidence suggests that the extent to which a membrane protein is embedded in the lipid bilayer might dictate the degree to which the Cdc48 complex is required (Carlson et al., 2006).

2.7 Proteasomal degradation

2.7.1 Proteasomal targeting

After the retrotranslocation, the substrates might be directly transported from the membrane-associated Cdc48 complex to the proteasome, as Cdc48 was found in association with the proteasomal cap (Verma et al., 2000). However, a series of ubiquitin-binding proteins escort the ubiquitylated substrates from the ER membrane to the proteasome (Nakatsukasa and Brodsky, 2008). These may include UBA and
UBL domain-containing proteins, which interact both with the proteasome and with ubiquitylated substrates in the Cdc48 complex. Rad23 and Dsk2 are such factors. However, is unknown, whether they are static members of the Cdc48 complex, or if they bind to it, and deliver from there the ERAD substrates to the proteasome, acting as mobile escorts (Vembar and Brodsky, 2008).

2.7.2 Deglycosylation and deubiquitylation

In the case of an ERAD glycoprotein, a unique additional step is the removal of residual sugars by the cytoplasmic peptide:N-glycanase (PNGase) (Hirsch et al., 2003; Suzuki et al., 2002), which probably occurs before degradation. In fact, fewer oligosaccharides are released from cytosolic glycoproteins when proteasomes are inhibited (Karaivanova and Spiro, 2000), suggesting that retrotranslocation and deglycosylation are intimately coupled to proteasomal degradation (Lederkremer and Glickman, 2005). Png1p associates with Cdc48p and Rad23p (Kim et al., 2006; Li et al., 2006). Like the N-glycans also the ubiquitin chains are removed prior to degradation. For this, ubiquitin-binding proteins which are associated with the proteasome or are integral proteasome subunits, catalyze the en bloc removal of the polyubiquitin moieties, or alternatively catalyze the trimming of the polyubiquitin chain (Amerik and Hochstrasser, 2004). The substrate is then delivered into the central pore of the proteasome. The central cavity harbors trypsin-like, chymotrypsin-like and post-glutamylpeptide hydrolazing enzymes, the proteases that will cleave the substrate into short peptides (Vembar and Brodsky, 2008).
Structural requirements of Nss membranes. I. Presence of logical...


Chapter 1


Chapter 1


Chapter 1


Trombetta, E.S., J.F. Simons, and Helenius. 1996. Endoplasmic reticulum glucosidase II is composed of a catalytic subunit, conserved from yeast to mammals, and a tightly bound non-catalytic HDEL-containing subunit. JBC, in press.


Chapter 1


Chapter 2

Screening for Htm1p interacting proteins
Introduction

In nearly all events that take place in a cell protein-protein interactions occur and most of the cellular processes are regulated by multiprotein complexes. Indeed, many of the ER-resident proteins are organized in distinct complexes and a nascent polypeptide chain encounters them as soon as it starts to enter the ER. First interactions occur with the translocon complex and soon after with the signal peptidase and the oligosaccharyltransferase complex (Evans et al., 1986; Johnson and van Waes, 1999; Silberstein et al., 1992). Specific interactions between these complexes mediate their close association, facilitating their coordinated and immediate action on the nascent polypeptide chains emerging into the ER (Chavan and Lennarz, 2006). In fact, signal peptide cleavage and STT3A OTase isoform mediated N-glycosylation occur mainly co-translocationally in higher eukaryotes (Hebert and Molinari, 2007; Ruiz-Canada et al., 2009). Chaperones and folding enzymes of the ER are estimated to be present in nearly millimolar concentrations and they also assemble to form complexes (Helenius and Aeby, 2004). However, their composition is variable and depends on the properties or requirements of the nascent proteins (Kuznetsov et al., 1997; Meunier et al., 2002; Tatu and Helenius, 1997). Thus, the folding machinery is dynamic and characterized by many highly transient interactions between different folding enzymes and chaperones, and only relatively small complexes (Snapp et al., 2006; Tatu and Helenius, 1997). While protein complexes and transient assemblies of the folding machinery support the synthesis, maturation, and productive folding of nascent chains, protein complexes are also involved in the recognition of misfolded proteins and their targeting to degradation. In S. cerevisiae two distinct multiprotein complexes have been identified by tandem affinity purification approaches (Carvalho et al., 2006; Denic et al., 2006; Gauss et al., 2006). These complexes mediate and coordinate the substrate recognition, extraction, and ubiquitylation of ERAD substrates. ERAD-C substrates (with lesions in the cytosolic domain) are degraded via the Doa10 complex. In this, the Doa10p E3 ligase assembles with the Ubx2p, Cdc48 complex (Cdc48p-Npl4p-Ufd1p) and with the Ubc7p via Cue1p. In contrast, ERAD substrates with lesions in their luminal or transmembrane regions are handled by the Hrd1p complex acting in the ERAD-L and ERAD-M pathway. Here, Hrd1p, like Doa10p, forms a complex with
the Ubx2p, Cdc48 complex, and with the Ubc7p, Cue1p. In addition, Usa1p, Der1p, and Hrd3p, Yos9p, Kar2p are bound (Carvalho et al., 2006; Denic et al., 2006; Gauss et al., 2006). The Hrd3p/Yos9p/Kar2p is believed to specifically recognize and target misfolded proteins carrying a specific N-glycan structure to the degradation.

Traditionally, for the investigation of the composition of protein complexes, biochemical methods such as co-immunoprecipitation, cross-linking and copurification by chromatography have been used. However, these techniques require harsh treatment for cell disruption and therefore weak and/or transient interactions may not be preserved (Thaminy et al., 2003). The characterization of protein-protein interactions and further identification of transient associations between well defined complexes will prove insightful (Collins et al., 2007). Alternatively to the traditional biochemical methods, the yeast two-hybrid system proved to be a powerful method for the investigation and discovery of protein-protein interactions in vivo (Fields and Song, 1989). It takes advantage of the observation that many eukaryotic transcription factors can be divided into two functionally distinct domains mediating first DNA binding and second transcriptional activation (Auerbach et al., 2002). In the classical two-hybrid system, two hybrid proteins are involved. One is termed bait, which is a hybrid of the protein of interest X and the DNA-binding domain. The other is the prey and is a hybrid of the interacting protein Y, fused to the activation domain of the transcription factor. Co-expression of the bait and prey protein in yeast, leads to the combination of the two domains, in the case of interaction between the proteins X and Y. This results in the reconstitution of the functional transcription factor, readily measured by the activity of reporter genes. Reporter genes commonly comprise loci that complement mutations in biosynthetic genes of the reporter strain and the lacZ gene, encoding the bacterial lactose catabolism enzyme β-galactosidase. The loci that complement mutations in biosynthetic genes allow positive growth selection of those cells that express interacting bait and prey proteins. Moreover, the β-galactosidase can be used in a colorimetric assay. In the classical yeast two-hybrid system, it is a prerequisite that the interaction of the proteins can occur in the nucleus of the cell. Thus not every protein-protein interaction in a cell will be accessible with this system. Difficulties can be encountered with proteins that are destined for the secretory pathway, and especially those which rely on post-translational modifications such as N-glycosylation or disulfide bond formation for the stability in their cognate protein complex or transient interaction. Another challenge offers the
detection of the interaction of membrane proteins. Due to the hydrophobic nature of the transmembrane domains, the nuclear targeting is hindered (Auerbach et al., 2002). To address this group of proteins, the split-ubiquitin membrane yeast two-hybrid system was developed (Stagljar et al., 1998). The basic concept of this system is that ubiquitin proteases (localizing to the cytosol or nucleus) hydrolyze specifically linkages between ubiquitin and C-terminally fused proteins (Figure 1A). Ubiquitin can be expressed independently as an N- and a C-terminal half (named Nub(l) and Cub) that spontaneously assemble to yield the ‘split-ubiquitin’ (Johnsson and Varshavsky, 1994). This quasi native ubiquitin is also readily recognized by the ubiquitin proteases. However, a point mutation in the Nub domain, exchanging isoleucine 13 by glycine (yielding NubG), abolishes the spontaneous assembly of the two halves (Figure 1A). Additional interactions are now required to bring the two halves into the required spatial proximity. This interaction can be provided by two proteins that are fused each to the NubG or the Cub domain. The established interaction of NubG with Cub is measured by the expression of reporter genes (here HIS3 and lacZ). Expression of these genes is induced upon cleavage of the artificial transcription factor, comprising the bacterial LexA-DNA binding domain and the Herpes simplex VP16 transactivator protein (Figure 1B) (Auerbach et al., 2002).

The putative ER-degradation lectin Htm1p, had been shown to act in the ERAD-L pathway (Vashist and Ng, 2004). Though, it was not detected as a component of the Hrd1p complex and proposed to act upstream of Yos9p (Denic et al., 2006; Gauss et al., 2006). In order to provide some further insight on the functional context of Htm1p, we set out to identify in vivo interacting partners of Htm1p, using the split-ubiquitin two hybrid system.
Figure 1. Detection of protein-protein interactions with the split-ubiquitin system.

**A)** Basic concept of the split ubiquitin. Ubiquitin proteases (UBP's) hydrolyze specifically linkages between ubiquitin and C-terminally fused proteins. Ubiquitin can be expressed independently as an N- and a C-terminal half (NubI and Cub). They can spontaneously assemble to yield the ‘split-ubiquitin’, which is again readily detected by UBP's. In contrast, exchanging isoleucine 13 by glycine (yielding NubG), abolishes the spontaneous assembly of the two halves.

**B)** The principle of the split-ubiquitin yeast two hybrid system. A prey protein is fused to NubG, while the protein of interest is fused to the Cub domain. If the proteins interact, they will provide the spatial proximity and the additional affinity to assemble the NubG-Cub split-ubiquitin. This is in turn recognized and processed by cytoplasmically localized UBPs. This results in the release of the reporter, comprising the LexA-DNA binding domain and the Herpes simplex VP16 transactivator protein, from the Cub domain. The reporter diffuses into the nucleus where transcription of reporter genes (here HIS3 and lacZ) is induced. Modified from (Auerbach et al., 2002).
Results

Design and functionality of the LV-Cub-Htm1 bait
Previously performed fractionation experiments in our group revealed the major portion of Htm1p in the membrane fraction, indicating that it is a membrane protein (Bielmann, 2003), (personal communication C. Jakob). EDEM1 was described as a type II ER transmembrane protein in COS cells (Hosokawa et al., 2001), and computational prediction suggested a similar topology for Htm1p. Therefore, the LexA-VP16-Cub (LV-Cub) reporter was fused to the N terminus of Htm1p.
To control the functionality of the Htm1-bait, complementation of the ∆htm1 ERAD phenotype was assessed with the ERAD model substrate CPY* (Finger et al., 1993; Wolf and Fink, 1975). Therefore, CPY* degradation rates of ∆htm1 cells transformed with the LV-Cub-Htm1 construct or empty vector were compared to degradation rates in wild type cells. As previously reported, degradation of CPY* in ∆htm1 is impaired. The LV-Cub-Htm1 complemented the CPY* degradation deficiency in ∆htm1 cells (Figure 2A). This indicated that the hybrid protein was functional, and that it had correctly reached its site of action.

Evaluation of the LV-Cub-Htm1 bait quality and topology
To evaluate the quality of the hybrid protein as bait in the split-ubiquitin assay, the self-activation of the construct was determined. For this, the L40 reporter strain was either transformed with the LV-Cub-Htm1 bait construct, the LV-Cub reporter alone, or empty vector, respectively. The self-activation was measured via the growth reporter HIS3 by streaking of the cells on synthetic minimal medium lacking histidine. Whereas the LV-Cub reporter control led to growth on histidine lacking medium, none of the four tested LV-Cub-Htm1 transformands grew (Figure 2B). This indicated, on one hand, that the bait did not self-activate; and on the other hand, corroborated that the bait had been targeted to the membrane, resulting in the tethering of the otherwise soluble LV-Cub reporter to the membrane.
For confirmation of the topology of the bait, the reporter strain was co-transformed with the LV-Cub-Htm1 bait, and prey constructs, NubG-Alg5, Nubl-Alg5, Ost1-NubG, and Ost1-Nubl, of two known ER localizing proteins Alg5p and Ost1p. Nubl is the wild type N-terminal portion of ubiquitin that spontaneously assembles with Cub, the
C-terminal portion of ubiquitin (Figure 1A), independently of further interactions between the hybrid proteins. Thus, Nubl-fusions with proteins of known cellular localization can be used as compartmental marker and indicate the topology of bait membrane proteins. The protein-protein interactions were analyzed by subsequent assessment of the expression of both reporter proteins: first, growth of the transformed reporter strain on synthetic medium lacking histidine; second, determination of the β-galactosidase activity of the grown yeast. As a positive control for the readout, the L40 reporter strain was co-transformed with the LV-Cub reporter and empty vector instead of bait and a prey. Activation of both reporter genes was readily detectable with the control transformants. Both, the Alg5-Nubl and the Ost1-Nubl hybrids interacted with the bait, and activation of the reporter genes appeared similarly to the control strain. This showed that the bait had reached its native compartment and inserted according to the predictions into the membrane. In contrast, the NubG hybrids of Alg5p and Ost1p did not interact (Figure 2C). This indicated that in vivo neither Alg5 nor Ost1 protein interact with Htm1p.
A) Functional analysis. CPY* cycloheximide chase degradation analysis over 90 min, of W303 (HTM1) or YG1582 (Δhtm1) transformed with LV-Cub-Htm1 or empty vector. Depicted is the immunoblot analysis probed with α-HA and subsequently with α-HXK.

B) Selfactivation analysis. Growth assay of L40 reporter strain transformed with either LV-Cub reporter R, LV-Cub-Htm1 bait construct (1-4) or no plasmid (−). The plasmids have LEU2 resistance marker. Growth on SD-His indicates activation of the reporter mediated HIS3 expression.

C) Topology analysis. NubI-NubG test. L40 co-transformed with LV-Cub-Htm1 with either Alg5-Nubl (Alg5-Nubl), or with Alg5-NubG (Alg5-NubG), or with Ost1-Nubl or Ost1-NubG. L40 co-transformed with LV-Cub and pRS414 (LV-Cub+pRS414); or co-transformed with empty LEU plasmid and empty TRP plasmid (pRS414). LV-Cub and LV-Cub-Htm1 plasmids carry a LEU marker, Nub-plasmids have a TRP marker.

Figure 2. Evaluation of the LV-Cub-Htm1-bait.
**Manual mini-screen with selected membrane-proteins**

A direct interaction analysis was performed with proteins that are involved in the degradation of misfolded proteins. These included members specifically involved in the ERAD-L/M pathway such as the E3 ubiquitin ligase Hrd1p and Der1p; a member specific for ERAD-C, Ubc6p; and members of both, Ubc7p and Cue1p. Moreover, proteins involved in protein translocation: Sec62p, Sss1p, Sbh1p; or in ER-Golgi transport Sed5p, Sec22p were used as prey (Scheper et al., 2003; Wittke et al., 1999). The L40 reporter strain was co-transformed with the NubG-constructs of the corresponding proteins, and with the LV-Cub-Htm1-bait. Analysis of the reporter gene activation did not reveal interactions with the tested proteins, belonging to the Hrd1p or Doa10p complex. Interactions were detected with Sbh1p and Sss1p, the β- and the γ-subunits of the Sec61p ER translocation complex. However, Sss1p did not interact when it was expressed from a single copy plasmid, but it interacted when it was expressed from a multi copy plasmid (Table 1).

**Table 1. Detected interactions in the manual mini screen.**

<table>
<thead>
<tr>
<th>Functional group</th>
<th>Prey-construct</th>
<th>Copy-number</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ubiquitylation</td>
<td>Hrd1-NubG</td>
<td>multi</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Der1-NubG</td>
<td>multi</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>NubG-Ubc6</td>
<td>single</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Ubc7-NubG</td>
<td>multi</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Cue1-NubG</td>
<td>multi</td>
<td>no</td>
</tr>
<tr>
<td>ER-Golgi transport</td>
<td>NubG-Sed5</td>
<td>single</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>NubG-Sec22</td>
<td>single</td>
<td>no</td>
</tr>
<tr>
<td>Translocation</td>
<td>NubG-Sec62</td>
<td>single</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>NubG-Sss1</td>
<td>single</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>NubG-Sss1</td>
<td>multi</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>NubG-Sbh1</td>
<td>single</td>
<td>yes</td>
</tr>
<tr>
<td>N-glycosylation</td>
<td>NubG-Alg5</td>
<td>single</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Ost1-NubG</td>
<td>multi</td>
<td>no</td>
</tr>
</tbody>
</table>

**Robotized screening of a global membrane protein prey library**

In order to screen for interacting proteins of Htm1p on a larger scale, the LV-Cub-Htm1 bait was subjected to an interaction screening with an ORF-HA$_2$-NubG- and a NubG-HA$_2$-ORF prey-library, comprising 705 proteins (Miller et al., 2005). The library was constructed on the basis of the annotation of the proteins as localizing to an...
‘integral membrane environment’ according to the YPD web page (Costanzo et al., 2001); and, further 63 proteins with similarity to these proteins were included. This high-throughput assay constitutes an ‘array approach’. In this, the bait and the prey-protein-library are introduced into isogenic reporter strains of opposite mating type. With the help of a robot, the bait-plasmid containing L40 strain was mated with the array of AMR70 transformants, containing each an assigned prey-plasmid. The protein-protein interactions in the diploid yeast were detected by growth on selective medium, lacking histidine as well as β-galactosidase activity. All detected 38 interactions are listed in Table 2 with a brief functional description of the targeted proteins.

Previous analysis of large-scale interaction data sets revealed significant numbers of false negative and false positive interactions (von Mering et al., 2002). In order to evaluate those hits with higher probability to be of physiologic relevance, the interacting proteins were sorted according to their cellular localization, based on the GO annotation of the ‘Saccharomyces Genome Database’ SGD project http://www.yeastgenome.org/. 29 of the totally 38 interactions were annotated ER residents, indicating that an enrichment of ER-proteins was achieved in the screen. A graphical representation of the detected interactions is shown in Figure 3.
Figure 3. Interactome detected in the robotic screen.

Htm1p is in the center of two concentric circles. The outer circle (black dots) depicts proteins that were annotated 'not resident in the ER'. The inner circle (colored dots), indicates proteins assigned to the ER compartment, or unknown localization. The color code used for the ER resident candidates specifies their GO molecular function if known; or are grey, if unknown. *localization unknown. **localization to the secretory pathway. The assignments were done according to gene ontology (GO) annotation used by BioGRID.
Besides the compartmental categorization of the hits, a semi-quantitative measure of the interaction was performed. For this, the interactions were measured at the moderately elevated temperatures of 32°C and 34°C. Those interactions that persisted up to 34°C might represent stronger and more stable interactions. The temperature increase led to a gradual reduction of the number of interactors (Figure 4 and Figure 5).

Figure 4. Example of a read out from the robotic screen.

L40 harboring the LV-Cub-Htm1 bait was crossed with the array of AMR70 array, harboring the NubG-prey-library. Each prey is twice present: once expressed from a single copy and once from a multicopy plasmid (localized next to each other in the array). Diploid selection was performed on SD-L-W medium at 30°C. Diploids were replica pinned onto selection plates (SD-L-W-H supplemented with 3mM 3-AT), and incubated at 30°C, 32°C or 34°C. Pictures were taken after 2 days (left panel). Subsequently, the β-galactosidase activity was determined in a qualitative, colorimetric assay with X-Gal substrate. Blue dots indicate yeast that produced β-galactosidase. Clones that grew in duplicate (interaction of the bait with the prey expressed from single- and multicopy plasmid); and turned both blue, were considered as interactors. Blue colored filter-papers are depicted in the right panel.
Furthermore, the detected interactions were compared with the interactome that was identified with this array by (Miller et al., 2005). In this large-scale screen 1,985 putative interactions involving 536 membrane proteins were identified. The interactions from this large-scale screen are shown and compared with the hits identified in the herein described screen in Figure 5.

**Figure 5. Comparison of the interactions with former interactions, detected with this array.**

Visualization of the interactions detected with the large-scale array approach by Miller et al. In the center (yellow dots) are the proteins grouped that did not interact with the Htm1p bait. The two concentric circles around the yellow pool depict proteins that interacted with Htm1p. The inner circle harbors the preys that interacted with one or more other proteins in the large-scale screening of Miller et al. (corresponding to the grey lines). The outer circle shows Htm1p-interactors that were not identified to interact with other baits in the large-scale screen. Black dots represent the interactions that were only detectable at the standard growth condition of 30°C. Orange depicts the interactions that were detectable at 32°C and in red those resisting 34°C.
Discussion

Analysis of previous large-scale interaction data sets revealed significant numbers of false negatives and false positives (von Mering et al., 2002). There is no doubt, that the same may be true for the herein described two-hybrid screen. Nevertheless, there may be relevant interactions among them. Possibly, they represent rather transient than stable associations, similar to the case of chaperones and folding enzymes (Kuznetsov et al., 1997; Meunier et al., 2002; Tatu and Helenius, 1997). In any case, it is evident that protein-protein interaction data obtained from two-hybrid systems require methodologically independent validation prior to final conclusions. Yet, some findings deserve comment as they may open room for some speculations and hypotheses for future research.

No interaction between Htm1p and the Hrd1p or the Doa10p complex
The herein described interaction analysis revealed no interaction of Htm1p with proteins that belong to the multiprotein complexes, the Hrd1p complex and the Doa10 complex. Hrd1p, Der1p, Ubc6, Ubc7p and Cue1p did not interact in the manual screen. Of these, Hrd1p, Ubc6p and Cue1p were also present in the array, and did as well not interact in the robotized screening. Furthermore, Doa10p which was included in the array did not interact. This is in line with the finding that Htm1p is not a component of the Hrd1p or the Doa10p complex, as was concluded from tandem affinity purification experiments (Carvalho et al., 2006; Denic et al., 2006; Gauss et al., 2006). Moreover, it suggests that also no transient interactions with these complexes occur.

Indications for Htm1p-translocon β subunit interaction
Htm1p was found to interact with the β subunits of both heterotrimeric translocon complexes, the Sec61 (Sec61p, Sbh1p, Sss1p) and the Ssh1 (Ssh1p, Sbh2, Sss1p). The interactions were readily detected in the robot screening, and the interaction with Sbh1p was also detected in the manual screen. Of note, Sbh1p from the yeast Yarrowia lipolytica was found to interact with calnexin. It was suggested that the role of Sec61β in the quality control could consist of maintaining the chaperone calnexin in the vicinity of the translocation pore. Sbh1p would thus function as a docking
protein to couple translocation and the quality control process (Boisrame et al., 2002). In accordance with this model, mammalian Secβ was found to interact with BAP31, which is a protein-sorting factor of newly synthesized membrane proteins. Moreover, microsomes from sbh1 sbh2 yeast are export competent for pαF in vitro, suggesting rather no involvement of the β subunits in retrotranslocation (Romisch, 2005). Noteworthy, a direct physical interaction of calnexin and EDEM was reported (Oda et al., 2003). However, here no interaction was detected between Htm1p and Cne1p. It should be recalled that all proteins pass the translocon during their synthesis. Protein-protein interactions detected with subunits of the translocon could also represent transient interactions during translocon passage instead of functional cooperativity.

**Interaction of Htm1p with Yet1p**

The function of Yet1p has not been characterized precisely in *S. cerevisiae*. Though, *YET1* was identified as a high-copy suppressor in a high-copy suppressor screen for oligosaccharyltransferase assembly mutations. In this screen also *UBC6*, encoding a ubiquitin conjugating enzyme that functions in ERAD was identified (Spirig, 1999). Similarly, *UBC6* overexpression also resulted in the suppression of a temperature sensitive sec61 mutation (Sommer and Jentsch, 1993). The mammalian homolog BAP31 has been suggested to function as an endoplasmic reticulum protein-sorting factor that may contribute to the delivery of client proteins from their site of synthesis to specific ER complexes that mediate subsequent export, retention, degradation or survival of the client protein (Wang et al., 2008; Wang et al., 2004). BAP31 was found in association with the Sec61β and translocation associated membrane protein TRAM (Görlich and Rappoport, 1993); as well as in physical and functional interaction with Derlin-1 (Wang et al., 2008).

**Indications for Htm1p interaction with peptidyl-prolyl cis-trans isomerases**

Of interest may also be the interactions that were detected with the peptidyl-prolyl cis-trans isomerases (PPIase) Fpr2p and Cpr4p. They appeared to be in a relatively stable interaction with Htm1p, as Cpr4p resisted 32°C and Fpr2p could yet be detected at 34°C. Interestingly, these proteins have so far not been reported in many other protein-protein interactions, what might be an indication that the interaction with
Htm1p is relatively specific. Cpr4p was found in interaction with Swp1p (Tarassov et al., 2008) and Fpr2p was found in interactions with three proteins, of which only Mpd2p (a protein disulfide isomerase) appears to be a potentially true positive hit, as it is the only protein that also localizes to the ER (Krogan et al., 2006; Tachikawa et al., 1997). Cpr4p belongs to the Cyp-family, and Fpr2p to the FKBP-family. PPIases increase protein folding rates in vitro, however their precise function in the maturation of polypeptides in vivo is poorly understood (Hebert and Molinari, 2007). However, Fpr2p and Cpr4p are both induced upon accumulation of unfolded proteins in the ER, what suggests that they may play a role in protein folding in vivo (Dolinski et al., 1997; Franco et al., 1991; Nielsen et al., 1992; Oliver et al., 1992; Partaledis and Berlin, 1993). FPR2 and CPR4 are not essential. Even a S. cerevisiae dodecuplet mutant lacking 12 identified members of the Cyp- and FKBP-family (including Δfpr2 and Δcpr4) is viable (Dolinski et al., 1997). It was suggested that immunophilins may perform specific functions through interactions with unique sets of restricted partner proteins that remain to be identified (Dolinski et al., 1997).

Comparison of the Htm1p-hits from this and earlier studies
General comparison of the hits obtained in this study with so far published interaction partners reveal no overlapping Htm1p-hits. The published interactors that were determined by different groups and methods (Gavin et al., 2002; Krogan et al., 2006; Uetz et al., 2000), did as well not reveal overlap among each other. This may be due to the different methodological set up, like varying expression strength or differences in selection or purification stringency. However, most of them appear to be false positives as they do not localize to the secretory pathway, let alone to the ER. However, most interestingly, the ER-protein Pdi1p was one of the Htm1p interactors. It had been identified in a high throughput proteome-wide purification of protein complexes (Krogan et al., 2006). Importantly, making use of a novel purification enrichment scoring system, Collins et al. merged the data sets from two independent affinity purification reports (Gavin et al., 2006; Krogan et al., 2006), and generated a high confidence protein-protein interaction data set. The interaction of Htm1p and Pdi1p was considered to be of high confidence (Collins et al., 2007).

As initially discussed, no more than speculations can be made on these results at this stage. Yet, alternative confirmation of the protein-protein interactions combined with direct functional investigations may reveal novel and relevant aspects in protein
folding, quality control and ERAD in the future.
Table 2: Htm1 protein-protein interactions detected at 30°C

<table>
<thead>
<tr>
<th>Gene</th>
<th>Cellular component</th>
<th>Description SGD</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMP3</td>
<td>Plasma Membrane</td>
<td>Small plasma membrane protein related to a family of plant polypeptides that are overexpressed under high salt concentration or low temperature, not essential for viability, deletion causes hyperpolarization of the plasma membrane potential</td>
</tr>
<tr>
<td>ERV14</td>
<td>ER</td>
<td>Protein localized to COPII-coated vesicles, involved in vesicle formation and incorporation of specific secretory cargo; required for the delivery of bud-site selection protein Axl2p to cell surface; related to Drosophila cornichon</td>
</tr>
<tr>
<td>STE14</td>
<td>ER</td>
<td>Farnesyl cysteine-carboxyl methyltransferase, mediates the carboxyl methylation step during C-terminal CAAX motif processing of a-factor and RAS proteins in the endoplasmic reticulum, localizes to the ER membrane</td>
</tr>
<tr>
<td>SBH1</td>
<td>ER</td>
<td>Beta subunit of the Sec61p ER translocation complex (Sec61p-Sss1p-Sbh1p); involved in protein translocation into the endoplasmic reticulum; interacts with the exocyst complex and also with Rtn1p; homologous to Sbh2p</td>
</tr>
<tr>
<td>EMP24</td>
<td>ER, COPII</td>
<td>Integral membrane component of endoplasmic reticulum-derived COPII-coated vesicles, which function in ER to Golgi transport</td>
</tr>
<tr>
<td>VPH2</td>
<td>ER</td>
<td>Integral membrane protein required for vacuolar H+ -ATPase (V-ATPase) function, although not an actual component of the V-ATPase complex; functions in the assembly of the V-ATPase; localized to the endoplasmic reticulum</td>
</tr>
<tr>
<td>YET1</td>
<td>ER</td>
<td>Endoplasmic reticulum transmembrane protein; may interact with ribosomes, based on co-purification experiments; homolog of human BAP31 protein</td>
</tr>
<tr>
<td>ERG2</td>
<td>ER</td>
<td>C-8 sterol isomerase, catalyzes the isomerization of the delta-8 double bond to the delta-7 position at an intermediate step in ergosterol biosynthesis</td>
</tr>
<tr>
<td>CPR4</td>
<td>Membrane</td>
<td>Peptidyl-prolyl cis-trans isomerase (cyclophilin), catalyzes the cis-trans isomerization of peptide bonds N-terminal to proline residues; has a potential role in the secretory pathway</td>
</tr>
<tr>
<td>SBH2</td>
<td>ER</td>
<td>Ssh1p-Sss1p-Sbh2p complex component, involved in protein translocation into the endoplasmic reticulum; homologous to Sbh1p</td>
</tr>
<tr>
<td>ZSP1</td>
<td>ER</td>
<td>Protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the ER; YBR287W is not an essential gene</td>
</tr>
<tr>
<td>PER1</td>
<td>Vacuole</td>
<td>Protein of the endoplasmic reticulum, required for GPI-phospholipase A2 activity that remodels the GPI anchor as a prerequisite for association of GPI-anchored proteins with lipid rafts; functionally complemented by human ortholog PERLD1</td>
</tr>
<tr>
<td>AVT6</td>
<td>Vacuole</td>
<td>Vacuolar amino acid transporter, exports aspartate and glutamate from the vacuole; member of a family of seven S. cerevisiae genes (AVT1-7) related to vesicular GABA-glycine transporters</td>
</tr>
<tr>
<td>ERG24</td>
<td>ER</td>
<td>C-14 sterol reductase, acts in ergosterol biosynthesis; mutants accumulate the abnormal sterol ignosterol (ergosta-8,14 dienol), and are viable under anaerobic growth conditions but inviable on rich medium under aerobic conditions</td>
</tr>
<tr>
<td>OST3</td>
<td>ER</td>
<td>Gamma subunit of the oligosaccharyltransferase complex of the ER lumen, which catalyzes asparagine-linked glycosylation of newly synthesized proteins; Ost3p is important for N-glycosylation of a subset of proteins</td>
</tr>
<tr>
<td>SSH1</td>
<td>ER</td>
<td>Subunit of the Ssh1 translocon complex; Sec61p homolog involved in co-translational pathway of protein translocation; not essential</td>
</tr>
</tbody>
</table>
VBA2  Vacuole  Permease of basic amino acids in the vacuolar membrane
GPI16  ER  Transmembrane protein subunit of the glycosylphosphatidylinositol transamidase complex that adds GPs to newly synthesized proteins; human PIG-Tp homolog
STT3  ER  Subunit of the oligosaccharyltransferase complex of the ER lumen, which catalyzes asparagine-linked glycosylation of newly synthesized proteins; forms a subcomplex with Ost3p and Ost4p and is directly involved in catalysis
FPR2  ER  Membrane-bound peptidyl-prolyl cis-trans isomerase (PPlase), binds to the drugs FK506 and rapamycin; expression pattern suggests possible involvement in ER protein trafficking
SPC1  ER  Subunit of the signal peptidase complex (SPC), which cleaves the signal sequence from proteins targeted to the endoplasmic reticulum (ER), homolog of the SPC12 subunit of mammalian signal peptidase complex
SHR3  ER  Endoplasmic reticulum packaging chaperone, required for incorporation of amino acid permeases into COPII coated vesicles for transport to the cell surface
SEC11  ER  18kDa catalytic subunit of the Signal Peptidase Complex (SPC; Spc1p, Spc2p, Spc3p, and Sec11p) which cleaves the signal sequence of proteins targeted to the endoplasmic reticulum
ERP4  ER  Protein with similarity to Emp24p and Env25p, member of the p24 family involved in ER to Golgi transport
PRM9  ER  Pheromone-regulated protein with 3 predicted transmembrane segments and an FF sequence, a motif involved in COPII binding; member of DUP240 gene family
SWP1  ER  Delta subunit of the oligosaccharyl transferase glycoprotein complex, which is required for N-linked glycosylation of proteins in the endoplasmic reticulum
CYT1  Mitochondrion  Cytochrome c1, component of the mitochondrial respiratory chain; expression is regulated by the heme-activated, glucose-repressed Hap2p/3p/4p/5p CCAAT-binding complex
GPI2  ER  Protein involved in the synthesis of N-acetylglucosaminyl phosphatidylinositol (GlcNAc-PI), the first intermediate in the synthesis of glycosylphosphatidylinositol (GPI) anchors; homologous to the human PIG-C protein
CSG2  ER  Endoplasmic reticulum membrane protein, required for mannosylation of inositolphosphorylceramide and for growth at high calcium concentrations
RIM2  Mitochondrion  Mitochondrial pyrimidine nucleotide transporter; imports pyrimidine nucleoside triphosphates and exports pyrimidine nucleoside monophosphates; member of the mitochondrial carrier family
KTR2  Golgi  Mannosyltransferase involved in N-linked protein glycosylation; member of the KRE2/MNT1 mannosyltransferase family
OST2  ER  Epsilon subunit of the oligosaccharyltransferase complex of the ER lumen, which catalyzes asparagine-linked glycosylation of newly synthesized proteins
YEA4  ER  Uridine diphosphate-N-acetylglucosamine (UDP-GlcNAc) transporter required for cell wall chitin synthesis; localized to the ER
CSH1  Vacuole  Probable catalytic subunit of a mannosylinositol phosphophorylceramide (MIPC) synthase, forms a complex with probable regulatory subunit Csg2p; function in sphingolipid biosynthesis is overlapping with that of Sur1p
ALG7  ER  Dol-P in the ER in the first step of the dolichol pathway of protein asparagine-linked glycosylation; inhibited by tunicamycin
LAC1 ER  Ceramide synthase component, involved in synthesis of ceramide from C26(acyl)-coenzyme A and dihydrosphingosine or phytosphingosine, functionally equivalent to Lag1p

HXT12  Possible pseudogene in strain S288C; YIL170W/HXT12 and the adjacent ORF, YIL171W, together encode a non-functional member of the hexose transporter family

SGE1 Plasma membrane  Plasma membrane multidrug transporter of the major facilitator superfamily, acts as an extrusion permease; partial multicopy suppressor of gal11 mutations
Material and Methods

Yeast strains and plasmids

Standard yeast media and genetic techniques were used (Guthrie and Fink, 1991). Yeast strains used in this study were: W303 MATa, leu2-3, 112, his3-11, trp1-1, ura3-1, can1-100, ade2-1 (gift D.T. Ng); haploid descendant of W303, YG1526 MATa, leu2-3, 112, his3-11, trp1-1, ura3-1, can1-100, ade2-1 Δhtm1. Reporter strains L40 (AMR69) MATa his3Δ200 trp1-901 leu2-3,112 ade2 lys2-801am URA3::(lexAop)8-lacZ; and AMR70 MATα his3Δ200 lys2-801am trp1-901 leu2-3,112 URA3::(lexAop)8-lacZ (Hollenberg et al., 1995). ORF-HA2-NubG/NubG-HA2-ORF prey-library (kindly provided by I. Stagljar (Miller et al., 2005)).

Plasmids used in this study were: pDN431-CPY*-HA (669); pRS314-Alg5-Nubl-(285), pRS314-Alg5-NubG (296), pAS-Ost1-Nubl (400) and pAS-Ost1-NubG (402) (Stagljar et al., 1998); pAS2-Hrd1-NUG (1034), pAS2-Der1-NUG (1036), pAS2-Ubc7-NUG (1037), pAS2-Cue1-NUG (1035) and pAS2-NUG-Sss1 (1038) (gift from K. Römisch; (Scheper et al., 2003)); pRS314-CUP-NUG-Sed5 (1028), pRS314-CUP-NUG-Sec22 (1033), pRS314-CUP-NUG-Sec62 (1032), pRS314-CUP-NUG-Sss1(1029), pRS314-CUP-NUG-Sbh1 (1030) and pRS314-CUP-NUG-Ubc6 (1031) (gift N. Johnsson; (Dünnwald et al., 1999; Wittke et al., 2000)); pRS414 (556); pRS315 (40); pMB2-LV-Cub (gift from I. Stagljar); LV-Cub-Htm1 bait was generated by homologous recombination, of HTM1 into the pMB2-LV-Cub vector, designed for the LV-Cub fusion for N-terminal tagging under TEF promoter and CYC1 terminator control. The HTM1 locus was amplified by PCR from pHTM1-7 (No. 622) with the forward primer LV-Htm1 (with the sequence: 5'-tatgcacagatcagctttgtcgacggtatcgataagcttggtttgctgcttatgggtgc-3'), thus omitting the atg start codon of HTM1; and reverse primer i-Htm1-cyc-rec-MB2 (sequence: 5'-atgtgggggggagggctgaatgtaagcgtgac-3') (bold indicate the regions homologous to the pMB2 vector). For the homologous recombination the yeast strain W303 was co-transformed with the PCR fragment and the PstI-Eco47III (Fermentas) linearized pMB2. Transformants were selected on synthetic minimal medium lacking leucine (SD-Leu) and the recombinated plasmid was isolated from the yeast cells and confirmed by sequencing analysis.
Analysis of CPY* degradation

The cycloheximide chase was performed as described previously (Jakob et al., 1998). In brief, yeast cells were grown at 30°C in appropriate medium to midlog phase, corresponding to an OD_{600nm} of 0.8-1.2. 3 x 10^8 cells were harvested and resuspended in medium. The chase was initiated by addition of cycloheximide (final concentration 100 µg/ml) and performed at 30°C. 1 x 10^8 cells were removed at each time point and transferred into NaN_3 (final concentration 0.1% [w/v]) on ice, immediately pelleted and flash-frozen in liquid nitrogen. Whole cell protein extracts were prepared using glass beads, 1% (w/v) SDS, 50mM Tris-HCL (pH 7.5), 2mM PMSF. Proteins were subjected to reducing 7% SDS-PAGE and electroblotted to nitrocellulose membranes. CPY*-HA was immunologically detected with anti-HA antibody (rabbit) at a dilution of 1:1000 (Santa Cruz) and goat anti-rabbit IgG-horseradish peroxidase at 1:3000 (Santa Cruz). Visualization was performed with ECL detection (GE Healthcare lifesciences). Reprobing of membranes with anti-hexokinase (HXK) antibody (rabbit) was done accordingly.

Robotized screening

The screening was performed similarly to the array screening methods earlier described (Miller et al., 2005; Uetz et al., 2000). A single colony of L40, transformed with the bait was grown overnight at 30°C in selective medium, and transferred into an Omnitray plate. A Hamilton robot with a 384-needle pinning tool was used to transfer the cell suspension onto SD-Leu plates. In parallel the array of the prey library was pinned onto YPD plates. The plates were incubated for 2 days at 30°C. The grown bait colonies were replica-pinned onto the array, grown on YPD plates. Mating was performed over night at 30°C, and spots were replica-pinned onto SD-Leu-Trp plates for diploid selection and grown for 2 days at 30°C. For determination of the interactions, the diploids were pinned onto SD-Leu-Trp-His plates supplemented with 3mM 3-amino-1,2,4-triazole and incubated for 4 days at 30°C. For determination of growth at elevated temperatures, the plates were replica pinned and incubated at 30°C, 32°C or 34°C for 4 days. Growth was visually analyzed and subsequently assayed by β-galactosidase activity, and interactions scored.
β-galactosidase assay

The filter lift assay has been described earlier (Mockli and Auerbach, 2004). In brief, cells were grown on SD-selection plates for 2 to 3 days. A piece of Whatman filter paper was directly placed onto the agar plate and incubated for 10 min to allow the yeast to stick to the filter. The transferred cells were lysed in a freeze-thaw cycle. For this, the filter was transferred with forceps to liquid N₂ for 2 min, and again into a Petri or Omnitray dish, with yeast side up where it was thawn for 5 min. The filter was overlaid with freshly prepared X-Gal agarose (1x PBS pH 7.4, 0.5% [w/v] agarose, 0.1% X-Gal [5-Bromo-4-chloro-3-indolyl-β-D-galactoside]). The plates were incubated at room temperature until agarose was polymerized and blue color developed.
References


Chapter 2


Chapter 3

Htm1 protein generates the N-glycan signal for glycoprotein degradation in the endoplasmic reticulum

Simone Clerc, Christian Hirsch, Daniela Maria Oggier, Paola Deprez, Claude Jakob, Thomas Sommer, and Markus Aebi

J. Cell Biol. 184:159-172

Contribution:
Figures: 1; 2; 3; 4; 5A, 5D, 5E; and 7
with Markus Aebi: writing of the manuscript
with Paola Deprez: development of the N-glycan analysis method
Htm1 protein generates the N-glycan signal for glycoprotein degradation in the endoplasmic reticulum

Simone Clerc¹, Christian Hirsch², Daniela Maria Oggier¹,³, Paola Deprez¹, Claude Jakob¹, Thomas Sommer² and Markus Aebi¹

¹Institute of Microbiology, Department of Biology, ETH Zürich, Wolfgang-Pauli Strasse 10, CH-8093 Zürich, Switzerland; ²Max-Delbrück-Center for Molecular Medicine Robert-Rössle-Strasse 10, D-13122 Berlin, Germany; ³present address: University of Applied Sciences Northwestern Switzerland, School of Life Sciences, Institute of Ecopreneurship, Gründenstrasse 40, CH-4132 Muttenz, Switzerland

Corresponding author:
Markus Aebi, Institute of Microbiology, ETH Zürich, Wolfgang-Pauli Strasse 10, HCI F407, CH-8093 Zürich

e-mail: aebi@micro.biol.ethz.ch
phone: ++41-44-632 64 13
fax: ++41-44-632 13 75

Running title: Htm1 protein is an α1,2-specific exo-mannosidase
Key words: ERAD, quality control, mannosidase, protein disulfide isomerase
Abstract

To maintain protein homeostasis in secretory compartments, eukaryotic cells harbor a quality control system that monitors protein folding and protein complex assembly in the endoplasmic reticulum (ER). Proteins that do not fold properly or integrate into cognate complexes are degraded by ER-associated degradation (ERAD), involving retrotranslocation to the cytoplasm and proteasomal peptide hydrolysis. N-linked glycans are essential in glycoprotein ERAD; the covalent oligosaccharide structure is used as a signal to display the folding status of the host protein. In this study, we define the function of the Htm1 protein as an $\alpha_{1,2}$-specific exomannosidase that generates the Man$_7$GlcNAc$_2$ oligosaccharide with a terminal $\alpha_{1,6}$-linked mannosyl residue on degradation substrates. This oligosaccharide signal is decoded by the ER-localized lectin Yos9p that in conjunction with Hrd3p triggers the ubiquitin-proteasome-dependent hydrolysis of these glycoproteins. The Htm1p exomannosidase activity requires processing of the N-glycan by glucosidase I, II and mannosidase I, resulting in a sequential order of specific N-glycan structures that reflect the folding status of the glycoprotein.
Introduction

In eukaryotic cells, the synthesis of asparagine-linked glycoproteins occurs in the ER. As nascent protein chains enter the ER lumen through the Sec61 translocon complex, the oligosaccharyltransferase transfers the \( \text{Glc}_3\text{Man}_9\text{GlcNAc}_2 \) glycan to asparagines in the consensus motif N-X-S/T, generating an N-linked glycan. Soon after the transfer, the N-linked glycan is processed by glucosidase I, glucosidase II, and ER mannosidase I. N-glycans increase the hydrophilicity of the protein and thereby affect protein folding directly. Importantly, these glycans also serve as ligands for proteins that modulate the folding (Helenius and Aebi, 2004). The \( \text{Glc}_3\text{Man}_9\text{GlcNAc}_2 \) structure is a signal for the interaction with the molecular chaperones calnexin or calreticulin in higher eukaryotes or Cne1p as proposed in \textit{Saccharomyces cerevisiae} (Xu et al., 2004). In mammalian cells, after the removal of the mannose-linked glucose by glucosidase II, the folding sensor UDP-Glucose:glycoprotein glucosyltransferase can readd a terminal glucose and thereby allows for another cycle of reassociation with calnexin/calreticulin. In contrast, in \textit{S. cerevisiae}, no reglucosylation of misfolded proteins has been found, which implies the absence of the calnexin/calreticulin cycle in this system (Fernandez et al., 1994).

Proteins that achieve the correct fold and integrate correctly into their cognate complexes can continue along the secretory pathway to their site of action. In contrast, proteins that do not achieve the proper fold in due time, as well as excess subunits of protein complexes, are retained in the ER and ultimately are degraded by the cytoplasmically located 26S proteasome. The process of recognition, retro-translocation and degradation of terminally misfolded proteins is referred to as ER-associated degradation (ERAD).

Although the N-glycans play a crucial role in the folding of N-glycoproteins, their remodeling by glucosidase I, glucosidase II and ER mannosidase I is also crucial for N-glycoprotein ERAD (Hitt and Wolf, 2004; Jakob et al., 1998a; Knop et al., 1996). The finding that ER mannosidase I is a relatively slow-acting enzyme (compared with glucosidase I and II), supported the mannose timer hypothesis (Helenius et al., 1997; Su et al., 1993) and suggested that the ER mannosidase I product glycan \( \text{Man}_8\text{GlcNAc}_2 \), when present on misfolded proteins, serves as the ERAD targeting
signal in mammalian cells and in *S. cerevisiae*. This led to the proposal that the ER contains a lectin or a lectinlike protein that recognizes Man$_8$GlcNAc$_2$ glycans on misfolded glycoproteins and by this initiates glycoprotein ERAD (Jakob et al., 1998a; Liu et al., 1997; Yang et al., 1998).

A group of proteins found throughout the eukaryotic kingdom with homology to class I mannosidases (family 47 glycosylhydrolases), termed Htm1p/ER degradation-enhancing $\alpha$-mannosidase-like protein (EDEM), was proposed to act as such degradation lectins in ERAD. Sequence comparison of ER mannosidase I, EDEM1/2/3, and Htm1p reveals a mannosidase homology region with significant sequence identity among the homologues. However, the proteins differ in their remaining domains. One particular characteristic is the presence of a C-terminal extension which varies among the EDEM/Htm1p homologues (Kanehara et al., 2007; Mast et al., 2005). In Htm1p, this domain constitutes almost one third of the whole protein. Although the essential residues of the processing mannosidases are conserved in Htm1p/EDEM1/2/3, initially neither *in vivo* (Jakob et al., 2001) nor *in vitro* (Hosokawa et al., 2001; Mast et al., 2005) mannosidase activity was reported, supporting their potential role as lectins. More recently, another protein besides Htm1p, Yos9p, has been suggested to act as ER-degradation lectin. Yos9p harbors a mannose 6-phosphate receptor homology domain essential for its function in N-glycoprotein ERAD. It was shown to be required for the efficient degradation of ERAD substrates that carry lesions in the ER lumen (Bhamidipati et al., 2005; Buschhorn et al., 2004; Kim et al., 2005; Szathmary et al., 2005).

In *S. cerevisiae*, three pathways, ERAD-L, ERAD-C and ERAD-M (L, C and M stand for Lumen, Cytosol and Membrane), have been described, engaging different factors for the degradation of proteins with misfolded lesions in the corresponding domains (Bonifacino et al., 1990; Taxis et al., 2003; Vashist and Ng, 2004). ERAD-L engages the Hrd1p E3 ubiquitin ligase, which is in a complex with the Ubc7p/Cue1p E2 ubiquitin-conjugating enzyme and the AAA-ATPase Cdc48p with the cofactors Npl4p, Ufd1p, and Ubx2p. Moreover, this complex includes Hrd3p, Yos9p, Kar2p, Der1p and Usa1p (Carvalho et al., 2006; Denic et al., 2006; Gardner et al., 2000; Gauss et al., 2006a; Gauss et al., 2006b). The Hrd3p/Kar2p/Yos9p is believed to specifically recognize and target misfolded proteins carrying Man$_8$GlcNAc$_2$ glycans of the ERAD-L pathway to the degradation. Interestingly, Htm1p, is also acting in the ERAD-L pathway *via* the Hrd1p complex (Vashist and Ng, 2004); however, it was not
found to be a member of the complex but proposed to act upstream of Yos9p (Denic et al., 2006; Gauss et al., 2006a; Kanehara et al., 2007).

Although in *S. cerevisiae* the Man$_8$GlcNAc$_2$ glycan as N-glycan ERAD determinant remained, it became evident that in mammalian cells, N-glycans of folding-deficient proteins are further trimmed to Man$_6$GlcNAc$_2$ and Man$_5$GlcNAc$_2$ structures (Ermonval et al., 2001; Foulquier et al., 2004; Foulquier et al., 2002; Frenkel et al., 2003; Hosokawa et al., 2003; Kitzmuller et al., 2003). By this extensive demannosylation, the terminal mannose of the A-branch, which is the acceptor mannose of UDP-Glucose:glycoprotein glucosyltransferase, is removed. Thus, the proteins can no longer associate with the chaperones calnexin and calreticulin and can be directly deviated into the disposal machinery (Molinari, 2007). Recently, EDEM3 was shown to have mannosidase activity *in vivo*, as its overexpression leads to mannose trimming *in vivo* to Man$_7$GlcNAc$_2$ and Man$_6$GlcNAc$_2$ (Hirao et al., 2006), and interestingly, EDEM1 was also proposed to act as processing mannosidase in the extraction of misfolded glycoproteins from the calnexin/calreticulin cycle rather than being a lectin (Olivari et al., 2006).

Reconsidering glycoprotein quality control and ERAD in *S. cerevisiae* in which the calnexin/calreticulin cycle is absent, no apparent need for enhanced N-glycan processing has emerged to date. Nevertheless, the essential residues of the processing mannosidases are conserved in Htm1p. Therefore, we reevaluated the function of Htm1p in the yeast ERAD system.
Results

HTM1 overexpression leads to altered N-glycan processing
To readdress the question whether Htm1p is an active mannosidase in vivo, the N-glycan structures of overall proteins from wild-type, Δhtm1, and wild-type cells overexpressing HTM1 were compared. N-linked sugars were metabolically labeled in vivo for 20 minutes with $^3$[H]-Mannose, and whole cell protein extracts were prepared. N-glycans were released by peptide N-glycosidase F treatment and analyzed by HPLC. As previously reported, the most abundant N-glycan structure in the ER of wild-type cells was Man$_5$GlcNAc$_2$ (Byrd et al., 1982; Jakob et al., 1998a). In addition, small amounts of oligosaccharides migrating at the positions of Man$_5$GlcNAc$_2$ and Man$_7$GlcNAc$_2$ glycans were also detected (Fig. 1 A). The conversion of the N-linked glycans to Man$_5$GlcNAc$_2$ was not impaired in Δhtm1 cells, and the N-glycan profile appeared similar both in wild-type and Δhtm1 cells (Jakob et al., 2001) (Fig. 1 A). However, when HTM1 was overexpressed we were able to detect an increase of a putative Man$_7$GlcNAc$_2$ glycan, indicating Htm1p induced N-glycan processing (Fig. 1, A and B).

We also tested the effect of HTM1 deletion and overexpression on the degradation of the ERAD model substrate carboxypeptidase Y* (CPY*) (Finger et al., 1993; Wolf and Fink, 1975). As previously reported, HTM1 deletion reduced the degradation rate of CPY*, but overexpression of the same protein had no effect on the stability of this ERAD substrate (Fig. 1 C), suggesting that Htm1p function is not the rate-limiting step in the degradation process.

HTM1 overexpression generates a Man$_7$GlcNAc$_2$ glycan
Next, we analyzed the kinetics of N-linked glycan processing by pulse-chase experiments. In wild-type cells, a 10-min pulse predominantly revealed the Man$_5$GlcNAc$_2$ structure. Small amounts of a Man$_9$GlcNAc$_2$ oligosaccharide were also present. Upon the chase, oligosaccharides with higher molecular weight became prominent, reflecting the processing of N-linked glycans by Golgi-localized mannosyltransferases (Byrd et al., 1982). In contrast, in HTM1-overexpressing cells the Man$_7$GlcNAc$_2$ was detected directly after the pulse, and the relative amount of
this oligosaccharide increased during the chase period (Fig. 2 A). In comparison with the empty vector control, the amount of the Man₇GlcNAc₂ glycan was 4-fold or 4.5-fold increased at 0 min or 30 min of chase, respectively. Concomitantly, the Man₈GlcNAc₂ was substantially reduced, indicating that it had served as physiological substrate for the trimming (Fig. 2 B). We concluded that overexpression of Htm1p results in an amplified level of the Man₇GlcNAc₂ N-linked glycan. This suggested a novel exomannosidase function of this protein in vivo.

_Htm1p mannosidase acts on Man₈GlcNAc₂ yielding Man₇GlcNAc₂ isomer C_

The aforementioned results indicated that the Man₇GlcNAc₂ oligosaccharide was generated by trimming of the Man₈GlcNAc₂ oligosaccharide. Therefore, we first tested whether the Man₈GlcNAc₂ oligosaccharide structure was required for the Htm1p-dependent processing. For this, we took advantage of processing-deficient mutant strains and of the relaxed substrate specificity of the yeast oligosaccharyltransferase to genetically tailor the structure of the glycans present on proteins (Jakob et al., 1998b). In wild-type cells, the activities of glucosidase I and II as well as ER mannosidase I result in the generation of the Man₈GlcNAc₂ structure (Kornfeld and Kornfeld, 1985; Moremen et al., 1994), and overexpression of Htm1p yielded the Man₇GlcNAc₂ glycan. The Glc₂Man₉GlcNAc₂ oligosaccharide in a glucosidase II-deficient strain is still processed by ER mannosidase I to the Glc₂Man₈GlcNAc₂ oligosaccharide (Jakob et al., 1998a), but _HTM1_-dependent processing was strongly impaired (Fig. 3, A and B). In contrast, the Glc₁Man₉GlcNAc₂ oligosaccharide transferred in Δalg8Δgls2 cells was processed by ER mannosidase I to the Glc₁Man₈GlcNAc₂ oligosaccharide, and this glycan was as well target for Htm1p-dependent processing to Glc₁Man₇GlcNAc₂. The deletion of the _MNS1_ locus resulted in the generation of Man₉GlcNAc₂ glycans and the absence of ER mannosidase I activity almost completely abolished Htm1p-dependent processing. From this result, we concluded that the sequential trimming of the N-glycan by glucosidase I, II, and ER mannosidase I, generating the Man₈GlcNAc₂ oligosaccharide in vivo, is a prerequisite for Htm1p-dependent processing. Moreover, because extension of the A-branch by one glucose did not prevent _HTM1_-dependent processing, this analysis suggested that the trimming involved either the B or C branch of the N-linked glycan target.

To elucidate whether the B or C branch was processed by Htm1p, we isolated the
radioactively labeled Man₇GlcNAc₂ product from HTM1-overexpressing cells and wild-type (vector control) cells by preparative HPLC. The glycans were digested in vitro with α-1,2-exomannosidase from Trichoderma reesei (Maras et al., 2000), and the products were reassessed by HPLC analysis to determine whether two or three α-1,2-exomannosidase-sensitive linkages were present yet (Fig. 4 A). The chromatograms showed a 3.5-fold increase of Man₅GlcNAc₂ over the Man₄GlcNAc₂ glycan in cells that had overexpressed HTM1 in comparison with the vector control (Fig. 4, B and C). From this, we deduced that the Htm1p mannosidase acted on an α-1,2-linked mannose residue of the Man₈GlcNAc₂ oligosaccharide. As the presence of a glucosyl residue on the A-branch of the oligosaccharide still allowed Htm1p-dependent processing, we concluded that overexpression of Htm1p results in the selective hydrolysis of the α-1,2-linked mannose of the C-branch.

Htm1p is an active mannosidase in vivo, and the essential residues are required for CPY* degradation
To determine whether the observed N-glycan processing activity was accomplished directly by Htm1p, mutant forms of Htm1p in which the conserved amino acid residues E222 and D279 were replaced by glutamine or asparagine, respectively, were overexpressed in yeast. The orthologous residues of E222 and D279 in ER mannosidase I are essential for the hydrolase function (Lipari and Herscovics, 1999). To assess the stability of the mutants, the overexpressed proteins were marked with a TPH7 tag (Knop et al., 1999) at the C terminus. The mutations did not affect the stability of the proteins, as was confirmed by immunoblot analysis of the TPH7-tagged proteins (Fig. 5 A). However, the N-glycan pulse-chase analysis with the overexpressed Htm1-E222Q-TPH7 as well as Htm1-D279N-TPH7 revealed that both mutations abolished the N-glycan processing (Fig. 5, B and C). The same result was obtained with the untagged mutants Htm1-E222Q and Htm1-D279N (unpublished data). This indicated that the mannosidase activity is a direct function of Htm1p.

To determine whether Htm1p ERAD function relies on its mannosidase activity, we assessed the functionality of the mutant Htm1 proteins in the ERAD process. Therefore, CPY* degradation rates of Δhtm1 cells transformed with HTM1-TPH7, Htm1-E222Q-TPH7, Htm1-D279N-TPH7, or empty vector were analyzed. The Htm1-TPH7 construct complemented the CPY* degradation deficiency in Δhtm1 cells. In
contrast, the point mutants Htm1-E222Q-TPH7 and Htm1-D279N-TPH7 did not rescue this phenotype (Fig. 5, D and E). The same result was obtained with the untagged point mutants (unpublished data). Therefore, the acidic residues E222 and D279 are essential for Htm1p ERAD function as well as for the observed mannosidase activity.

As noted in the Introduction, both Mns1p and Htm1p contain a mannosidase homology region, but Htm1p is characterized by a C-terminal extension that accounts for almost one third of the whole protein. We generated a truncated version of the HTM1 locus encoding the mannosidase domain only (Htm1-Δ517-TPH7). Immunoblot analysis confirmed that Htm1-Δ517-TPH7 was expressed (Fig. 5 A) but did not result in N-glycan processing (Fig. 5, B and C) and did not complement the Δhtm1 CPY* degradation phenotype (Fig. 5, D and E). Therefore, the mannosidase homology region was necessary but not sufficient for Htm1p function. We propose that Htm1p functions as an α-1,2-mannosidase on the C-branch of the N-linked glycan. This mannosidase activity is essential for efficient degradation of the CPY* ERAD substrate and requires the C-terminal domain of the protein.

*Htm1p physically interacts with Pdi1p via its C-terminal domain*

To address the function of the C-terminal domain more closely, Htm1p-interacting proteins were identified. We generated a functional C-terminally Myc-tandem affinity purification (TAP)-tagged Htm1p construct (unpublished data). Microsomes were prepared, and solubilized Htm1-Myc-TAP protein was precipitated from the supernatant with IgG beads. The purified complexes were resolved by SDS-PAGE, and proteins were stained (Fig. 6 A). The two prominent bands that were visualized around 62 kDa were isolated from the gel and subjected to mass spectrometry. Both bands were identified as protein disulfide isomerase 1 protein (Pdi1p; Table S1). We wondered whether the C-terminus of Htm1p might be required for this interaction. Therefore, Myc-tagged constructs of Htm1p were generated of different truncates within the C terminus, namely at the positions 749, 683, 647, 582, and 547. We also generated Pdi1-HA-HDEL and a C-terminal truncate of Pdi1p at the position 459, Pdi1-Δ459-HA-HDEL. Immunopurification of Pdi1-HA-HDEL or Pdi1-Δ459-HA-HDEL was performed, and the presence or absence of full-length or truncated Htm1p-Myc was determined by immunoblot analysis. The results confirmed the interaction of the
full-length Pdi1p with Htm1p. However, no interaction could be detected between Pdi1-Δ459-HA-HDEL and Htm1-Myc (Fig. 6 B). When we pulled down full-length Pdi1p-HA-HDEL and assessed the binding of the Htm1p truncates, the interaction was readily detectable with the Htm1-Δ749-, -Δ683-, and -Δ647-Myc-truncates, but not with Htm1-Δ582 and -Δ547 (Fig. 6 C). From this, we deduced that the C-terminus is involved in the interaction with Pdi1p.

A terminal α1,6-linked mannose as a glycan determinant directing misfolded glycoproteins to degradation

Having identified the main product of the Htm1p activity as a specific Man₇GlcNAc₂ isomer, we were wondering whether this glycan, rather than the postulated Man₈GlcNAc₂ oligosaccharide, might function as the putative glycoprotein degradation signal in *S. cerevisiae*. Generation of this isomer on a misfolded protein might target its host protein for degradation. In contrast to Man₈GlcNAc₂, the Man₇GlcNAc₂ C structure is characterized by a terminal α-1,6-linked mannose of the C-branch (Fig. 7 A). Within the framework of our hypothesis, we assumed specifically that a terminal α-1,6-linked mannosyl residue functioned as a signal required for degradation. Yos9p, the putative ER-degradation lectin with a mannose 6-phosphate receptor homology domain essential for glycoprotein ERAD, might act as the receptor for this structure (Bhamidipati et al., 2005; Buschhorn et al., 2004; Kim et al., 2005; Szathmary et al., 2005). To address this hypothesis directly, we took advantage of the possibility to alter N-glycan structures by manipulating the biosynthetic pathway of lipid-linked oligosaccharide assembly (Jakob et al., 1998b). It is noteworthy that the putative product of Htm1p, the Man₇GlcNAc₂ oligosaccharide is not a biosynthetic intermediate in the pathway of N-glycoprotein biosynthesis. However, Δalg3 mutant cells produce a Man₅GlcNAc₂ glycan that also carries a terminal α1,6-linked mannose linked to an α1,3-substituted mannose, as is the case in the Man₇GlcNAc₂ C isomer. In contrast, Δalg9 and Δalg12 cells do not produce a terminal α1,6-linked mannose (Fig. 7 A). In support of our hypothesis, wild-type and Δalg3 cells degrade CPY* efficiently, whereas Δalg9 and Δalg12 cells show a strongly reduced degradation rate (Fig. 7, B and C; Jakob et al., 1998a). Moreover, the rapid degradation in wild-type and Δalg3 cells was dependent on the presence of the putative lectin Yos9p (Fig. 7, F and G), but Htm1p was only required in wild-type...
but not in $\Delta alg3$ cells (Fig. 7, D and E). This reinforced our hypothesis that Htm1p is essential in wild-type cells for the generation of a specific glycan signal, a terminal $\alpha 1,6$-linked mannose. This signal is always present in the $\Delta alg3$ cell, making degradation of malfolded proteins independent of Htm1p function but still dependent on the Yos9p lectin.
Discussion

Initial findings suggested that the Man\textsubscript{8}GlcNAc\textsubscript{2} B glycan may serve as glycan signal for degradation in \textit{S. cerevisiae} and mammalian cells. Htm1p/EDEM1 and more recently Yos9p were proposed to act as lectins or lectinlike proteins that specifically recognize misfolded glycoproteins carrying Man\textsubscript{8}GlcNAc\textsubscript{2} glycans and by this initiate glycoprotein ERAD. In this study, we analyzed the function of Htm1p in more detail.

\textit{Htm1p is an exomannosidase involved in ERAD}

We detected the removal of one mannose upon \textit{HTM1} overexpression, yielding the Man\textsubscript{7}GlcNAc\textsubscript{2} C glycan, but this did not increase the degradation rate of the model protein CPY*. This suggests that Htm1p is an $\alpha1,2$-exomannosidase acting on the C branch of the N-linked glycan of glycoproteins. Importantly, overexpression of mutant Htm1p did not result in an increase in the overall Man\textsubscript{7}GlcNAc\textsubscript{2} glycan level and these mutant forms were inactive in ERAD function. From this, we propose that the Man\textsubscript{7}GlcNAc\textsubscript{2} glycan generated by Htm1p on ERAD substrates serves as the N-glycan degradation signal, but Htm1p activity is not limiting in the degradation pathway. In this study, we addressed several aspects of this model: (a) what is the nature of the glycan signal; (b) how is the signal specifically generated on an ERAD substrate; and (c) what is the receptor of the glycan signal in the ERAD pathway?

\textit{A terminal $\alpha1,6$-linked mannosyl residue is the product of Htm1p activity}

The combined actions of ER mannosidase I and Htm1p result in a unique protein-bound high mannose structure not found during the biosynthesis of the lipid-linked oligosaccharide substrate for N-linked protein glycosylation (Burda and Aebi, 1999). ER mannosidase I produces a terminal $\alpha1,3$-linked mannose on the B-branch, whereas the action of Htm1p removes the capping $\alpha1,2$-mannose of the C-branch, resulting in a terminal $\alpha1,6$-linked mannose. Interestingly, a similar structure is present on the Man\textsubscript{5}GlcNAc\textsubscript{2} oligosaccharide found in $\Delta$alg3 mutant cells. Both contain a terminal $\alpha1,6$-linked mannose (Fig. 7 A). As reported earlier (Jakob et al., 1998a) and in this study, degradation of CPY* is independent of glycan processing in $\Delta$alg3 mutant cells, suggesting that the signal for degradation is present on the unprocessed $\Delta$alg3 glycan. Therefore, we postulate that Htm1p-dependent
processing unmasks the $\alpha_{1,6}$-linked mannose residue, and this structure is an essential functionality in the recognition of misfolded glycoproteins. It has been proposed that mammalian homologs of Htm1p (EDEM1 and EDEM3) also act on the $\alpha_{1,2}$-linked mannose residues of the A branch (Hirao et al., 2006; Olivari et al., 2006). Our data do not allow us to formally exclude the possibility that this is also the case for Htm1p. However, analysis of the free N-linked glycans, the putative products of ERAD present in the cytoplasm of yeast reveal that $\text{Man}_8\text{GlcNAc}_2$ and $\text{Man}_7\text{GlcNAc}_2$ are the predominant forms (Chantret et al., 2003; Suzuki and Funakoshi, 2006) (unpublished data).

**Substrate specificity of Htm1p**

Our model suggests that Htm1p has a dual substrate specificity; it requires a well-defined oligosaccharide structure present on an unfolded protein. Our data showed that the processing of the N-glycan by ER mannosidase I was essential for efficient generation of the $\alpha_{1,6}$-linked mannose by Htm1p (Fig. 3). We propose that substrate recognition by Htm1p involves a terminal $\alpha_{1,3}$-mannosyl residue generated by ER mannosidase I. This explains the observation that glycoprotein ERAD requires trimming by ER mannosidase I both in yeast and in mammalian cell culture systems (Jakob et al., 1998a; Liu et al., 1997; Yang et al., 1998).

Our analysis did not allow us to address the recognition of the substrate protein domains by Htm1p directly. Detailed work by Ng and others revealed that glycans localized on specific sites of ERAD substrates are required for degradation whereas others are dispensable (Kostova and Wolf, 2005; Spear and Ng, 2005) (D. Ng, personal communication). Therefore, only a minor portion of N-linked glycans might be processed by Htm1p. Our observation that overexpression of this protein was required to visualize small amounts of processing products is in line with this hypothesis. It is tempting to speculate that the C-terminal domain of Htm1p is required for the recognition of the protein part of the substrate. An appealing explanation could be that this part of the protein directly or indirectly binds misfolded peptides or hydrophobic patches and thus mediates substrate selection. In support of this proposal is our finding that the C terminus was required for the interaction with Pdi1p (Fig. 6 C). This interaction had previously been reported from a high throughput proteome-wide purification of protein complexes from *S. cerevisiae*. 
(Collins et al., 2007; Krogan et al., 2006). Interestingly, Pdi1p has been shown to recognize terminally misfolded secretory proteins and target them to the retrotranslocation (Gillece et al., 1999).

**Yos9p recognizes the α1,6-linked terminal mannose**

The specifically modified N-glycan as a signal for ERAD requires decoding by a lectin. Several groups have reported on the role of the putative lectin Yos9p in the process of ERAD (Bhamidipati et al., 2005; Buschhorn et al., 2004; Carvalho et al., 2006; Denic et al., 2006; Gauss et al., 2006a; Kim et al., 2005; Szathmary et al., 2005). Yos9p is part of a large complex that among other factors comprises the Hrd1p/Hrd3p E3-ligase. The interaction of a nonfolded protein with Hrd3p is believed to provide the first signal for degradation and Yos9p further inspects the glycans of the protein. If this glycan signals terminal misfolding, the dual recognition triggers the ubiquitination and proteasomal degradation of the substrate protein (Denic et al., 2006; Gauss et al., 2006a). Our genetic data support the hypothesis that Yos9p specifically binds to the terminal α1,6-linked mannosyl residue, generated by the Htm1p activity. Work by the group of J.S. Weissman corroborates this proposal by showing that the Yos9p lectin indeed prefers glycan structures that contain a terminal α1,6-linked mannose (Quan et al., in press). In addition, the in vivo immunoprecipitation experiments performed with CPY*-HA and Yos9-ProA-His7 protein revealed physical interaction of these proteins in wild-type cells where the signal is generated by Htm1p and in the Δalg3 mutant cells that expose the signal due to incomplete N-glycan biosynthesis (Szathmary et al., 2005).

**Comparison of Htm1p activity and EDEM function in mammalian cells**

Recent studies suggest that EDEM1 as well as EDEM3 act as enzymes rather than as lectins, and enhance the demannosylation of ERAD substrates in vivo (Hirao et al., 2006; Olivari et al., 2006). In view of the conservation of essential residues of the processing mannosidases, our finding that Htm1p has mannosidase activity in vivo is in accordance with the described activities of the mammalian orthologues.

Although direct structural analysis of processing products are missing, indirect evidence suggests that mammalian EDEM proteins have a dual function; they process the terminal mannose residues of the A-branch and thereby remove the
UDP-Glucose:glycoprotein glucosyltransferase acceptor, thus extracting the proteins from the calnexin/calreticulin cycle. However, N-glycans of misfolded glycoproteins are trimmed further down to Man$_5$GlcNAc$_2$ and Man$_5$GlcNAc$_2$ in mammalian cells. Besides the EDEM-family, other processing mannosidases of the glycosyl-hydrolase 47 family, like Golgi mannosidases and ER mannosidase I have been shown to contribute to the enhanced N-glycan processing of misfolded glycoproteins (Molinari, 2007; Moremen and Molinari, 2006).

We extend the model of glycan-signaling in ERAD of mammalian cells and propose that the extensive mannose trimming serves two distinct goals: the removal of the terminal $\alpha$$1,2$-linked mannose of the A-branch ensures the exit from the calnexin/calreticulin cycle, whereas the additional trimming of the C-branch generates the terminal misfolding signal and targets the N-glycoprotein for degradation. This dual activity explains the biphasic mode of glycoprotein degradation in mammalian cells: a lag phase where the protein remains in the calnexin/calreticulin cycle and is protected from degradation, followed by the actual degradation phase (Molinari, 2007). Mannose trimming of the A-branch by EDEM mannosidases releases the protein from the protective calnexin/calreticulin cycle and allows degradation. Indeed, overexpression of these enzymes was shown to result in an earlier onset of degradation (Molinari, 2007; Molinari et al., 2003). In view of our results, we propose that degradation itself requires the processing of the C-branch to generate the $\alpha$$1,6$-linked mannose as a degradation signal. It is possible that EDEM proteins are able to trim both the A- and the C-branch of the N-linked glycan. The absence of a calnexin/calreticulin cycle in S. cerevisiae made it possible to analyze the generation of this degradation signal directly.

It is a unique feature of the eukaryotic N-linked protein glycosylation that a complex oligosaccharide of a defined structure can be transferred to multiple sites, characterized by a short consensus sequence, N-X-S/T, of very diverse secretory proteins. Therefore, this posttranslational modification therefore tags many different proteins with a unique structure that is processed in the ER by a set of hydrolases and glycosyltransferases. They act in a sequential order and some of these enzymes modify only N-glycans of unfolded proteins. This orchestrated processing results in a set of defined glycan signals that can be interpreted by a general folding and degradation machinery (Helenius and Aebi, 2004). Our work reveals the identity of
one of these signals, the terminal $\alpha_{1,6}$-linked mannose residue generated by the exomannosidase activity of Htm1p. Our data suggest that this signal is interpreted by the lectin Yos9p, and we propose that this interaction communicates to the multicompontent degradation platform (Carvalho et al., 2006; Denic et al., 2006; Gauss et al., 2006a) that an unfolded, bound protein is destined for destruction. The complete, bipartite signal leads to the retrotranslocation, ubiquitination, and proteasomal degradation of the substrate protein. However, we do not know how Htm1p activity is regulated by the presence of an unfolded protein domain, and how the folding status of many different glycoproteins can modulate this mannosidase activity. A detailed biochemical analysis of the EDEM/Htm1 protein family will clarify this issue.
Materials and Methods

Yeast strains and plasmids
Yeast strains used in this study are listed in Table S2 (Biederer et al., 1997; Jakob et al., 2001; Jakob et al., 1998a; Vijayraghavan et al., 1989). Standard yeast media and genetic techniques were used (Guthrie and Fink, 1991). YG695, YSC8, YSC17, YSC24, YG1330 and YDO1 are haploid descendants of the wild-type strains SS328/SS330. YG1330 and YDO1 were generated using the pYM C-terminal epitope-tagging method (Knop et al., 1999). In brief, the TEV-ProA-7His tag together with the KanMX6 cassette was amplified by PCR. The PCR fragment was inserted into the genome of YG618 by homologous recombination at the 3’ end of the HTM1 locus, or at the 3’ end of the mannosidase homology region between Asn517 and Asn518, respectively. YTX140, YCH030, YCH036, YCH046, YCH133, YCH135, YCH137, YCH175, YCH176, YCH177, YCH178 are haploid descendants of DF5. PCR-based strategies were used to introduce the epitope tags to Htm1p and Pdi1p in these strains. All epitope-tagged proteins were expressed from their natural chromosomal locus under control of the endogenous promoter.

The YEp352-HTM1 (pHTM1u-1) plasmid was constructed by transferring the HTM1 gene from pRS316-HTM1 (pHTM1-7) to the high copy number vector YEp352 plasmid using XhoI-SacI restriction endonucleases. YEp352-HTM1-TPH7 (pDOG3) and YEp352-htm1Δ518-TPH7 (pDOG4) were generated by PCR amplification of the modified HTM1 loci from YG1330 and YDO1, and subcloning into YEp352. Mutations in full-length YEp352-HTM1-TPH7 were introduced by the QuikChange method (Stratagene) yielding YEp352-htm1-E222Q-TPH7 (pDOG5) and YEp352-htm1-D279N-TPH7 (pDOG6). Plasmids were confirmed by DNA sequence analysis.

Analysis of CPY* degradation
The cycloheximide chase was performed as described previously (Jakob et al., 1998a). In brief, yeast cells were grown at 30°C in appropriate medium to mid-log phase, corresponding to an OD_{600nm} of 0.8 - 1.2. 3 x 10^8 cells were harvested and resuspended in medium. The chase was initiated by addition of cycloheximide (final concentration 100 μg / ml) and performed at 30°C. 1 x 10^8 cells were removed at each time point and transferred into NaN₃ (final concentration 0.1% [wt/vol]) on ice,
immediately pelleted, and flash-frozen in liquid nitrogen. Whole cell protein extracts were prepared using glass beads, 1% (wt/vol) SDS, 50mM Tris-HCl, pH 7.5, 2mM PMSF. Proteins were subjected to reducing 7% SDS-PAGE and electroblotted to nitrocellulose membranes. CPY* was immunologically detected with anti-CPY antibody (rabbit) at a dilution of 1:1000 (Zufferey et al., 1995) and goat anti-rabbit IgG-horseradish peroxidase at 1:3000 (Santa Cruz Biotechnology, Inc.). Visualization was performed with ECL detection (GE Healthcare lifesciences). Reprobing of membranes with anti-hexokinase (HXK) antibody (rabbit) was performed accordingly. The x-ray films (super RX; Fujifilm) were scanned, and the CPY* protein amounts were determined densitometrically with a molecular imager (FX; Bio-Rad Laboratories) using the Quantity One program (Bio-Rad Laboratories).

\[ ^3H \]mannose labeling and analysis of N-linked glycans

Cells were grown at 30°C in appropriate medium to mid-log phase, corresponding to an OD\(_{600}\)nm of 0.8-1.2. \( n \times 5 \times 10^8 \) cells (\( n = \) number of analyzed time points) were harvested and washed in YP0.1%D (YPD containing 0.1% glucose). Labeling was performed in \( n \times 200 \mu l \) YP0.1%D containing \( n \times 100 \mu Ci \) D-2-\[^3H\]$-mannose (500GBq/mmol; Hartmann Analytic) and incubated for 10 (Figs. 2-5) or 20 min (Fig. 1) at 30°C. The radioactivity was chased by the addition of D-mannose to a 2% (wt/vol) final concentration at 30°C incubation. The chase was stopped by removing 5 \( \times 10^8 \) cells into TCA on ice (final concentration 10% [vol/vol]). TCA precipitates were washed twice with cold acetone and air dried for 20 min. The pellets were resuspended in buffer S (20 mM NaP, pH 7.5, 0.5% SDS, 40 mM DTT) and vortexted with glass beads at 50°C for one hour. Lysates were cleared by centrifugation, and iodoacetamide was added to 50-mM final concentration. Samples were incubated for 30 min at 37°C. Oligosaccharides were cleaved from proteins by digestion with peptide N-glycosidase F according to the recommendations of the manufacturer (BioConcept). Cleanup of oligosaccharides was modified from (Grubenmann et al., 2004). Prepacked C\(_{18}\) Sep Pak columns (Waters) were connected to columns (extended-volume empty reservoirs; Socochim SA) that were packed with Supelclean ENVI-Carb 120/400 (Sigma-Aldrich). The combined columns were equilibrated with methanol, acetonitrile, acetonitrile/H\(_2\)O (25:75; vol/vol), and acetonitrile/H\(_2\)O (2:98; vol/vol). Samples were loaded onto the columns after supplementation with
acetonitrile/H$_2$O (2:98; vol/vol). Columns were washed with acetonitrile/H$_2$O (2:98; vol/vol) and glycans eluted from the ENVI-Carb column with acetonitrile/H$_2$O (25:75; vol/vol). The solvent was evaporated in a speed vacuum.

For HPLC analysis of the oligosaccharides, a liquid chromatography (LC)-NH$_2$ column (250 x 4.6mm; Sigma-Aldrich) (Cacan et al., 1993) including an LC-NH$_2$ guard column was used. Oligosaccharide samples in acetonitrile/H$_2$O (70:30; v/v) were filtered through a 0.45-µm filter (Millipore) and injected on the equilibrated (acetonitrile/H$_2$O [70:30; vol/vol], 30 min) system using an autosampling device (Merck/Hitachi AS-2000). The gradient was acetonitrile/H$_2$O (70:30; vol/vol) to acetonitrile/H$_2$O (45:50; vol/vol) over 90 min, 5 min at acetonitrile/H$_2$O (45:50; vol/vol), returning to acetonitrile/H$_2$O (70:30; vol/vol) over 5 min and washing for 20 min at acetonitrile/H$_2$O (70:30; vol/vol) before injection of the following sample using a pump (Merck/Hitachi L-2600A; Sigma Aldrich). The eluate from the column was mixed continuously with scintillation fluid (FLO-Scint A; Packard) in a ratio of 1:1.5 (eluate:scintillation mix, vol/vol), and radioactivity was monitored with a flow monitor (FLO-ONE A-525, Packard) as described previously (Zufferey et al., 1995). Quantification of the peaks was done using PeakFit 4.06 program (SPSS) using the second derivative method.

Preparative purification of $[^3]$Hmannose-labeled N-linked glycans

$[^3]$Hmannose labeling was performed as described in the previous section with the following modifications: 4 x $10^9$ cells were labeled with 800 µCi D-2-$[^3]$Hmannose in 2 ml YP0.1%D. The labeling was stopped by addition of TCA, to the final concentration of 10% (vol/vol). The samples were worked up each in 2 batches and after the elution from the ENVI-Carb column were pooled again. An analytical HPLC run was performed with the N-glycans from 5 x $10^8$ cells. For isolation of individual N-glycans, the LC-NH$_2$ column was disconnected from the flow monitor and elution fractions of 500 µl were manually collected. The radioactivity in the fractions was determined using a β-counter (Tri-Carb 2800TR; PerkinElmer) with the QuantaSmart program (PerkinElmer). The counts were displayed in Excel (Microsoft), and the reconstituted profiles were compared with the analytical run for identification of the peaks.
Fractions constituting one peak were pooled, and the solvent was evaporated by speed vac. The oligosaccharides were resuspended in 50 mM sodium acetate buffer, pH 5, and digested with 1 μl of α-1,2 exo-mannosidase from T. reesei (provided by N. Callewaert, Ghent University, Ghent, Belgium) (Maras et al., 2000) or mock at 37°C overnight. The products were reassessed by HPLC analysis (see previous section).

**Determination of Htm1p-TPH7 protein expression**

Yeast cells were grown at 30°C in appropriate medium to mid-log phase, corresponding to an OD₆₀₀nm of 0.8-1.2. 4 x 10⁷ cells were harvested and lysed with glass beads in 1% (wt/vol) SDS, 50 mM Tris-HCl, pH 7.5, 2 mM PMSF, and 1 x PIC (protease inhibitor complete cocktail, Roche) by vortexing at 4°C for 10 min. Samples were heated at 95°C for 5 min. Supernatants were cleared by centrifugation and subjected to endoglycosidase H treatment according to the recommendations of the manufacturer (BioConcept). Proteins were subjected to reducing 7% SDS-PAGE and electroblotted to nitrocellulose membranes. TPH7-tagged proteins were immunologically detected with peroxidase-antiperoxidase (PAP) antibody at a dilution of 1:2000 (Sigma-Aldrich). Visualization and reprobing of membranes with anti-HXK antibody (rabbit) were performed as stated in the Analysis of CPY* degradation section.

**Purification of Htm1p**

1500 OD of yeast cells expressing C-terminally Myc-TAP-tagged Htm1p were washed with water + 1 mM PMSF and disrupted in 3 ml of IP32 (50 mM Hepes-NaOH, pH 7.2, 50 mM NaCl, 125 mM KOAc, 2 mM MgCl₂, 1 mM EDTA, and 3% glycerol) using glass beads. Upon addition of 45 ml IP32 + 1 mM PMSF, the lysate was centrifuged (1000 g for 8 min at 4°C). The resulting supernatant was centrifuged (32000 g for 1 h at 4°C), and the pellet lysed in 50 ml of IP32 + 0.25% NP-40. After clearance of the insoluble material by centrifugation (32000 g, 30 min at 4°C), Htm1p was precipitated from the supernatant by the addition of 50 μl IgG beads (GE
Healthcare) overnight. Beads were washed three times with IP32 + 0.25% NP-40. Bound proteins were liberated by the addition of TEV protease (Invitrogen) in a volume of 100 µl IP32 + 0.25% NP-40 followed by incubation for 3 h at 16°C. Samples were lyophylized and analyzed by SDS-PAGE followed by Coomassie staining and mass spectrometry.

**Mass spectrometry**

Proteins were separated on 10% polyacrylamide gels and stained with Coomassie brilliant blue R-250. Protein bands were excised and in-gel digested with sequencing-grade trypsin (Promega). The peptide mixture was separated on a reverse-phase column (75 µm PepMap C18; Dionex) connected to a capillary liquid chromatography system delivering a gradient of 5 to 50% acetonitrile. Eluting peptides were ionized by electrospray ionization on a mass spectrometer (Q-TOF1; Micromass). MS/MS analyses were conducted by using collision energy profiles chosen on the basis of the mass to charge ratio value and the charge state of the parent ion. The generated mass data were processed into peak lists containing mass to charge value, charge state of the parent ion, fragment ion masses, and intensities using the MassLynx 4.1 software (Micromass), and were correlated with protein databases using the Mascot 2.2 software (Matrix Science) (Perkins et al., 1999). Nonredundant protein databases NCBI (National Center for Biotechnology Information) and SwissProt (entry version 92) were searched without applying any constraints on M_r or species. Precursor ion accuracy was +/-0.1 D, and fragment ion accuracy +/-0.2 D. The results were manually validated.

**Co-immunoprecipitation**

To precipitate HA- and Myc-tagged proteins, logarithmically growing cells were harvested and washed with water + 1 mM PMSF. 50 OD cells were disrupted with glass beads in 400 µl IP32 (50 mM Hepes-NaOH, pH 7.2, 50 mM NaCl, 125 mM KOAc, 2 mM MgCl_2, 1 mM EDTA, and 3% glycerol) + 1 mM PMSF. After lysis, 1 ml
IP32 buffer was added, and a low speed centrifugation (1000 g, 5 min at 4°C) was performed. From the supernatant, a microsomal fraction was generated by high speed centrifugation (20000 g for 20 min at 4°C). After solubilisation in IP32 + 0.5% NP-40 and 1 mM PMSF, insoluble material was cleared from the lysate by centrifugation (20000 g for 15 min at 4°C). Tagged proteins were precipitated from the supernatant by addition of 1 μl of anti-HA (Sigma-Aldrich) or anti-Myc (Sigma-Aldrich) antibodies and 10 μl protein A-Sepharose beads (GE Healthcare) at 4°C overnight. Beads were washed three times with IP32 + 0.5% NP-40. Bound proteins were eluted by the addition of 100 μl SDS sample buffer and analyzed by Western blotting using the indicated antibodies.

**Online supplemental material**

Table S1 shows the peptide analysis of Htm1p-binding protein. Table S2 details the strains used in this study.

**Acknowledgement**

We thank the members of the Aebi lab for fruitful discussion and J.S. Weissman and D.T. Ng for the communication of data before publication. We are grateful to N. Callewaert for the enzyme gift. This work was supported by the Swiss National Science Foundation (grants to C. Jakob and M. Aebi) and the Eidgenössische Technische Hochschule Zurich.
Abbreviations list

CPY*, carboxypeptidase Y*
EDEM, ER degradation-enhancing α-mannosidase-like protein
ERAD, ER-associated degradation
HXK hexokinase
LC, liquid chromatography
TAP tandem affinity purification
## Supplemental material

### Supplemental tables

#### S1. Peptide analysis of Htm1p binding protein

<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Mass/charge ratio</th>
<th>Mr (expt)</th>
<th>Mr (calc)</th>
<th>Delta</th>
<th>Miss</th>
<th>Peptides identified</th>
<th>Ions score</th>
</tr>
</thead>
<tbody>
<tr>
<td>202</td>
<td>219</td>
<td>1006.50</td>
<td>2010.99</td>
<td>2011.00</td>
<td>-0.01</td>
<td>0</td>
<td>LSIYLPSAMDEPVYNGK (Ox. M)</td>
<td>59</td>
</tr>
<tr>
<td>367</td>
<td>384</td>
<td>1027.99</td>
<td>2053.97</td>
<td>2054.00</td>
<td>-0.03</td>
<td>0</td>
<td>SQEIFENQDSSVQLVGK</td>
<td>111</td>
</tr>
<tr>
<td>444</td>
<td>460</td>
<td>881.49</td>
<td>1760.97</td>
<td>1760.97</td>
<td>-0.01</td>
<td>0</td>
<td>GVVIEGYPTIVLPGGK</td>
<td>72</td>
</tr>
<tr>
<td>472</td>
<td>481</td>
<td>592.86</td>
<td>1183.71</td>
<td>1183.61</td>
<td>0.10</td>
<td>0</td>
<td>SLDSLDFIK</td>
<td>54</td>
</tr>
</tbody>
</table>

Summarized are the peptide sequences of Pdi1p identified by mass spectrometry. Start and end indicate the first and last amino acid residue of a detected peptide within the sequence of Pdi1p. Mass/charge ratio denotes the mass to charge ratio of the observed ions. Mr (expt) is the experimentally determined mass, and Mr (calc) the calculated mass of the peptide. Delta represents the difference between Mr (expt) and Mr (calc). Miss indicates the number of missed trypsin cleavage sites; and the column peptides identified lists the actual sequence of the identified peptide. Ox. M means that the methionine in the corresponding peptide is oxidized. The ions score is \(-10^* \log(P)\), where P is the probability that the observed match is a random event. Swiss-Prot accession no. <PDB>P17967</PDB>. Score = 295. (PDI_YEAST) Protein disulfide-isomerase precursor. Found in search of hirsch-3.pkl. Nominal mass (Mr) = 58191. Sequence coverage is 13%.
### S2. Strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS328</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801</td>
<td>(Vijayraghavan et al., 1989)</td>
</tr>
<tr>
<td>YG618</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1</td>
<td>(Jakob et al., 1998a)</td>
</tr>
<tr>
<td>YG796</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1Δalg9::KanMX</td>
<td>(Jakob et al., 1998a)</td>
</tr>
<tr>
<td>YG777</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1Δmms1::KanMX</td>
<td>(Jakob et al., 1998a)</td>
</tr>
<tr>
<td>YCJ1</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1Δhtm1::KanMX</td>
<td>(Jakob et al., 2001)</td>
</tr>
<tr>
<td>YG624</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1Δgls2::KanMX Δalg8::HIS3</td>
<td>(Jakob et al., 1998a)</td>
</tr>
<tr>
<td>YG695</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1Δgls2::KanMX</td>
<td>This study</td>
</tr>
<tr>
<td>YSC8</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1Δalg3::HIS3</td>
<td>This study</td>
</tr>
<tr>
<td>YSC17</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1Δalg3::HIS3 Δhtm1::KanMX4</td>
<td>This study</td>
</tr>
<tr>
<td>YSC24</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 tyr1 prc1-1Δalg3::HIS3 Δlyos9::HIS3</td>
<td>This study</td>
</tr>
<tr>
<td>YG1330</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1HTM1::HTM1-TEV-ProA-His7-KanMX</td>
<td>This study</td>
</tr>
<tr>
<td>YDO1</td>
<td>MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1HTM1::htm1Δs18-796-TEV-ProA-His7-KanMX</td>
<td>This study</td>
</tr>
<tr>
<td>YTX140</td>
<td>MATα prc1-1, trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>(Biedener et al., 1997)</td>
</tr>
<tr>
<td>YCH030</td>
<td>MATα HTM1-13myc(HIS3), prc1-1, trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH036</td>
<td>MATα HTM1-13myc(HIS3), 6xHA-HRD3, prc1-1, trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH046</td>
<td>MATα HTM1-13myc-TAP-Tag(TRP1), prc1-1, trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH133</td>
<td>MATα PDI-3xHA-HDEL(TRP1), HTM1-13myc(HIS3), prc1-1, trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH135</td>
<td>MATα pdi1Δ459::3xHA-HDEL(TRP1), HTM1-13myc(HIS3), prc1-1, trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH137</td>
<td>MATα htm1Δ749::13myc(HIS3), PDI-3xHA-HDEL(TRP1), prc1-1, trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH175</td>
<td>MATα htm1Δ582::13myc(HIS3), PDI-3xHA-HDEL(TRP1), trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH176</td>
<td>MATα htm1Δ647::13myc(HIS3), PDI-3xHA-HDEL(TRP1), trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH177</td>
<td>MATα htm1Δ683::13myc(HIS3), PDI-3xHA-HDEL(TRP1), trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
<tr>
<td>YCH178</td>
<td>MATα htm1Δ547::13myc(HIS3), PDI-3xHA-HDEL(TRP1), trp1-1(am), his3Δ200, ura3-52, lys2-801, leu2-3, -112</td>
<td>This study</td>
</tr>
</tbody>
</table>
References


Figure Legends

Figure 1. Overexpression of *HTM1* leads to altered N-glycan processing.

(A) $[3^H]$ mannose-labeled N-linked glycan profiles from strain YG618 transformed with YEp352 (wild type [wt] + empty vector) or YEp352-HTM1 (wild type + *HTM1*) and strain YCJ1 carrying YEp352 ($\Delta htm1$ + empty vector). M8, M7, and M5 mark the elution positions of Man$_8$GlcNAc$_2$, Man$_7$GlcNAc$_2$, and Man$_5$GlcNAc$_2$, respectively. Labeling of the glycans was performed for 20 min. N-glycans were released after the preparation of whole cell protein extracts with peptide N-glycosidase F, purified, and analyzed by HPLC analysis.

(B) Relative intensity of the Man$_9$GlcNAc$_2$ (M8), Man$_7$GlcNAc$_2$ (M7), and Man$_5$GlcNAc$_2$ (M5) peaks in the cells from A. Deviations represent the experimental error of three measurements of the experiment depicted in A.

(C) CPY$^*$ degradation analysis in strain YG618 transformed with YEp352 (wild type + empty vector) or YEp352-HTM1 (wt + *HTM1*) and strain YCJ1 carrying YEp352 ($\Delta htm1$ + empty vector). The graph represents the amount of remaining CPY$^*$ at the indicated chase time points. Each time point represents the mean +/- standard deviation of three independent experiments.

Figure 2. *HTM1* overexpression generates a Man$_7$GlcNAc$_2$ glycan.

(A) Pulse-chase analysis of the N-glycan processing in YG618 transformed with YEp352 (empty vector) or YEp352-HTM1 (*HTM1* overexpression). Cells were pulsed for 10 min with $^3$[H]-mannose and chased for the different times indicated on the left. M9, M8 and M7 mark the elution positions of Man$_9$GlcNAc$_2$, Man$_8$GlcNAc$_2$, and Man$_7$GlcNAc$_2$, respectively.

(B) Quantification of the relative glycan intensities at 0 and 30 min of chase. The values represent the mean of seven independent experiments + standard deviation. *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.001$, determined by using paired Student’s *t* test. HEX, hexose.
Figure 3. Processing of N-glycans by glucosidase I, glucosidase II, and mannosidase I is required for efficient trimming by Htm1p.

(A) N-glycan pulse-chase analysis over two time points (10-min pulse; 0- and 30-min chase). Graphs display N-glycan profiles obtained in YG618 (wild type [wt]), YG695 (Δgls2), YG624 (Δalg8Δgls2) and YG777 (Δmns1), transformed each with YEp352 (empty vector) or YEp352-HTM1 (HTM1). The symbolized N-glycans shown at the left represent the structure of the main N-glycan, which is produced in the ER of the corresponding cells. M7, Man7GlcNAc2; M7G1, Glc1Man7GlcNAc2; M8, Man8GlcNAc2; M8G2, Glc2Man8GlcNAc2; M8G1, Glc1Man8GlcNAc2; M9, Man9GlcNAc2.

(B) Quantification of HTM1-dependent processing products in the four different strains shown in A. The values represent the relative level + standard deviation of the HTM1-dependent processing products from three independent measurements after a 30-min chase.

Figure 4. Analysis of the Man7GlcNAc2 oligosaccharide by α1,2-exomannosidase digestion.

(A) Schematic representation of the three possible Man7GlcNAc2 product isoforms A, B, and C (M7A, M7B and M7C) upon removal of one mannose unit from Man8GlcNAc2 (M8). Depending on whether an α1,2-linked mannose of branch A or C or the α1,3-linked mannose of branch B had been trimmed, two or three α1,2-exo-mannosidase-sensitive mannoses remain. The products A and C can thus be distinguished from glycan B. M, mannose.

(B) HPLC profiles of the radiolabeled isolated Man7GlcNAc2 (M7) peaks from YG618 containing YEp352 (empty vector) or YEp352-HTM1 (HTM1 overexpression). Isolated glycans were left untreated (-) or treated (+) with the α1,2-exomannosidase from T. reesei and analyzed by HPLC. M5 represents the elution peak of a Man5GlcNAc2 that was used as a standard (third trace).

(C) Quantification the Man5GlcNAc2 (M5) oligosaccharide obtained from the Man7GlcNAc2 oligosaccharide after digestion with α1,2-exomannosidase. The Man4GlcNAc2 (M4) and the Man5GlcNAc2 oligosaccharide levels in the elution profiles depicted in B were quantified, and the level of the Man4GlcNAc2 (M4) oligosaccharide was set to 1. The value represents one measurement of
oligosaccharides obtained from strain YG618 transformed with YEp352 (empty vector) and the mean of two independent experiments for YG618 transformed with YEp352-HTM1 (HTM1 overexpression).

Figure 5. Structure function analysis of Htm1p.

(A) Strain SS328 was transformed with plasmids expressing the TPH7-tagged Htm1p (Htm1-TPH7; lane 1), and the mutant forms of Htm1p given at the right (lanes 2–4) or empty vector DNA (lane 5). Extracts were prepared, separated by SDS-PAGE, and transferred to nitrocellulose. Blots were probed with peroxidase-antiperoxidase-specific antibodies (top) and HXK-specific serum (bottom) to verify equal loading of the gel. The position of molecular-size markers is given at the left. The blot shows equivalent expression of wild-type and mutant forms of Htm1p fusions.

(B) N-glycan analysis after a pulse of 10 min and a chase of 0 and 30 min. Graphs display N-glycan HPLC profiles obtained in YG618 transformed with empty vector or plasmid DNA expressing the different mutant forms of tagged Htm1p given at the left. Only the overexpression of tagged wild-type Htm1p resulted in enhanced production of Man$_7$GlcNAc$_2$ (M7) oligosaccharide. M8, Man$_8$GlcNAc$_2$.

(C) Quantification of the Man$_7$GlcNAc$_2$ glycan levels obtained from the different strains given in B. N-glycan profiles of a 30-min chase as shown in B were quantified, and the relative level of the Man$_7$GlcNAc$_2$ is given. Relative intensities of the peaks depicted in B are plotted + standard deviations of two independent experiments.

(D) Mutations in the catalytic domain of Htm1p and truncation of the C-terminus abolishes Htm1p function in CPY* degradation. Strain YCJ1 was transformed with empty vector ($\Delta$htm1 + empty vector) and the different mutant constructs indicated above the lanes. The cells were chased with cycloheximide for the time indicated, extracts were prepared, and the proteins were separated by SDS-PAGE. After transfer to nitrocellulose, CPY* was detected with serum directed against CPY. The position of the band representing CPY* is given at the right. The secondary antibody used in this experiment also reacted with the protein A in the TPH7 tag of the Htm1 proteins (*, position of the band representing full-length protein; **, band representing truncated Htm1 protein).
Detection of HXK levels made a quantification of the CPY* degradation in these experiments possible.

(E) Values given represent the mean value of remaining CPY* levels from three independent experiments. The standard deviation is given. MW, molecular weight.

Figure 6. Htm1p physically interacts with Pdi1p.

(A) Microsomes were prepared from control cells or cells expressing Htm1-Myc-TAP. After solubilization, both lysates were incubated with IgG beads. The bound material was subsequently liberated by digestion with TEV protease and analyzed by SDS-PAGE. Visualization of the protein content by silver staining revealed Htm1-Myc and two bands around 62 kD that were clearly absent from the control lane (compare lane 2 with lane 3). Proteins were excised from the gel, trypsinized, and subjected to mass spectrometry. Both bands at 62 kD were identified as Pdi1p. The identity of the Htm1-Myc band was also confirmed by mass spectrometry. For size comparison, a molecular weight (MW) standard was loaded in lane 1.

(B) Detergent-solubilized microsomes from yeast strains expressing full-length Pdi1-HA-HDEL (lanes 1 and 2) or Pdi1-Δ459-HA-HDEL (lane 3) together with Htm1-Myc (lanes 2 and 3) were analyzed by Western blot (WB) to confirm the presence of the tagged proteins (left). HA-tagged proteins were immunoprecipitated, and analysis of the anti-HA precipitates by Western blotting (right) revealed that Htm1-Myc interacted with full-length Pdi1-HA-HDEL (lane 5) but not with Pdi1-Δ459-HA-HDEL (lane 6).

(C) Microsomes were isolated from yeast strains expressing either full-length Htm1-Myc or the indicated truncated variants of Htm1p together with Pdi1-HA-HDEL (lanes 2–7). Immunoblotting of the total lysates confirmed the expression of the tagged proteins (left, top and middle), and immunoprecipitation of Pdi1-HA-HDEL was confirmed (top right). Probing for the Myc-tagged Htm1p constructs revealed a readily detectable interaction between Pdi1-HA-HDEL and Htm1-, Htm1-Δ749-, Htm1-Δ683- and -Δ647-Myc truncates (lanes 9-12) but not with shorter variants of Htm1p (lanes 13 and 14). Probing of the supernatants after immunopurification with serum against
HA confirmed the quantitative depletion of Pdi1-HA-HDEL from the lysates (bottom left). The secondary antibody used in the Western blot also reacted with the heavy chain (HC) of the HA antibody, used in the immunopurification (IP).

Figure 7. HTM1-independent but YOS9-dependent degradation of CPY* in Δalg3 cells.

(A) N-glycan structures of the wild-type Man₈GlcNAc₂ (Mns1p product), Htm1p product Man₇GlcNAc₂ C and the N-gycans prevailing in Δalg3, Δalg9 or Δalg12 cells. Man₇GlcNAc₂ C and Man₅GlcNAc₂ in Δalg3 both expose an α1,6-linked terminal mannose (circles).

(B-G) Analysis of CPY* degradation in different mutant strains. The cells were chased with cycloheximide for the time indicated, extracts were prepared, and the proteins were separated by SDS-PAGE. After transfer to nitrocellulose, CPY* was detected with serum directed against CPY. The position of the band representing CPY* is given at the right. Detection of HXK levels made a quantification of the CPY* degradation in these experiments possible. (B and C) Analysis of wild-type (wt), Δalg3, and Δalg9 cells. (D and E) Analysis of wild-type, Δhtm1, Δalg3, and Δalg3Δhtm1 cells. (F and G) Analysis of wild-type, Δyos9, Δalg3, and Δalg3Δyos9 cells. The values given in the relative quantification experiments are the mean of three independent experiments (B and C are two independent experiments). Standard deviations are indicated.
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5
Figure 6
Figure 7
Chapter 4

Analysis of the substrate specificity of Htm1p
Introduction

Class I α1,2-mannosidases constitute a group of enzymes that belong to the family GH47 of the glycoside hydrolase classification (http://www.cazy.org) (Henrissat and Bairoch, 1996). Members of this protein family reveal significant amino acid similarity, share characteristic signature motifs and X-ray crystallography suggests that they have a similar overall three-dimensional structure consisting of a (αα)7-helix barrel. They are calcium-dependent inverting glycosidases and are inhibited by both 1-deoxymannojirimycin and kifunensine (Tremblay et al., 2007). Despite the similarities in protein sequence and presumed protein fold, they differ in their enzymatic activities. Accordingly, they are classified into subgroups. The first subgroup comprises the yeast and human endoplasmic reticulum Class I α1,2 mannosidases, Mns1p and ER mannosidase I respectively. Primary, they process Man9GlcNAc2 glycans to form the Man8GlcNAc2 isomer B (Camirand et al., 1991; Gonzalez et al., 1999; Tremblay and Herscovics, 1999). The second subgroup includes the Golgi α1,2 mannosidases IA (Bause et al., 1992; Lal et al., 1994), IB (Herscovics et al., 1994; Lal et al., 1998; Schneikert and Herscovics, 1994), and IC (Tremblay and Herscovics, 2000), as well as the α1,2-mannosidases from filamentous fungi (Ichishima et al., 1999; Yoshida et al., 1993) and from insect cells (Kawar et al., 2000). The function of the mammalian enzymes in this subfamily is to form the Man5GlcNAc2 intermediate glycan which is necessary for the generation of complex and hybrid N-glycans in vivo (Herscovics, 1999). The third subgroup includes the EDEM/Htm1p family.

Comparison of the substrate specificities of the ER (subgroup 1) and Golgi (subgroup 2) α1,2 mannosidases revealed that they have differences in both, the extent of mannose trimming, as well as branch specificity for substrate recognition. ER mannosidase I preferentially cleaves the α1,2 mannose on the central branch of the Man9GlcNAc2 structure to produce the Man8GlcNAc2 B isomer (Herscovics, 1999; Lipari and Herscovics, 1994; Puccia et al., 1993; Ziegler and Trimble, 1991). In contrast, the Golgi mannosidase IA and IC preferably cleave the terminal mannoses of the A branch, and Golgi mannosidase IB, first cleaves the C branch of the Man9GlcNAc2 (and also Man8GlcNAc2) structure. As second action Golgi mannosidase IA and IB cleave vice versa the terminal mannose of the C (Golgi
mannosidase IA) or the A branch (Golgi mannosidase IB). Third, the remaining α1,2-linked mannose on the A branch is removed, and in a last reaction the terminal B branch mannose is removed. Golgi mannosidase IC trims first the A branch. However, in the second reaction it can trim the second mannose on the A branch before the terminal C branch mannose. In any case, Golgi mannosidase IC also removes the terminal B branch mannose in the last reaction. Therefore, the Golgi mannosidase I families have complementary specificity to that of ER mannosidase I (Lal et al., 1998; Tremblay and Herscovics, 2000).

As we had found that Htm1p has α1,2 mannosidase activity, we determined the N-glycan specificity of this GH47 subgroup 3 member in more detail.
Results

Residual Htm1p activity requires the C-branch $\alpha_1,6$ mannose

The observation that CPY* degradation in the $\Delta$alg3 background was independent of HTM1 suggested that the ER-degradation signal is already present in this structure. Moreover, genetically tailored Man$_7$GlcNAc$_2$ C glycan renders CPY* degradation HTM1 independent (Quan et al., 2008). This suggests no requirement for A branch processing by Htm1p for efficient degradation. To analyze whether Htm1p was able to process other mannoses than the terminal C branch mannose, we challenged Htm1p to process N-glycans with a complete A branch, or with a complete A and B branch, but lacking the C branch. Therefore, we used $\Delta$alg9 and $\Delta$alg12 mutant cells. $\Delta$alg9 cells produce a Man$_6$GlcNAc$_2$ glycan with an intact A branch and HTM1 overexpression did not result in N-glycan processing of this structure. In $\Delta$alg12 cells, the A branch and the B branch are both completely synthesized but the C branch is lacking, resulting in a Man$_7$GlcNAc$_2$ isomer which is different from the Htm1p Man$_7$GlcNAc$_2$ C product (Figure 1). Due to the missing C branch, this glycan is a poor substrate for ER-mannosidase I (Cipollo and Trimble, 2002; Ziegler and Trimble, 1991). Indeed, we detected only slow production of a Man$_6$GlcNAc$_2$. HTM1 overexpression did not lead to faster or further processing of the N-glycan processing in this cell (Figure 1). This indicated that neither the terminal mannose of the A branch nor the terminal mannose of the B branch in these structures, were targets of Htm1p. In contrast, when we overexpressed HTM1 in the ERAD deficient $\Delta$ubc7 mutant, besides Man$_7$GlcNAc$_2$, some Man$_6$GlcNAc$_2$ and to lesser extent Man$_5$GlcNAc$_2$ glycans were detectable, indicating weak residual specificity of Htm1p for the A branch (Figure 1). Since in the $\Delta$alg9 and $\Delta$alg12 cells ERAD substrates also accumulate (Jakob et al., 1998) (Chapter 3), we deduced that Htm1p requires the $\alpha_1,6$-linked mannose of the C branch for this residual activity.
Figure 1. Residual A branch processing activity requires the C branch α1,6-linked mannose.

N-glycan pulse chase analysis over two time points (10 minutes pulse; 0 and 30 minutes chase). Graphs display N-glycan profiles obtained in wild type (wt), Δalg9, Δalg12 and Δubc7, transformed each with YEp352 (empty vector) or YEp352-HTM1 (HTM1). The symbolized N-glycans shown at the left represent the structure of the main N-glycan, produced in the ER of the corresponding cells. MX is ManαGlcNAc₂.
HTM1 overexpression in Δmns1 cells generates the Man₈GlcNAc₂ C isomer

When Htm1p was overexpressed in Δmns1 cells, we detected greatly reduced N-glycan processing activity, leading to the conclusion that prior processing by mannosidase I was required for Htm1p function in vivo (Chapter 3). However, we were able to detect small amounts of a Man₈GlcNAc₂ glycan. As we had concluded that in wild type cells, Htm1p removes one mannose from the C branch, we assumed that the Man₈GlcNAc₂ glycan product, generated by Htm1p in Δmns1 was the Man₈GlcNAc₂ C isomer. Since, a terminal α1,6 linked mannose constitutes the degradation signal, we tested whether the generated Man₈GlcNAc₂ isomer in the Δmns1 background was able to promote ERAD. Therefore, we analyzed the ERAD competence of Δmns1 cells that overexpressed HTM1, in order to compensate for the reduced specificity of Htm1p for the Man₈GlcNAc₂ glycan. We assayed both, CPY* degradation rates, and the ERAD dependent temperature sensitive growth phenotype of the sst3-7 mutant membrane protein. Sst3-7 cells which are proficient in ERAD do not grow (Spirig et al., 1997), while ERAD deficient cells like Δmns1 or Δhtm1 grow (Jakob et al., 2001), (Figure 2A upper panel). HTM1 overexpression abolished the Δhtm1 dependent growth rescue of sst3-7 mutants at 30°C. Importantly, HTM1 overexpression complemented also the Δmns1 and to great extent the Δhtm1Δmns1 mutant phenotype in the sst3-7 background (Figure 2A lower panel). As previously reported, CPY* degradation is impaired in Δmns1 cells, and we found that overexpression of HTM1 complemented also the CPY* degradation phenotype. In contrast, HTM1 overexpression did not increase degradation rates in the Δalg12 mutant (Figure 2B). Taken together, these results confirmed our assumption that HTM1 overexpression in the Δmns1 background generated the Man₈GlcNAc₂ C isomer. Moreover, the result indicates that the intact B branch which is retained in this isomer, does not interfere with Yos9p recognition or function, respectively. In line with this is the finding that Yos9p binds in vitro the Man₈GlcNAc₂ C isomer with a similar affinity like Man₇GlcNAc₂ C isomer (Quan et al., 2008).
Figure 2. HTM1 overexpression complements the Δmns1 ERAD phenotypes.

A) Stt3-7 temperature sensitive growth phenotype. Serial dilutions of wild type (wt), Δhtm1, Δmns1, single and double mutants, stt3-7, stt3-7 Δhtm1, stt3-7 Δmns1, stt3-7 Δhtm1 Δmns1 were transformed with either YEp352 (empty vector) or with YEps52-HTM1 (HTM1 overexpression). A 1:10 serial dilution of exponentially growing cells, starting at 5 x 10^4 cells was spotted onto plates containing minimal medium, lacking uracil, as indicated on the left. Plates were incubated at the permissive temperature of 23°C (left panel) or at the restrictive temperature of 30°C (right panel) for 2 days. B) CPY* degradation analysis. Degradation was analyzed in Δmns1 and Δalg12 cells, transformed with empty vector (plasmid) or HTM1 and/or MNS1 or ALG12 plasmid. The graph represents the amount of remaining CPY* at the indicated chase time points.
Discussion

The results from Chapter 3 suggest that the first and the second glucose, as well as the terminal B-branched mannose negatively regulate Htm1p activity. Therefore, processing of the Glc$_3$Man$_8$GlcNAc$_2$ glycan to Man$_7$GlcNAc$_2$ is required for full Htm1p activity. As we show in this chapter, prolonged exposure of the primary product Man$_7$GlcNAc$_2$ C may lead to further but slow processing by Htm1p. This activity requires the presence of the $\alpha$1,6-linked mannose of the C branch (Figure 3). Of note, Mns1p has also been shown to require the C branch for full functionality (Cipollo and Trimble, 2002; Ziegler and Trimble, 1991). Htm1p shows poor activity on Man$_9$GlcNAc$_2$. Yet, the minute fraction that is processed in this background seems to be devoid of the terminal mannose of the C branch.

Figure 3. Sugar determinants for Htm1p action.
A) The arrow indicates the primary target of Htm1p, which is the C branch. Red dots indicate that the glucoses n, m, and the mannose i (Chapter 1, Figure 2) negatively regulate Htm1p activity. B) Shows the $\alpha$1,6-linked mannose, that positively regulates residual Htm1p activity for the A branch.

But for the specificity for $\alpha$1,2-linked mannoses, Htm1p activity does not show a clear resemblance to one of the other two defined GH47 family subgroups. In terms of extent of mannose removal, Htm1p appears similar to the subgroup 1 or Mns1p, respectively, as it removes one mannose, and for enhanced trimming prolonged N-glycan exposure is required (Herscovics et al., 2002). However, in contrast to Mns1p,
Htm1p has specificity for the C branch which appears similar to Golgi mannosidase IB, a member of the subgroup 2, which has primary preference for the C branch on both, the $\text{Man}_9\text{GlcNAc}_2$ and the $\text{Man}_8\text{GlcNAc}_2$ B glycan (Lal et al., 1998). Moreover, subgroup 1 and subgroup 2 members act complementary and independently of each other, while Htm1p is functionally dependent on Mns1p.

The structural basis for the branch specificity of ER mannosidase I was previously described for *S. cerevisiae* Mns1p (Vallee et al., 2000). The crystal structure of Mns1p revealed that it forms an ($\alpha\alpha$)$_7$-helix barrel (Vallee et al., 2000). The nine highly conserved acidic residues and the calcium ion, all of which are essential for the catalytic activity (Lipari and Herscovics, 1999) are located at the bottom of a 15Å deep cavity within the barrel. It has been postulated that the specificity of ER mannosidase I subgroup to remove the B branch mannose is conferred by the dimension of the carbohydrate binding site. The arginine R273 plays a critical role in reducing the degree of freedom of the oligosaccharide entering the active site (Romero et al., 2000; Vallee et al., 2000). Mutation of R273 to leucin resulted in a low efficiency enzyme with a hybrid activity between ER mannosidase I (it cleaved first the B branch) and Golgi mannosidase I (cleaving proceeded to $\text{Man}_5\text{GlcNAc}_2$) (Romero et al., 2000). Moreover, structural investigation of the Class I $\alpha$-mannosidases from *Penicillium citrinum* and *Trichoderma reesei*, which have similar substrate specificity as Golgi mannosidases revealed smaller and uncharged residues at this position. This, combined with a more spacious substrate binding site, was proposed to provide the basis for the extended specificity of *Penicillium citrinum* (Lobsanov et al., 2002). Interestingly, amino acid sequence alignment reveals a serine in Htm1p at the Mns1p R273 position. Serine is smaller than arginine, but in contrast to leucin it is polar. This indicates that this S277 might be a key residue in the regulation of Htm1p specificity. In contrast, the mammalian EDEM proteins have in this position a glycin. Therefore, they are similar to the Golgi subfamily (Figure 4). Indeed, EDEM3 has been shown to process N-glycans more extensively (Hirao et al., 2006). However, investigation of the murine Golgi mannosidase IA revealed a different situation. It has a constricted binding cleft, in which an alternative oligosaccharide conformation is bound. This leads to the exposure of the terminal C branch mannose instead of the B branch to the catalytic activity (Tempel et al., 2004).
Figure 4: Amino acid sequence alignment of Htm1p and Mns1p homologs from different organisms.

Blue arrows indicate the position of Mns1p R273 or Htm1p S277, respectively. E222 and D279 (Htm1p numbering) are framed and represent residues that are conserved throughout the GH47 family. Mutagenesis of these residues abolished Htm1p function (Chapter 3). Modified from (Oggier, 2006).

Aiming to unravel the structure of Htm1p, thorough in vitro investigation of the N-glycan specificity and the substrate protein determinants for Htm1p and EDEM proteins will extend the knowledge on structure-function relationships in GH47 α1,2-mannosidases.
Material and methods

Yeast strains and plasmids

Standard yeast media and genetic techniques were used (Guthrie and Fink, 1991). Yeast strains used in this study were: YG618 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1) (Jakob et al., 2001); YG840 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 Δalg12::kanMX4), YG414 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 Δalg9::KanMX) (Burda et al., 1996), YG1197 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 Δubc7::KanMX); YCJ1 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1 Δhtm1::KanMX) (Jakob et al., 2001); YG777 (MATα ade2-101 ura3-52 his3Δ200 lys2 Δmns1::KanMX prc1-1) (Jakob et al., 1998); Δyhr204Δmns1 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 prc1-1 Δmns1::KanMX Δhtm1::KanMX); YG543 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 leu2 stt3-7) (Spirig et al., 1997); YCJ82 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 stt3-7 Δhtm1::KanMX), YG950 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 stt3-7 Δmns1::KanMX). YCJ88 (MATα ade2-101 ura3-52 his3Δ200 lys2-801 stt3-7 Δhtm1::KanMX Δmns1::KanMX) (Jakob et al., 2001); YG808 (MATα ade2-101 his3Δ200 lys2-801 ura3-52 Δalg12::KanMX prc1-1).

Plasmids used in this study were: YEpl52 (80), YEpl52-HTM1 (626), pRS316 (41), pRS316-MNS1 (636), pRS316-ALG12 (467).

Analysis of CPY* degradation

The cycloheximide chase was performed as described previously (Jakob et al., 1998). In brief, yeast cells were grown at 30°C in appropriate medium to mid-log phase, corresponding to an OD₆₀₀nm of 0.8 - 1.2. 3 x 10⁸ cells were harvested and resuspended in medium. The chase was initiated by addition of cycloheximide (final concentration 100 µg/ml) and performed at 30°C. 1 x 10⁸ cells were removed at each time point and transferred into NaN₃ (final concentration 0.1% [wt/vol]) on ice, immediately pelleted and flash-frozen in liquid nitrogen. Whole cell protein extracts were prepared using glass beads, 1% (wt/vol) SDS, 50mM Tris-HCl (pH 7.5), 2mM PMSF. Proteins were subjected to reducing 7% SDS-PAGE and electroblotted to nitrocellulose membranes. CPY* was immunologically detected with anti-CPY
antibody (rabbit) at a dilution of 1:1000 (Zufferey et al., 1995) and goat anti-rabbit IgG-horseradish peroxidase at 1:3000 (Santa Cruz).

$[^3]H$-mannose labeling and analysis of N-linked glycans

Cells were grown at 30°C in appropriate medium to mid-log phase, corresponding to an OD$_{600nm}$ of 0.8-1.2. $n 	imes 5 	imes 10^8$ cells ($n =$ number of analyzed time points) were harvested and washed in YP0.1%D (YPD containing 0.1% glucose). Labeling was performed in $n 	imes 200$ μl YP0.1%D containing $n 	imes 100$ μCi D-2-$[^3]H$-mannose (500GBq/mmol; Hartmann Analytic) and incubated for 10 min at 30°C. The radioactivity was chased by the addition of D-mannose to a 2% (wt/vol) final concentration at 30°C incubation. The chase was stopped by removing $5 	imes 10^8$ cells into TCA on ice (final concentration 10% [vol/vol]). TCA precipitates were washed twice with cold acetone and air dried for 20 min. The pellets were resuspended in buffer S (20 mM NaP, pH 7.5, 0.5% SDS, 40 mM DTT) and vortexed with glass beads at 50°C for one hour. Lysates were cleared by centrifugation, and iodoacetamide was added to 50-mM final concentration. Samples were incubated for 30 min at 37°C. Oligosaccharides were cleaved from proteins by digestion with peptide N-glycosidase F according to the recommendations of the manufacturer (BioConcept). Cleanup of oligosaccharides was modified from (Grubenmann et al., 2004). Prepacked C$_{18}$ Sep Pak columns (Waters) were connected to columns (extended-volume empty reservoirs; Socochim SA) that were packed with Supelclean ENVI-Carb 120/400 (Sigma-Aldrich). The combined columns were equilibrated with methanol, acetonitrile, acetonitrile/H$_2$O (25:75; vol/vol), and acetonitrile/H$_2$O (2:98; vol/vol). Samples were loaded onto the columns after supplementation with acetonitrile/H$_2$O (2:98; vol/vol). Columns were washed with acetonitrile/H$_2$O (2:98; vol/vol) and glycans eluted from the ENVI-Carb column with acetonitrile/H$_2$O (25:75; vol/vol). The solvent was evaporated in a speed vacuum. For HPLC analysis of the oligosaccharides, a liquid chromatography (LC)-NH$_2$ column (250 x 4.6mm; Sigma-Aldrich) (Cacan et al., 1993) including an LC-NH$_2$ guard column was used. Oligosaccharide samples in acetonitrile/H$_2$O (70:30; v/v) were filtered through a 0.45-μm filter (Millipore) and injected on the equilibrated (acetonitrile/H$_2$O [70:30; vol/vol], 30 min) system using an autosampling device.
(Merck/Hitachi AS-2000). The gradient was acetonitrile/H$_2$O (70:30; vol/vol) to acetonitrile/H$_2$O (45:50; vol/vol) over 90 min, 5 min at acetonitrile/H$_2$O (45:50; vol/vol), returning to acetonitrile/H$_2$O (70:30; vol/vol) over 5 min and washing for 20 min at acetonitrile/H$_2$O (70:30; vol/vol) before injection of the following sample using a pump (Merck/Hitachi L-2600A; Sigma Aldrich). The eluate from the column was mixed continuously with scintillation fluid (FLO-Scint A; Packard) in a ratio of 1:1.5 (eluate:scintillation mix, vol/vol), and radioactivity was monitored with a flow monitor (FLO-ONE A-525, Packard) as described previously (Zufferey et al., 1995).
References


Cipollo, J.F., and R.B. Trimble. 2002. The *Saccharomyces cerevisiae* alg12D mutant reveals a role for the middle-arm a1,2Man- and upper-arm a1,2Man-a1,6Man- residues of Glc3Man9GlcNAc2-PP-Dol in regulating glycoprotein glycan processing in the endoplasmic reticulum and Golgi apparatus. *Glycobiology.* 12:749-62.


