1 Introduction

The 26S proteasome is a cellular protein degradation machine responsible for eliminating incorrectly folded proteins from the cell. The proteasome works as a part of cellular protein quality control, which is operated by two interacting mechanisms: 1) the unfolded protein response (UPR) and 2) ER-associated degradation (ERAD; Figure 1). The unfolded protein response is activated by the stress in the endoplasmic reticulum (ER) caused by misfolded or misassembled proteins. From the ER, the incorrectly folded proteins are transferred into the cytoplasm and destroyed by ER-associated degradation by the proteasome.

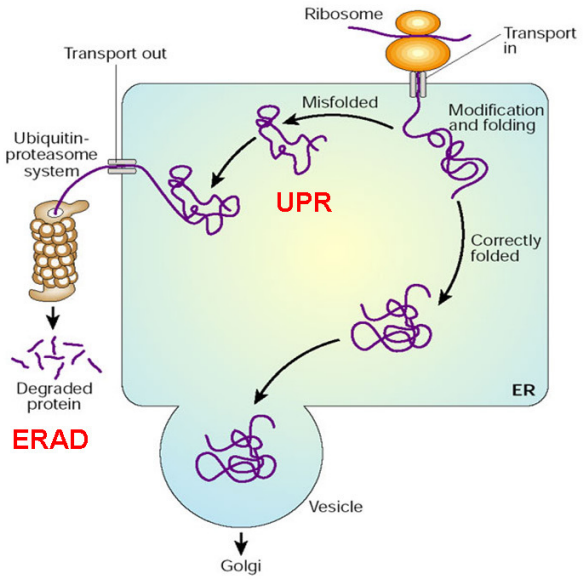

Figure 1. UPR and ERAD are part of protein quality control in the cell. After translocation into ER, correctly folded proteins are transferred into Golgi. Misfolded proteins are translocated from ER into the cytoplasm and degraded by the proteasome (modified from Dobson 2003).

The proteasome is a cylindrically-shaped complex of protein subunits and was first observed in 1968 in a human erythrocyte sample using electron microscopy (Harris, 1968). Recently, the structure and function of the 26S proteasome has been the target of studies with several different organisms including microbes. The 26S proteasome has been widely studied in human cells and in the yeast *Saccharomyces cerevisiae*. However, limited information is available on its structure and function in filamentous fungi. The filamentous fungus *Trichoderma reesei* used in the current study is utilised widely in production of proteins for industrial applications. While the yield of homologous proteins secreted into

the culture medium by *T. reesei* has been very high, up to 100 g/L (Cherry and Fidantsef, 2003), yields of heterologous proteins have remained low. Generally, it is believed that protein mis-folding and incorrect processing through the secretory pathway results in the low production levels of heterologous proteins in *T. reesei*. Protein quality control in filamentous fungi is relatively poorly understood although there have been several studies into the function of the unfolded protein response (UPR) in *Aspergillus* species (Mackenzie *et al*., 2005 Valkonen *et al*., 2003; Sims *et al*., 2005) and *T. reesei* (Collén *et al*., 2005; Valkonen *et al*., 2004). The release of the *T. reesei* genome sequence has broadened the scope of investigations into the quality control mechanism in this organism at a whole genome level allowing the combined use of microarray and proteomic techniques.

1.1 Protein quality control in the cell

The eukaryotic cell is a highly organised structure with distinct subcellular compartments. For a cell to function efficiently, each of its numerous proteins must be localised to the correct cellular membrane or cellular compartment. Proteins destined for the plasma membrane, cell surface or secretion outside the cell are all translocated from the cytoplasm into the ER for post-translational modifications before delivery to their next destination. Newly synthesised proteins in the ER must pass a tight protein quality control (PQC) process to prevent incorrectly folded proteins accumulating in the cell (Dobson 2003). Therefore, PQC plays a vital role in maintaining a proper balance between protein synthesis, maturation and degradation, which is crucial for cell survival (Mast *et al*., 2005).

Protein quality control in the ER is able to provide a balance between retaining and degrading potentially harmful protein products without preventing the export of biologically active proteins (Sitia and Braakman, 2003). The PQC systems operate in two interacting ways to minimise harmful effects in the cell. First, accumulation of incorrectly folded proteins in the ER activates a transcriptionally-regulated pathway called the unfolded protein response, or UPR. Secondly, proteins which fail to fold into their correct three-dimensional structure are retained within the ER and degraded by a mechanism termed ER-associated degradation or ERAD (Sitia and Braakman, 2003; Friedlander *et al*., 2000).

The protein quality control mechanism can be described as a three-step process (Trombetta and Parodi 2003). The first step involves recognition of the aberrant ER proteins. This is followed by the action of ER chaperones and folding enzymes as well as the glycosylation machinery (Chevet *et al*., 2001). Molecular chaperones refold misfolded proteins and sufficiently matured proteins will be transported to the Golgi apparatus in the secretory pathway (Vashist *et al*., 2001). Misfolded or misassembled proteins will be sorted for retention in the ER (Hiller *et al*., 1996; Cabral *et al*., 2001). In the second step, misfolded proteins destined for destruction are relocated from the ER into the cytosol (Plemper *et al*., 1999; Wiertz *et al*., 1996). This retrograde transport step involves the action of the Sec61 translocation channel, which also mediates the import of proteins into the ER. In the final step, misfolded proteins are polyubiquitinated and proteolytically degraded by the 26S proteasome (Spiro, 2004; Werner *et al*., 1996).

The UPR and ERAD pathways are tightly interconnected (Friedlander *et al*., 2000; Figure 2). If protein folding in the ER is inefficient, removal of misfolded proteins will be performed jointly by UPR and ERAD (Travers *et al*., 2000).

In the absence of the ER degradation machinery, UPR has to deal with protein misfolding by activating the expression of factors involved in protein folding, vesicular transport or alternative degradation, such as activation of proteases (Fewell, *et al*., 2001). For ERAD to perform efficiently, it requires an intact UPR, and the consequent UPR induction increases ERAD capacity.

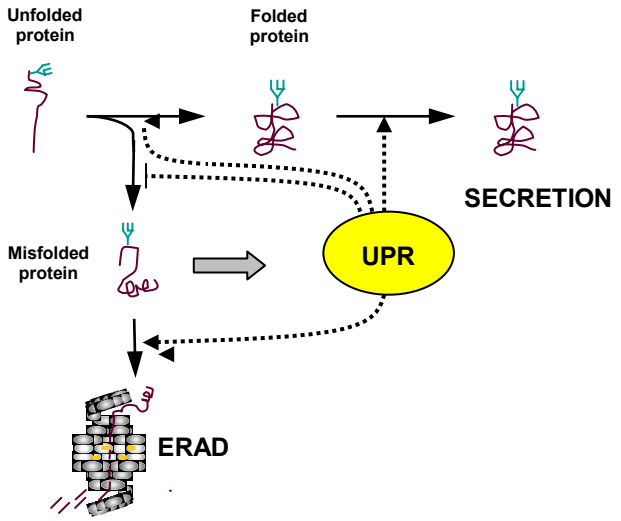

Figure 2. Diagram illustrating the interaction between UPR and ERAD. Proteins enter the ER in an unfolded form, where they fold and continue on to the secretory pathway but irreversibly misfolded proteins are eliminated by the ERAD machinery. Either cellular stress or loss of ERAD activity results in the accumulation of misfolded proteins and activation of the UPR (grey arrow). UPR acts to reduce the level of misfolded proteins by restoring protein folding to the native state, promoting secretion and enhancing the rate of ERAD. At the same time, the expression of misfolded proteins is reduced (Travers *et al*., 2000).

However, induction of UPR is not necessary under normal growth conditions, during which ERAD activity is sufficient and can cope with the low amounts of unfolded proteins. In contrast, when the ER is under stress, the UPR is activated, resulting in increased expression of ER folding enzymes (Friedländer *et al*., 2000; Travers *et al*., 2000). Stress in the cell can be caused by misfolded proteins produced by mutation, transcriptional or translational errors, incorrect folding, imbalanced subunit synthesis, improper trafficking, or damage caused by environmental conditions or metabolic by-products (Goldberg, 2003). Saloheimo *et al.* (2003) observed that the UPR pathway was induced during foreign protein expression in *T. reesei,* and it also led to a rapid down-regulation of genes encoding secreted proteins. Furthermore, a loss of ERAD can lead to strong UPR induction, and the loss of function of both components is lethal to the cell (Travers *et al*., 2000; Fewell *et al*., 2001).

1.1.1 Role of chaperones in protein quality control

Approximately one third of all proteins in a eukaryotic cell are translocated into the ER, which makes the ER lumen a crowded place with a protein concentration of more than 100 mg/mL (Kleizen and Braakman, 2004). Although the native structure of a protein is

encoded in its amino acid sequence, the ER greatly enhances protein folding efficiency because of the unique oxidising potential that supports disulfide bond formation during protein folding (Tu and Weissman, 2004). In the ER, the protein quality control maintains the balance of proteins with the help of molecular chaperones, which recognise and repair misfolded and aggregated proteins (Goldberg, 2003). Kleizen and Braakman (2004) reviewed several studies in which chaperones have been shown to function in multiprotein complexes. Examples include a binding protein (BiP) in complexes with Grp94, Grp170 and protein disulfide isomerase (PDI) and also lectin chaperones, CNX and CRT in complexes with BiP and Grp70.

Binding protein, BiP

Newly-synthesised proteins start to fold immediately during translation and translocation. In the ER lumen an unfolded nascent protein interacts and binds to the eukaryotic binding protein, BiP (also called glucose regulated protein, Grp78/ Kar2p). BiP is the most abundant ER chaperone and is closely related to the cytosolic Hsp70 (Nguyen 1991). BiP has an ATPase domain and a peptide binding domain that co-ordinate a repetitive cycle of ATP hydrolysis and ADP exchange, and thus stimulate the binding and release of the unfolded protein (Gething, 1999; Bateman, 2007). BiP also is bound to the lumenal domain of the regulators of UPR in eukaryotic cells under normal cell growth conditions. These include Ire1 (serine-threonine kinase / endoribonuclease), AFT6 (activating transcription factor 6) and PERK (protein kinase-like ER kinase). In the event of an increase in the amount of unfolded proteins within the ER, BiP is released from the lumenal domains of the regulatory proteins and performs its function, preferentially binding to misfolded proteins and thus preventing misfolded proteins to aggerate in the ER (Umebasyshi *et al*., 1999; Bertolotti *et al*., 2000). BiP resides in the ER and exists in two states, an adenoside diphosphate (ADP-) and an ATP-bound state. The ATP-ADP exchange reaction induces conformational changes in BiP. Following from this, BiP has low affinity to unfolded substrates in the ATP-bound form and high affinity in the ADP-bound form. ATP hydrolysis and the ADP-ATP exchange reaction of Bip are stimulated by other co□ chaperones such as DnaJ, GrpE and Lhs1p. The regulation of BiP chaperone activity consumes most of the ATP in protein folding (Schröder, 2008; Geysens *et al*., 2009).

Protein disulfide isomerase, PDI

Protein disulfide isomerase (PDI) primarily resides in the ER and catalyses the formation, and either the reduction or isomerisation of disulfide bonds. These reactions are dependent on the redox conditions of the substrate protein. The activity of PDI depends on a pair of cysteines, which are found in the motif Cys-X_a-X_b-Cys (Cys represents cysteine and X_a and X_b can be any amino acid) within a substrate, in a domain homologous to thioredoxin. PDI can catalyse the formation of disulfide bonds on the substrate proteins when the active cysteines of a substrate are present in an oxidised form (Frand and Kaiser 1999). PDI is also capable of catalysing the post-translational disulfide rearrangement of disulfide bonds in a substrate protein (Turano *et al*., 2002).

Calnexin, calreticulin and EDEM

Most of the proteins made in eukaryotic cells acquire *N*-linked glycans co-translationally in the ER and may interact with a special glycoprotein-specific chaperone system called the calnexin-calreticulin cycle. The lectin-like chaperone calnexin (CNX; membrane bound) and soluble calreticulin (CRT) bind carbohydrate moieties and together with ERp57, a thiol oxidoreductase (member of PDI family), they promote the proper folding of glycoproteins through a binding and release cycle (Ritter and Helenius, 2000). As proteins are translocated across the ER membrane, three glucose residues are added to the core glycans of the polypeptide chain by monosaccharyltransferase (see Fig. 3; Helenius and Aebi, 2004). The first glucose is then rapidly removed by glucosidase I, followed by removal of the second glucose by glucosidase II. The monoglucosylated core ligand then binds the glycoprotein to CNX or CRT. If the glycoproteins contain cysteine residues, disulfide bonds are transiently formed by ERp57 (Russel *et al*., 2004; Oliver *et al*., 1999; Helenius and Aebi, 2004). The complex dissociates, releasing the protein from CNX and CRT when the remaining glucose residue is trimmed by glucosidase II. The glycoprotein is now free to leave the ER, unless it is incorrectly folded and thus it is recognised by UDP-glucose glucosyltransferase (UGGT) which reglucosylates the *N*-glycan. The reglucosylation induces a new round of CNX/CRT binding, preventing the escape of the glycoprotein from the ER. A glycosylated protein stays in the cycle until it is either properly folded and oligomerised or enters the ERAD degradation pathway (Molinari *et al*., 2004; Cabral *et al*., 2002). Misfolded proteins targeted for degradation thus are marked with the *N*-linked

oligosaccharide mannose-N-acetylglucosamine ($Man_8GlcNAc_2$) by mannosidase. ER-degradation enhancing α-mannosidase-like protein (EDEM) associates with CNX (see Fig. 3). Htm1/Mnl1 is a mannosidase homologue which is thought to serve as a mannose lectin and to be responsible for directing misfolded glycoproteins into the retrotranslocation through the Sec61 channel and degradation pathway (Kanehara *et al*., 2007). Once the protein exits the ER, it undergoes ubiquitination and degradation by the 26S proteasome (Schrag *et al*., 2003; Wang and Hebert, 2003; Helenius and Aebi, 2004).

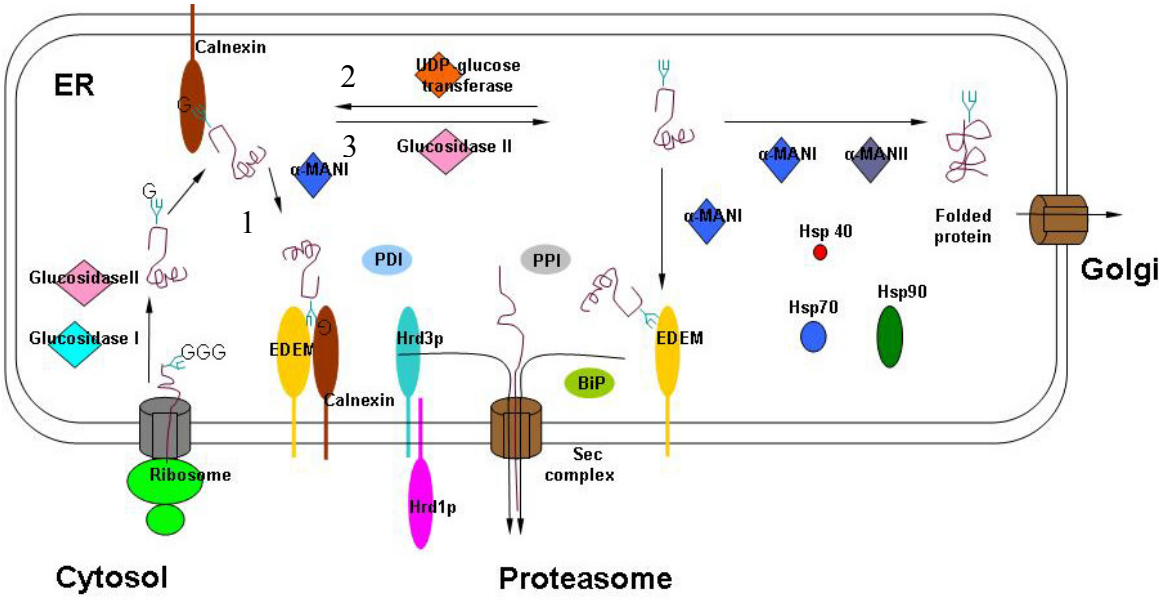

Figure 3. Diagram of participants during protein folding/unfolding in the ER. After translocation into the ER, substrate proteins are monoglycosylated following the interaction of glucosidases I and II. The monoglycosylated proteins could go three different ways. (1) The binding to calnexin or calreticulin is followed by the last glycan trimming with slow acting mannosidase I (ManI). Following the release of glycan from the substrate protein, the CNX/CRT complex allows EDEM to take the glycoprotein out of the CNX binding cycle. The EDEM-bound substrate then is targeted for translocation into the cytosol, where it will be deglucosylated and ubiquinated and is subjected to proteasomal degradation. (2) UDP-glucose transferase re□ glucosylates incompletely folded glycoproteins and rebinds them to CNX, keeping them in the cycle until they are either properly folded or sent for degradation. (3) If the substrate protein is folded correctly, it will be released from CNX and glucosidase II will remove the remaining glucose residue and the substrate will be free to leave the ER (modified from Williams 2005).

The EDEM homolog in yeast (Htm1p/Mnl1p) also has been shown to accelerate ER-associated glycoprotein degradation. EDEM was shown to function as a chaperone and thus showed indications of maintenance of the retrotranslocation of misfolded alpha1-antitrypsin by inhibiting aggregation so that unstable misfolded proteins could be accommodated by the dislocon for ERAD (Hosokawa *et al*., 2006). Recently, Oda *et al*. (2006) found two

more transmembrane proteins in mammalian cells, Derlin-2 and Derlin-3, which are up-regulated by UPR stress and are targets of the IRE1-mediated ERAD regulation.

Peptidylprolyl isomerase, PPI

Peptidylprolyl isomerases (PPIs) comprise another family of chaperones involved in protein quality control. They are found in all cellular compartments where protein folding occurs. PPIs catalyse *cis/trans* isomerisation of the peptide bond and are induced by UPR. PPIs improve the efficiency of protein disulfide isomerase as a catalyst of protein folding (Schönbrunner and Schmid, 1992; Bose and Freedman, 1994).

Heat shock proteins, Hsps

Many proteins of the heat shock family are involved in the protein quality control under ER stress. Some of the proteins belonging to the family of 70 kDa heat shock proteins (Hsp70) are only expressed under stress conditions, while some are present in cells under normal growth conditions and are not heat-inducible. The Hsp70 can bind and release polypeptides in an ATP-dependent cycle to prevent protein aggregation and promote proper folding. Hsp70 binds to a wide range of newly synthesised polypeptides comprising about 15-20 % of total protein in the cell (Bukau and Horwich, 1998). The 40 kDa heat shock protein (Hsp40) regulates Hsp70 chaperone activity by delivering specific substrates to Hsp70 at the ER or mitochondrial membrane (Glover and Lindquist, 1998). Hsp40 and Hsp70 as a complex also recognise abnormal proteins and conjugate them to the ubiquitin ligase enzyme, leading to protein degradation by the 26S proteasome (Goldberg, 2003). The highly abundant ER chaperone, 90 kDa heat shock protein (Hsp90; Grp94 homologue) binds to proteins during folding and after the substrate proteins have been released from BiP. Hsp90 has ATPase activity and is presumed to regulate interaction with substrate polypeptides. Hsp90 can prevent aggregation of substrates and enhance both Hsp70- and Hsp40-mediated folding (Glover and Lindquist, 1998).

All components discussed above participate in ER quality control. Proteins are allowed to exit the ER and enter the secretory pathway only after they have been properly folded and modified into their biologically active conformation. The continuous interaction between protein folding, retrotranslocation machinery of the ER and the ubiquitin-proteasome

system in the cytosol is essential for the maintenance of healthy cellular activity and metabolism in both unicellular and multicellular organisms (Ellgaard *et al*., 1999; Haigh and Johnson, 2002; Kostova and Wolf, 2003).

Several groups have isolated genes encoding ER chaperones and foldases from filamentous fungi. *Bip* genes have been isolated from *Neurospora crassa* (Techel *et al*., 1998), *A. awamori* (Hijarrubia *et al*., 1997) and *T. reesei* (Saloheimo *et al*., 2003). Genes for *pdi* have been cloned from *T. reesei* (Saloheimo *et al*., 1999) and have similar features to their *S. cerevisiae* homologues. The gene *prpA* isolated from *A. niger var. awamori* does not seem to belong to any previously recognised PDI family (Wang and Ward, 2000). A homologue to the mammalian calnexin gene has been cloned from *A. niger,* and orthologues identified in *A. nidulans* (Wang *et al*., 2003). Proteins with PPI activity have been identified from *A. nidulans* (Joseph *et al*., 1999; Conesa *et al*., 2002; Derkx and Madrid, 2001), but no calreticulin is known from fungi.

1.1.2 Molecular mechanism of unfolded protein response in the endoplasmic reticulum

The ER provides an environment for glycoproteins to fold and form disulfide bonds; it also houses the mechanism to detect problems in folding. A cascade of events in the signalling pathway is triggered by the accumulation of misfolded proteins in the ER, sensed by a transmembrane kinase Ire1p localised in the ER/nuclear envelope (Fig. 4). As a result, BiP is released from Ire1p (Urano *et al*., 2000).

Release of BiP molecules from the lumenal site of Ire1p results in the dimerisation of Ire1p. This conformational change (dimerisation) transmits a signal across the membrane and triggers cytoplasmic kinase activity. The kinase induces a splicing event in the *HAC1* mRNA (encodes a basic leucine-zipper [bZIP]-containing transcription factor), allowing synthesis of the transcription factor, Hac1p, which up-regulates the transcription of all genes containing an UPR Element (UPRE). This cascade of events leads to an increase in the levels of proteins required for folding and protein quality control (Fig. 5; Kawahara *et al*., 1997; Virgilio *et al*., 1999; Bertolotti *et al*., 2000; Travers *et al.*, 2000; Sidrauski *et al.*, 2002; Ogawa *et al.*, 2004). Induction of 381 genes has been shown to be dependent on the

IRE1 and *HAC1* pathway in *S. cerevisiae.* These include genes involved in secretion related processes, such as protein folding and disulfide bond formation, ER-associated degradation, ubiquitin/proteasome pathway and vesicle trafficking/transport (Travers *et al*., 2000).

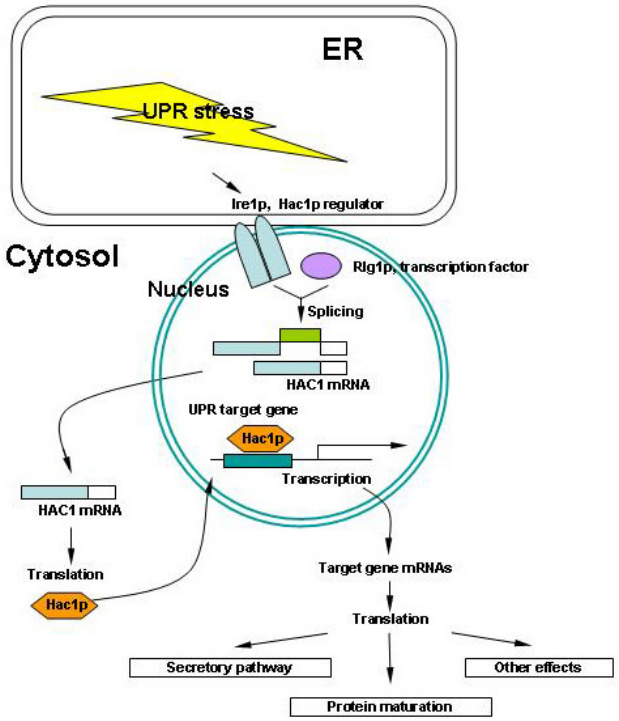

Figure 4. Ire1p-mediated regulation of the unfolded protein response in yeast. After BiP is released from the Ire1p lumenal domain, Ire1p dimerises and together with Rlg1p induces the splicing of HAC mRNA. This splicing leads to the translation of HAC1 mRNA. After translation, Hac1p migrates into the nucleus and binds upstream of the UPR element, thereby triggering the transcription of target genes (Adapted from McCracken and Brodsky, 2000).

Regulation of the UPR pathway by Ire1p seems to be very similar in different species of filamentous fungi. For example, functional homologues of the *S. cerevisiae HAC1* have been found in *T. reesei* (Hac1p; Saloheimo *et al*., 2003) and *A. nidulans* (HacA; Ngiam *et al*., 1997). However, the process of mRNA splicing is different from *S. cerevisiae* (see Fig. 5). Induction of *hac1* in *T. reesei* first involves the splicing of an intron of 20 bp and secondly, truncation of the 5□upstream region of the mRNA (Saloheimo *et al*. 2003). Removal of the upstream open reading frame region from the mRNA probably enhances translation initiation (Valkonen, 2003).

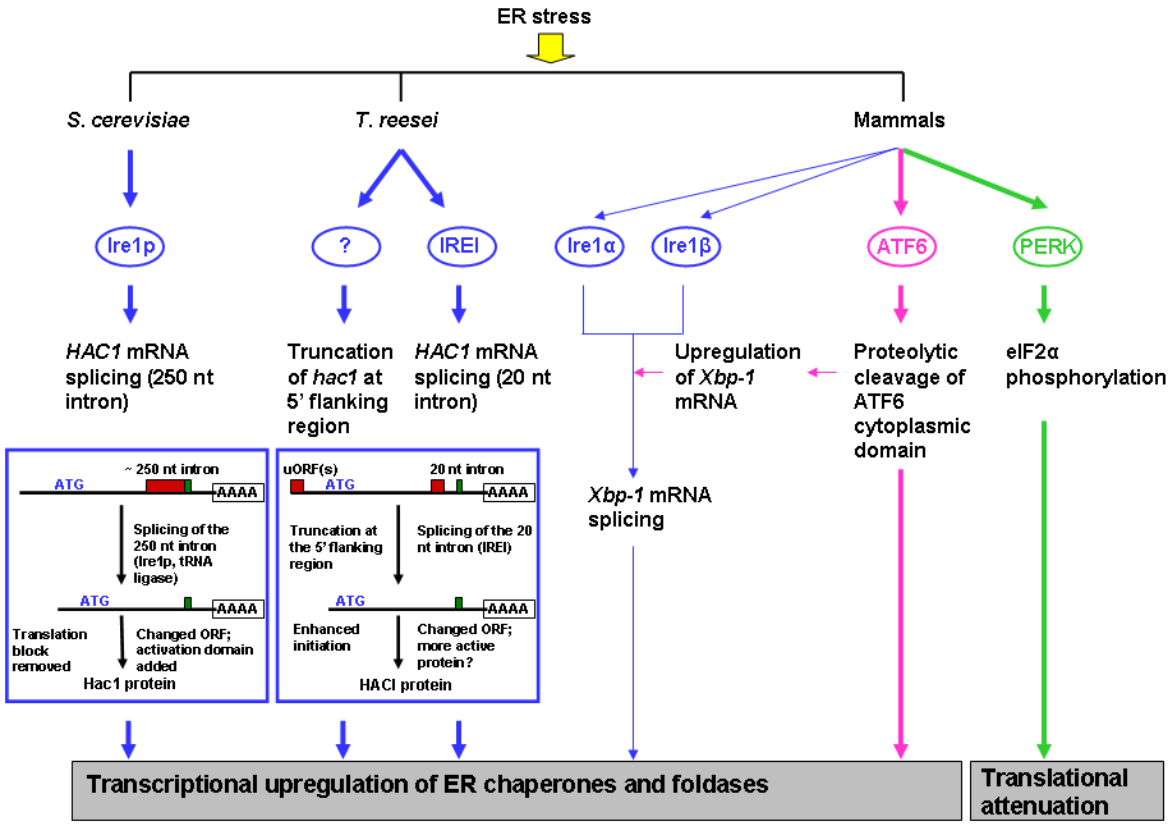

Figure 5. Comparison of induction of the HAC1 coding gene in *T. reesei*, *S. cerevisiae* and mammalian cells. Activated Ire1p promotes splicing of HAC1 mRNA by a different mechanism in *S. cerevisiae* (left) and filamentous fungi (middle) and then HAC is actively translated and capable of activating its downstream genes. In mammalian cells, in addition to IRE1-regulated induction, there are two other regulatory systems, ATF6-and PERK- mediated regulation. The width of arrows indicates the most prominent of the pathways in activating UPR (modified from Valkonen, 2003).

In addition to IRE1 mediated regulation of UPR in mammalian cells, there are three other regulatory pathways which function in three distinct ways: 1) X-box DNA-binding protein-mediated (XBP1; Yoshida *et al*., 2001; Zhang and Kaufman, 2004, Ma and Hendershot, 2001), 2) activating transcription factor 6-mediated (ATF6; Yoshida *et al.* 2001; DenBoer *et al.*, 2005) and 3) protein kinase-like endoplasmic reticulum kinase-mediated pathways (PERK; Bertoletti *et al*., 2000; Harding *et al*., 2000). Homologues for PERK or ATF6 have not yet been identified in fungi (Guillemette *et al*., 2007). However, there is indirect evidence that such homologues may exist based on the finding of an *IRE1* -independent pathway in *S. cerevisiae,* induced by stress. This *IRE1*-independent pathway in yeast cells could be homologous for mammalian PERK or ATF6-mediated induction of UPR (Leber, *et al*., 2004).

As part of protein quality control, the proteins that fail to fold to their proper conformational structure are retained within the ER and degraded to prevent the accumulation of aberrant proteins that would otherwise impair the vital functions of the cell (Caldwell *et al*., 2001). In ERAD, misfolded proteins are removed from the ER by retrotranslocation into the cytosol where they are degraded by the ubiquitin proteasome system (Knop *et al*., 1996; Scott and Schekman, 2008).

1.1.3 ER-associated protein degradation (ERAD)

Misfolded proteins from the ER are targeted to the cytoplasm for proteasomal degradation by the ubiquitin-proteasome system. In *S. cerevisiae* the key proteins in this process are ER membrane-bound ubiquitin proteins Hrd1p/Der3p and Hrd3p (shown in Fig. 3). Hrd1p/Der3p has been found to mediate ubiquitination of misfolded ER proteins. Ubiquitination is a compulsory covalent modification of the substrate for degradation and it is coupled to retrotranslocon of the protein. Hrd1p/Der3p, the multi-spanning ER-membrane protein, that has a cytosolically-positioned RING-H2 domain, which binds to ubiquitin1 (Ubc1p) or ubiquitin 7 (Ubc7p), mediating the covalent attachment between the ubiquitin and target substrate. The localisation of Ubc7p and Hrd1p/Der3p to the ER membrane and their involvement in ubiquitination suggest that they have a role in the retrograde transport process together with the Sec61 complex (Plemper *et al*., 1999; Mayer *et al*., 1998; Gauss *et al*., 2006).

The Sec61 import channel also may be the export channel (Brodsky and McCracken, 1999; Kostova and Wolf, 2002). The retrotranslocation channel differs in its active components from the import channel. The Sec61 translocon is composed of three subunits: pore forming α subunit, and the β and γ subunits (Sec61p, Sbh1p and Sss1p in yeast). In *S. cerevisiae* the targeting of the signal-sequence-bearing proteins to the Sec61 complex is achieved either co- or post-translationally (Hann and Walter, 1991; Knight and High, 1998; Brodsky *et al*., 1999). The signal sequence seems to play a critical role in the choices between co- and post-translocation. Once the signal sequence of the target protein is recognised by the signal recognition particle (SRP), they form a complex which is bound to the Sec61 component. This complex assembly activates a translocation channel and allows the polypeptide to be transferred through to the ER lumen. The BiP protein is suggested to

have an active role in translocation by providing a pulling force to the polypeptide moving across the ER membrane (Glick, 1995; Wang and Johnsson, 2005).

Post-translational protein translocation occurs after the polypeptide has been fully translated and does not require SRP recognition. The only requirement is that the Sec62□ Sec63-Sec71p-Sec72p protein complex recognises the target protein and dissociates to Sec61 complex to form a functional unit together with the assistance of the lumenal ATPase BiP, which finalises the translocation requiring ATP (Brodsky *et al*., 1993; Müsch *et al*., 1992; Ng *et al*., 1996; Plath *et al*., 2004; Wang and Johnsson, 2005). The presence of two subsets of translocons can explain how this complicated two-way traffic across the membrane is regulated (Plemper and Wolf, 1999; Friedlander *et al*., 2000).

1.1.4 Ubiquitin-proteasome system

Nearly every protein in a cell is a potential substrate for the ubiquitin system. Most of the transcriptional regulatory proteins in the cell are relatively short-lived and targeted for proteolytic degradation. In addition, proteins which reside or travel through the ER can be targeted for ubiquitination and degradation. Misfolded proteins in the cytoplasm and ER need to be tagged with ubiquitin for proteasomal degradation (Hampton, 2002; McDonough and Patterson, 2003; Trombetta and Parodi, 2003).

Ubiquitin

Ubiquitin (Ub) is a small (8 kDa) protein composed of 76 amino acids. This protein is found only in eukaryotic organisms. Ubiquitin is highly conserved among eukaryotes. Prokaryotes have no molecule analogous to ubiquitin, but there are some proteins that share structural and functional similarities with ubiquitin, including the sulphur transferases ThiS and MoaD (Hochstrasser, 2000; Wang *et al*., 2001).

Ub is found throughout the cell and can exist either in a free form or as part of a complex with other proteins. It conjugates to proteins through an isopeptide bond between lysine (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 or Lys63) and the ε-amino group of lysine side chains on the substrate proteins (Zhou and DeWille, 2007). The first ubiquitin chains identified were those linked via lysine 48. However, more recent work has uncovered a

wide variety of linkages involving all possible lysine residues which affect the target location of the tagged protein. (Kirisako *et al*., 2006; Xu and Peng, 2008). Single Ub molecules can be conjugated through the lysine residue to the target protein and serve other functions such as activation of transcriptional regulators or routing of proteins to the lysosome/vacuole. Polyubiquitin chains are formed when isopeptide linkages between the C - terminus of ubiquitin and the ε -amino group of lysine to other ubiquitin molecules are formed (Pickart, 1997). Successive addition of activated ubiquitin molecules to internal Lys48 of the previously conjugated molecule leads to generation of the polyubiquitin chain that is the degradation signal recognised by the 26S proteasome. Ubiquitination is a process that is dependent on ATP hydrolysis (Wilkinson, 2000; Babbitt *et al*., 2005).

The location of the lysine residue at which the ubiquitination occurs is important and the different lysine residue binding combinations provides a versatile post-translational signaling system to the cell. Mono-ubiquitination regulates the activity of the substrate protein, offering a signaling function in diverse processes, such as endocytosis, protein sorting, subnuclear trafficking and gene expression. Through polyubiquitination, different ubiquitin chains can be formed depending on the lysine linkage used between the single Ub molecule and substrate protein. For example, the formation of polyubiquitin chains by linkage at Lys-48 and Lys-29 of ubiquitin can act as a signal for proteasomal degradation, whereas ubiquitination at other lysine residues (e.g. Lys 63) act as a signal for non-proteolytic functions in the DNA repair pathways, activation of transcription factors, endocytosis and ribosomal functions (Weissman, 2001; Pickart and Fushman, 2004; Kuhlbrodt *et al*., 2005). In higher eukaryotes, ubiquitin is one of a group of ten protein homologues, ubiquitin-like proteins (Ubls) such as SUMO, Nedd8, UCRP, FAT10, HAB, ISGN15, Apg8, Apg12, URM1, and AN1 (Jensch and Pyrowolakis, 2000; Schwartz and Hochstrasser, 2003).

Very little research has been carried out with ubiquitin-like proteins in filamentous fungi. Recently, Wong *et al.* (2008) have characterised a SUMO-encoding gene from *A. nidulans.* Conjugation of SUMO to the target proteins encompasses several cellular activities including cell-cycle control, nuclear transport and cell responses to viral infections. It is thought that sumoylation might hinder ubiquitination, because in some cases ubiquitination

of a protein leads to its degradation whereas sumoylation stabilises the protein (Welchman *et al*., 2005).

Ubiquitination is the mechanism that regulates protein turnover in a cell. By regulating protein degradation, cells can quickly eliminate a protein that in turn regulates another function e.g. a transcription factor that is needed for expression of a particular gene (Herschko and Ciechanover, 1998). During the past decade a surprising number of other ubiquitin-like modifiers has been discovered, which all look like ubiquitin in their mechanisms of substrate conjugation. By tagging proteins with ubiquitin or ubiquitin-like proteins, the cell can create a diverse assortment of modified proteins that can be identified by downstream protein receptors or interactors and used to control many regulatory pathways in the cell (Hochstrasser, 2000; Schwartz and Hochstrasser, 2003).

The ubiquitin-proteasome system (UPS) is highly conserved in all eukaryotes. UPS is a cascade of enzymatic reactions (shown in Fig. 6) that can be separated into three steps. First, the single conserved ubiquitin-activating enzyme, E1, hydrolyses ATP and forms a high-energy thioester linkage between its active site cysteine and the carboxy-terminus of ubiquitin (Ub; Attaix *et al.*, 2002). In the second step, ubiquitin is transferred from E1 to a member of the family of ubiquitin-conjugating enzymes, E2 or Ubc. In the third step, the E2 enzymes together with ubiquitin protein ligases or E3s attach ubiquitin to lysine residues from the substrate protein. In most cases, E3 functions to align the substrate and E2 in a way that facilitates ubiquitination (Oka *et al*., 2004).

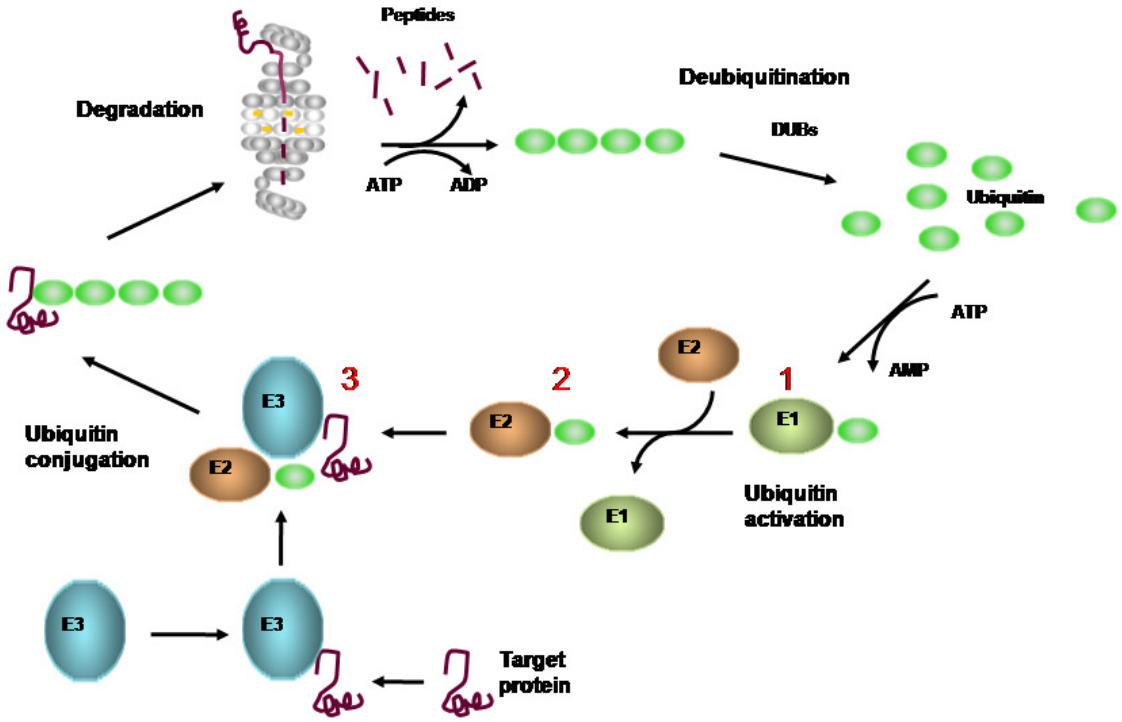

Figure 6. The ubiquitin proteasome system (UPS). A cascade of enzymatic reactions leads to ubiquitination of lysine residues of the substrate. First, ubiquitin is activated by E1 (1) and transferred to the E2 enzyme (2) and finally conjugated to substrate proteins (3) with a specific E3 ligase. Proteasomal degradation requires poly-ubiquitination of the substrate. Deubiquitin enzymes (DUBs) cleave ubiquitin from the substrate and from each other (Adams, 2003; Kostova and Wolf, 2003).

Eleven E2s have been found in yeasts (Pickart, 2001) and 20-30 E2s in mammals (Scheffner *et al*., 1998). Despite being structurally similar, E2s are responsible for many biological functions such as DNA repair and the degradation of cyclin as a part of the cell-cycle (Pickart, 2001; Welchman *et al*., 2005; Ho *et al*., 2006). In yeast, only three play a role in the formation of the polyubiquitin degradation signal. The other yeast E2s are involved in the formation of mono-, di-, and triubiquitin conjugates which are not targeted for degradation (Pickart, 2001; Gerards *et al*., 1998).

E3s are responsible for the selective recognition of protein substrates. There are two main types of E3 ligases, HECT-domain type (**H**omology to **E**6AP **C** terminus) and RING-domain type (contain the sequence motif with a characteristic spacing of cysteine (C) and histidine (H) residues). Both utilise different mechanisms to transfer Ub onto substrates (Ho *et al*., 2006; Pickart, 2001). Twenty different human HECT E3s which interact with two classes of E2 have been described (Schwarz *et al*., 1998), but in the mammalian genome

numerous potential uncharacterised HECT E3s have been identified by sequencing projects (Pickart, 2001; Pickart and Eddins, 2004; Ho *et al*., 2006).

So far, no genes encoding ubiquitin, ubiquitin activating or conjugating enzymes or ubiquitin ligases have been isolated from *T. reesei*. However, in the current study, eight *S. cerevisiae* ubiquitin homologues were found in the *T. reesei* genome indicating that similar genes are present in filamentous fungi.

The substrate must be conjugated to a polyubiquitin chain that comprises at least four ubiquitin moieties in order to be efficiently degraded by the proteasome. A family of conjugating factors, called E4s, regulates elongation of the polyubiquitin chain. The first E4 to be described was the yeast protein UFD2. The other protein discovered that is an E4 family member is called CHIP (**C**-terminus of the **H**sp70-**i**nteracting **p**rotein), which has been shown to regulate the ubiquitination activity of the E3 parkin gene. The mutations in the *parkin* gene have been found to cause a juvenile Parkinson disease (Kahle *et al*., 2000). A polyubiquitin chain is formed on the substrate by successive addition of ubiquitin molecules to lysine residues of the previously attached ubiquitin. Polyubiquitin chains can be formed also with isopeptide linkages involving different ubiquitin lysines (Pickart, 1997; Meusser *et al*., 2005). The formation of the polyubiquitin chain can be catalysed by E1, one E2, and one E3. The protein substrate can be ubiquitinated by different combinations of E2 and E3. E2 can interact with a limited number of E3 proteins, which in turn recognise their specific protein substrates (Pickart, 2001; Gerards *et al*., 1998).

Deubiquitination

Ubiquitination seems to be a reversible process. Specific hydrolytic enzymes, deubiquitinases (DUBs; **d**e**ub**iquitinating enzymes) can remove the ubiquitin molecules. Based on sequence similarity, deubiquitinating enzymes may be classified into two families: ubiquitin-specific proteases (UBPs or USPs) and ubiquitin-carboxy-terminal hydrolases (UCHs). Some of the DUBs are associated with the 19S cap of the 26S proteasome, cleaving ubiquitin from the substrates before degradation (Botchler *et al*., 1999; Meusser *et al*., 2005; Kim *et al*., 2003; Wing, 2003; Attaix *et al*., 2002). In addition to the role in ubiquitin-proteasome system, DUBs participate in post translational protein modifications. For example, Amerik *et al*. (2000, 2006) found the deubiquitinating enzyme

Doa4 was also involved in the endocytosis and vacuolar targeting of proteins in *S. cerevisiae*.

1.1.5 Tagging proteins for degradation in the ubiquitin-proteasome pathway

Proteins can contain multiple forms of the signal that is recognised by the ubiquitin proteasome machinery. Pickart (1997) refers to three types of signals. First, a correlation between the half-life of a protein and the nature of the N-terminal residue can be a signal for the ubiquitin-proteasome degradation pathway. Second, certain amino acid sequences are signals for degradation. One such sequence is known as a PEST sequence (in which a short stretch of about eight amino acids is **e**nriched with **p**roline, glutamic acid, **s**erine, and **t**hreonine; Rogers *et al*., 1986; Wilkinson, 2000). Third, the signals may also be in the hydrophobic core of the protein (Hershko and Ciechanover, 1998; Pickart, 2001). When proteins exist in their native state the signals are hidden and the protein can escape degradation. However, in a partially unfolded state, the signals may be seen by the ubiquitin machinery, causing the protein to become tagged by Ub and degraded by the proteasome (Thrower *et al*., 2000).

Substrate proteins without Ub interact with the proteasome but quickly dissociate. Proteins containing a single Ub dissociate from the proteasome more slowly while the substrate protein with Ub-chains is thought to interact with a proteasome for a longer period of time, thus increasing the probability that the proteasome will degrade it (Benaroudj *et al*., 2001). Proteins tagged with short Ub chains are poor substrates for the proteasome, thus the short Ub chain is cleaved off by deubiquitinating enzymes, resulting in premature release of the target protein from the proteasome (Eytan *et al*., 1993; Lam *et al*., 1997; Thrower *et al*., 2000; Guterman and Glickman, 2004).

1.2 26S proteasome, a molecular machine designed for controlled proteolysis

Proteasomes are large multisubunit proteases that are found in the cytosol and are sometimes attached to the ER. Proteolytically-active proteasomes are also located in the nucleus of eukaryotic cells (Rockel *et al*., 2005; Scharf *et al*., 2006). The 26S proteasome is an ATP- dependent proteolytic system which functions in the degradation of short-lived proteins under normal metabolic conditions as well as the degradation of long-lived proteins. Studies by Rock *et al.* (1994) have shown that over 30 % of all proteins at any given moment produced in the cell are degraded by the 26S proteasome.

1.2.1 Structure of the 26S proteasome

The 26S proteasome has a molecular mass of approximately 2.5 Million Daltons (MDa) and is composed of at least 32 different subunits. It is formed by the assembly of a 20S proteolytic complex, called the core particle (CP) and two 19S regulatory particles (RP) known as the □cap□ The 19S regulatory complexes are attached at both ends of the 20S proteasome (Fig. 7; Glickman and Adir, 2004).

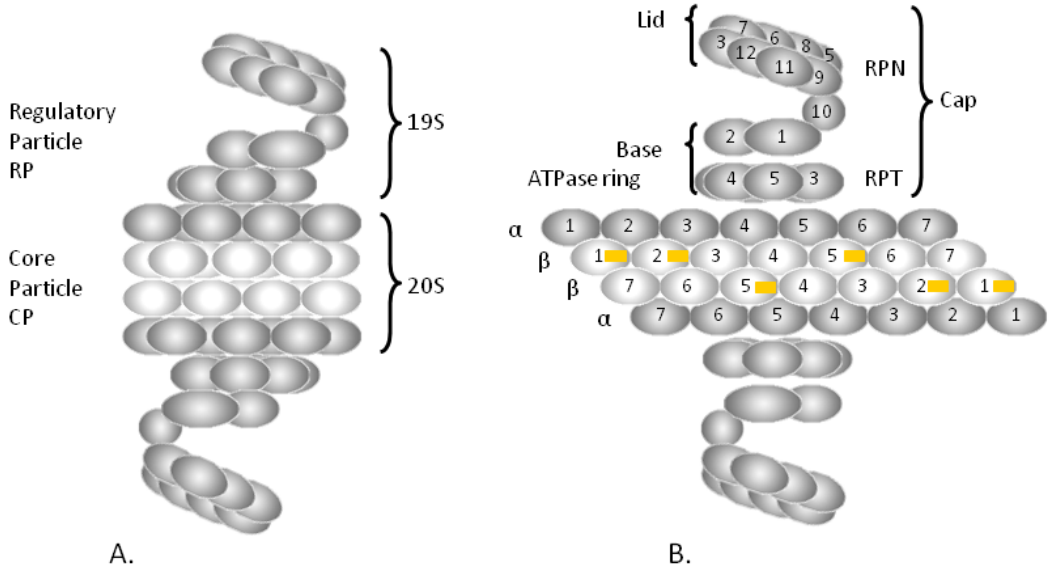

Figure 7. The assembly and structure of the 26S proteasome. A: The 20S core particle (CP) assembles with two 19S regulatory particles (RP) attached at either end of the CP. B: Seven α-subunits make up the two outer layers and seven β-subunits the two inner rings in the four-layer 20S cylinder structure. The yellow colored β subunits have protease activities: β1-subunits peptidylglutamylpeptide hydrolyzing (PGPH) activity, β2 subunits trypsin-like (TL) activity and β5-subunits chymotrypsin-like (ChTL) activity (modified from Wolf and Hilt, 2004).

The 20S proteasome core particle forms a barrel, measuring 15 nm in length and 11 nm in diameter (Bochtler *et al*., 1999). The complex has a molecular mass of 700 kDa and contains 28 subunits in both prokaryotes and eukaryotes, although the complexity differs. The cylindrical structure, as first demonstrated by immuno-electron microscopy (Zwickl *et al*., 1992), is formed by four heptameric rings, each containing seven subunits (Fig. 7). Löwe *et al*. (1995) obtained the first X-ray crystal structure of the 20S proteasome from *Thermoplasma acidophilium* and showed the architecture of the CP at 3.4Å resolution. The two outer layers of the 20S proteasome consist of seven alpha subunits while the two central rings are composed of seven beta subunits (Coux *et al*., 1996; Yerlikaya, 2004).

In the centre of the multicatalytic protease chamber, there are three subunits performing the proteolytic function. Subunit β1 has peptidyl-glutamyl peptide-hydrolysing activity (PGPH; cleavage after acidic amino acids), subunit β2 has trypsin-like activity (TL; cleavage after basic amino acids) and subunit β5 has chymotrypsin-like activity (ChTL, cleavage after hydrophobic amino acids; Kopp *et al*., 1997; Heinemeyer *et al*., 1997; Rechsteiner, 1998;

Staszczak 2002; Heinemeyer *et al*., 2004; Nandi *et al*., 2006). These proteolytic activities can be assayed with fluorogenic substrates. Two additional peptide hydrolysing activities are also known, the first cleaving after branched chain amino acids (BrAAP, β1 and β5) and the second after small neutral amino acids (SNAAP, β5) (Groll *et al*., 1999; Nandi *et al.*, 2006). How the proteasome chooses to cleave the substrates into the peptides is unclear, but it has been shown that the proteasome can cleave substrates at almost every peptide bond (reviewed in Groll and Huber, 2003). Site-directed mutagenesis studies with all β-subunits of *Thermoplasma* showed that the N-terminal the threonine residue (Thr1) was essential for autolytic and proteolytic functions (Seemüller *et al.*, 1995). Peptides generated can vary from 3 - 25 amino acids in length suggesting that proteasomes process peptides until they are small enough to diffuse out of the proteolytic chamber (Nandi *et al*., 2006).

To activate the 20S proteasome, the ATP-dependent association of 19S regulatory particle is necessary and induces the opening of alpha rings in the 20S complex, allowing the substrate to enter into the proteolytic chamber (Groll *et al*., 2000). The exact number of 19S particle subunits in different organisms is unknown. In *S. cerevisiae*, 18 subunits have been identified ranging from 15 to 112 kDa in molecular mass (Glickman *et al*., 1998). The 19S RP associates with the 20S proteasome in an ATP-dependent reaction to form an active 26S proteasome particle providing substrate specificity and mediating ATP-dependence. The 19S RP is involved in the recognition, binding and unfolding of ubiquitylated proteins and in regulation of the opening of the 20S core for protein degradation. The core RP can further dissociate into lid and base, together called a cap (Fig. 7).

The base of the 19S RP consists of a ring of six ATPases (Rpt1 Rpt6), which dock onto the α-rings of the 20S core, and three non-ATPase subunits (Rpn1, Rpn2 and Rpn10; Adams *et al.*, 1997; Voges *et al.*, 1999). All six ATPase subunits belong to a larger family termed AAA (**A**TPases **A**ssociated with a variety of cellular **A**ctivities)-ATPases or CADs (**C**onserved **A**TPase **D**omain), a very diverse family of ATPases with members involved in vesicle fusion, peroxisome assembly, and transcriptional activation (Bochtler *et al*., 1999). Specific functions of the 19S ATPase subunits involve binding to substrate and unfolding it prior to leading it into the cavity of the 20S core particle for degradation (Braun *et al*., 1999). For other ATPases, Rpt5 binds to the ubiquitinated substrate (Lam *et al*., 2002), and Rpt2 is believed to control both substrate entry and product release from the 20S channel

(Köhler *et al*., 2001). Rpn1 interacts with Rad23 and Dsk2, two proteins harboring ubiquitin-like domains (Ubl) and is capable of binding and delivering ubiquitinated substrates to the proteasome. Elsässer *et al*. (2002; 2004) suggested that Rpn10 also contributed to the binding of ubiquitin chains.

The lid complex forms the distal mass of the 19S RP and is necessary for degradation of ubiquitinated target proteins. It is comprised of eight subunits: Rpn3, Rpn5, Rpn6, Rpn7, Rpn8, Rpn9, Rpn11 and Rpn12 (Fig. 7; Glickman *et al*., 1998; Saeki *et al*., 2000). Five of these subunits, Rpn3, Rpn5, Rpn6, Rpn7 and Rpn9, contain a PINT/PCI domain at their terminal ends. Subunits of this type are found in other large complexes in the cell involved in diverse functions such as the COP9-signalosome complex and eukaryotic initation factor 3, eIF3. Sequence similarity has been found between these complexes and the 19S lid particle (Wei and Deng, 1998; Seeger *et al*., 1998; Asano *et al*., 1997). Rpn8 and Rpn11 contain a highly conserved metalloisopeptidase domain (MPN) and this activity is necessary for deubiquitination and thus proteasomal proteolysis of substrates. Currently, it is believed that Rpn11 is the deubiquitinating enzyme that separates the substrate from the polyubiquitin chain and allows it to unfold and translocate into the 20S proteasome channel (Glickman *et al*., 1999; Verma *et al*., 2002; Yao and Cohen, 2002; Glickman and Adir, 2004). This movement is irreversible and leads to proteolysis of the substrate. Following release from the substrate, the polyubiquitin chain is hydrolysed into single ubiquitin moieties which can take part in a new round of protein degradation. The Rpn12 subunit, known to be essential for growth in yeast cells, interacts with the other RP subunit, Rpt1. A combination of mutations in both Rpn12 and Rpt1 was found lethal to the cell (Takeuchi and Toh-e, 1999a). Rpn4 is known to act as a transcriptional regulator not only for proteasomal genes but also numerous genes involved in different cellular processes which have the same motif in their promoter regions. The Rpn4 is known to be positive and negative transcriptional regulator of the ubiquitin-proteasome system The Rpn4 binding site, PACE (**p**roteasome **a**ssociated **c**ontrol **e**lement) region is 9-bp motif in the promoters of proteasome genes (Wang *et al*., 2000; Karpov *et al*., 2008).

The first subunit isolated from 26S proteasome from *T. reesei, prs12,* has been cloned by Goller *et al.* (1998). The *prs12* genome sequence and cDNA cloned sequence from *T. reesei* were shown to have a high identity to the homologous counterparts from humans,

mouse, *Drosophila* and *S. cerevisiae*, and therefore, was predicted to be a fungal homologue of Rpn12 of 26S proteasome. In addition, during this work, our homology searches showed that the *T. reesei* genome contains as many proteasome genes as *S. cerevisiae*. Therefore, a similar subunit composition may be expected.

1.2.2 Assembly of the 26S proteasome

Assembly of 20S

The assembly of eukaryotic proteasomes is not fully understood. However, it requires a highly organised process to ensure proper positioning of each of the 14 different subunits. Several groups have reported that proteasome assembly depends on additional factors such as molecular chaperones that are only transiently associated with the 26S complex (Schmidtke *et al*., 1996; Schmitdt *et al*., 1997; Ramos *et al*., 1998). The *in vitro* assembly of the bacterial *Rhodococcus* proteasome has been used as a model of the assembly pathway (Fig. 8). When the α- and β-subunit precursors are allowed to interact, they first form an α/β heterodimer (I in Fig. 8). These heterodimers assemble further into half□ proteasomes built from the α- and β-subunit precursors. The assembly then proceeds via dimerisation of half-proteasomes (II) and is completed by the final conversion of preholoproteosomes (III and IV). The last step involves cleavage of the propeptide, which is a rate-limiting step in the assembly pathway (Schmidtke *et al.*, 1997; Baumeister *et al*., 1998). In the case of *S. cerevisiae,* it was found that the Ump1 (ub-mediated proteolysis) protein was required for maturation of 20S proteasomes and was degraded after completion of assembly (Ramos *et al*., 1998). Hirano *et al.* (2005) have identified a proteasome assembling complex (Pac) consisting of two chaperones, Pac1and Pac2, which binds to α-subunits and ensures proper formation and stability of the ring of α subunits. Therefore, the Ump1 and the Pac complex play distinct roles during proteasome assembly.

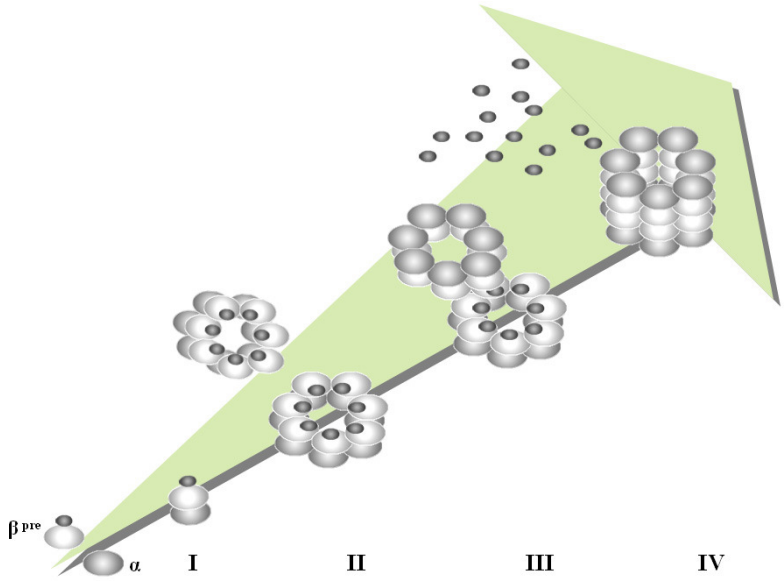

Figure 8. Assembly pathway of the bacterial *Rhodococcus* proteasome. The first step is the formation of α/β heterodimers with the help of propeptide (I), followed by the assembly of half proteasomes (II). They dimerise to form preholoproteasomes (III), which are converted into holoproteasomes (IV) through autocatalytic removal of the propeptide (modified from Baumeister *et al*., 1998).

Assembly of 19S

So far, very little is known about the mechanism of the 19S particle assembly or the factors that facilitate the process. The exact stoichiometry of subunits in the 19S is also unknown. It is more heterogeneous than the 20S core particle, and therefore it requires multiple chaperones for assembly. Many proteins have been shown to interact with the proteasome in affinity-purified protein samples (Verma *et al*., 2000) or by high-throughput proteomics (Querrero *et al*., 2006), which reflects the dynamic nature of the 26S proteasome complex. Several of the interacting proteins have been shown to be chaperones, such as Hsp90. Inactivation of Hsp90 causes dissociation of the 26S proteasome into the 19S base, 19S lid and 20S core particles. Hsp90 has been shown to interact with the 19S regulatory particle and therefore, it is suggested to play a role also in 19S proteasome assembly (Imai *et al*., 2003).

Two other chaperones are presumed to facilitate 26S proteasome assembly: Blm10 and Nob1. Blm10 has been found attached to the α-ring of the 20S particle. However, in the mutant 20S, Blm10 was present in high amounts on the surface which lead to inefficient assembly of the 26S proteasome. Nob1 is the other protein that may assist in 26S proteasome assembly and it is abundant in growing cells. Nob1 forms a complex with the

19S RP and is degraded just after the active 26S proteasome is formed (Tone *et al*., 2000). The 26S proteasome has been shown to dissociate into discrete 20S CP and 19S RP components *in vitro*, and to reassociate in an energy-dependent manner. The ratio of free 20S CP to 26S proteasome holoenzyme has been found to vary in eukaryotic cells responding to cellular conditions, suggesting the possibility of dissociation and reassociation *in vivo* (Kurucz *et al*., 2002; Rosenzweig and Glickman, 2008). Although the 19S particle is well defined and the role of the subunits in the substrate degradation process is known, an arrangement and interactive model of the subunits and the distinctive function of each subunit is still to be proposed.

1.2.3 Diverse functions of the 26S proteasome

Proteasomes were thought initially to be assembled only for degradation of damaged or misfolded proteins but over the last decade they have been found to be involved in the control of cell proliferation, the cell cycle, activation of transcriptional repressors and activators, in multiple aspects of gene expression, apoptosis and other cellular processes (Glickman *et al*., 1999). Table 1 shows examples of the diverse functions of proteasomes described in the literature. The importance of proteolytic degradation inside cells and the role of ubiquitin in proteolytic pathways was acknowledged in the award of the 2004 Nobel Prize in Chemistry to Aaron Ciechanover, Avram Hershko and Irwin Rose (http://nobelprize.org/nobel prizes/chemistry/laureates/2004).

Table 1. Examples of diverse functions of the 26S proteasome beyond misfolded protein degradation.

Function of the 26S proteasome	***Organism***	***Mechanism***	***Reference***
Activation of transcriptional repressors and activators (transcriptional activator Gal4). Transcription factor GCN4 destruction by ubiquitin (Ub) □ mediated proteolysis.	*S. cerevisiae*	Degradation	Lee *et al*., 2005 Kornitzer *et al*., 1994 Tansey, 2001
Stimulation of transcription. Repression of gene transcription (MATα2).	*S. cerevisiae*	Degradation	Johnsson *et al.*, 1998 Laney and Hochstrasser, 2003
DNA repair and RNA polymerase II interaction.	*S. cerevisiae*	Degradation	Krogan *et al*., 2004 Beaudenou *et al*., 1999
RNA polymerase II interaction.	*S. cerevisiae*	Degradation	Reid and Svejstrup, 2004
26S subunits binding to DNA, CDC20 and chromatin.	*S. cerevisiae*	Unclear	Morris *et al*., 2003 Muratani and Tansey, 2003 Ezhkova and Tansey, 2004 Aquilar and Wendland, 2003
Interaction with nucleotide extension repair protein Sec23.	*S. cerevisiae*	Interaction	Baker and Grant, 2005 Ferdous *et al*., 2001
Interaction with transactivators, transcription elongation factor, CDC68.	*S. cerevisiae*	Interaction	Gillette *et al.*, 2004 Archer *et al*., 2008
Cyclins, cell cycle control (CDC6).	Human, *S. cerevisiae* Rat hepatocyte	Degradation	Hershko and Chiechanover, 1998 Voges *et al*., 1999 Elsässer *et al*., 1999 Perkin *et al*., 2001
Liver alcohol dehydrogenase.	Rat hepatocyte	Degradation	Mezey *et al*., 2001
Apoptosis.	*S. cerevisiae*	Interaction	Madeo *et al*., 1997
Inactive precursor proteins processed to the active form.	Human	Limited proteolysis	Hilt and Wolf, 1999 Palombella *et al*., 1994

1.2.4 Proteasome-interacting proteins

While protein ubiquitination and proteasome-mediated protein degradation are essential to the regulation of many important biological processes, the number of ubiquitinated substrates is enormous. The proteins that recognise the ubiquitinated substrates and translocate them to the 26S proteasome for degradation are called polyubiquitin receptors or proteasome interacting proteins (PIPs). Proteins, such as Rad23 and Dsk2, with ubiquitin-like (UBL) or ubiquitin associated (UBA) domains are known to function as polyubiquitin receptors (Saeki *et al*., 2002). Some more recently found ubiquitin receptors provide a view into the wide range of different proteins involved in the ubiquitin proteasome system. Hsp90 is known to interact with the 26S proteasome but also assists in its assembly in an ATP-dependent manner (Imai *et al*., 2003). The COP9 signalosome is known to participate in diverse cellular and developmental processes and it shares two similar domains with the 26S proteasome, PCI/PINT (**P**roteasome, **C**OP9, **I**nitiation factor) and the MPN (**M**pr1[acetyl transferase]-**P**ad1[Phenylacrylic acid decarboxylase] **N**-terminal). These two modules are associated specifically with multiprotein complexes. PCI/PINT is believed to be incorporated into the subunits in both the proteasome and COP9 complexes. It is suggested also to be involved in mediating protein-protein interactions (Hofmann and Bucher, 1998; Ciccarelli *et al.*, 2003). A controlling role over ubiquitin proteasome-mediated protein degradation has been suggested for the COP9 signalosome due to its involvement in de-ubiquitination (Wei and Deng, 2003).

Guerrero *et al.* (2006) identified 64 proteasome-interacting proteins using a QTAX (Quantitative analysis of tandem-affinity purified cross-linked protein complexes) method, which involves two-step affinity purification under fully denaturing conditions using histidine-biotine-histidine (HBH)-tagged proteasome subunits (Rpn1, Rpn5, Rpn10 and Rpn11). Analysis of purified proteins was performed by LC MS/MS and PiPs were identified based on their stable isotope labelling of amino acids in cell culture (SILAC) ratios. In addition the authors extended the network analysis to categorise the proteins identified to their association to gene ontology (GO). This analysis revealed eight functional groups of PIPs, such as chromatin modeling, tRNA aminoacylation, metabolism, transport, translation, DNA replication, endosytosis and protein folding. Of these, 42 were not known previously function as ubiquitin receptors. Recent research by Guerrero *et al.*

(2008) using the QTAX method provides the most comprehensive report of the proteasome-interacting network of *S. cerevisiae*, which is composed at least 471 proteins. The use of different tags gave different PIP identifications suggesting a subunit-specific interaction (Guerrero *et al*., 2008). Since the known PIPs represent only a small subclass of substrates, additional ubiquitin receptors targeting various classes of ubiquitinated substrates still remain to be found.

1.2.5 Proteasome inhibition

Proteasome inhibitors can be categorised into two groups, synthetic analogs and natural products. Synthetic inhibitors are peptide-based compounds with diverse pharmacophores that react with the active site of the proteasome. Synthetic inhibitor analogs have been created by combining a peptide moiety with a reactivity group, an aldehyde, a borate or a vinyl sulfone. Synthetic inhibitors of 20S proteasomes include compounds such as carbobenzoxyl-leucinyl-leucinyl-leucinal (MG132), calpain inhibitor carbobenzoxyl□ leucinyl-leucinal (MG135), N-asetyl-leucinyl-leucinyl-norleucinal (MG101) or carbobenzoxy-leucinyl-leucinyl-norvalinal (MG115). All of these inhibitors are chymotrypsin-like selective inhibitors which are widely used in studies of the role of the proteasome in various cellular processes (Lee and Goldberg, 1996; Kawazoe *et al*., 1998; Myung *et al*., 2001; Adams, 2003a).

Natural products used as proteasome inhibitors include the *Streptomyces lactacystinaeus* metabolite, lactacystin that targets the 20S proteasome by an irreversible modification of the amino-terminal threonine of the β-subunits (Fenteany *et al.*, 1995). In addition to its high proteasome specificity, lactacystin also inhibits other cellular proteases (Myung *et al*., 2001). Bush *et al*. (1997) showed that by using proteasome inhibitors, the accumulation of misfolded proteins in the cell markedly increased the levels of expression of heat shock proteins and ER chaperones, such as Hsp90, Hsp70, Hsp47, Hsp40, Hsp25 and Grp94, BiP, ERp72. Inhibition of proteasome function was shown to result in programmed cell death (apoptosis) in mammalian cells (MacLaren *et al*. 2001).

1.2.6 The role of proteasome in human diseases

Intracellular protein turnover occurs extensively and the stability of many proteins is regulated individually and can vary under different conditions. Proteolysis of cellular proteins is a very complex, temporarily controlled and highly regulated process. Due to the large number of protein pathway interactions in an organism, it is not surprising that the proteasome can be linked to several human diseases associated with protein misfolding events. Cystic fibrosis is one example where mutations in the gene encoding a crucial transport protein results in the CFTR (**C**ystic **F**ibrosis **T**ransmembrane Conductance **R**egulator) folding improperly and not being secreted in the sufficiently high quantities required for proper function (Riordan, 1999). Considerable attention has been paid to a group of diseases where proteins or fragments of proteins are converted from their normally soluble forms to insoluble fibrils or plaques which accumulate in a variety of organs including the liver, spleen and brain. This group of diseases, of which nearly 20 have been described, includes neurodegenerative diseases such as Alzheimer's and Parkinson's diseases, certain malignancies, type II diabetes, and disorders of the immune and inflammatory system (Dobson, 1999; Chung *et al*., 2001; Ciechanover, 2006).

1.2.7 Approaches for proteasome purification

The unicellular yeast *Saccharomyces cerevisiae* is an acknowledged model of a eukaryotic cell and the majority of published purification methods for the 26S proteasome come from yeast studies (Glickman *et al*., 1998; Verma *et al*., 2000). However, the 20S proteasome particle was first purified from rabbit erythrocytes in 1987 (Hough *et al*.). Since then proteasomes have been isolated from higher plants (e.g. Udvardy, 1993; Fujinami *et al*., 1994; Yanagawa *et al*., 1999), mammalian cells (e.g. Ugai *et al*., 1993; Tripler *et al*., 1997) and from yeast (Glickman *et al*., 1998; Kimura *et al*., 2003). Isolation of the 20S proteasome particle from a filamentous fungus also has been reported recently (Grinyer *et al*., 2007). While the earlier proteasome purification methods relied on glycerol gradients, PEG precipitation with multistep chromatography or tags combined with immunochromatography, more sophisticated high affinity chromatography systems have been applied recently to simplify the purification process. A simple and relatively rapid isolation and purification method for the entire 26S proteasome from *T. reesei* was developed during this study.

While the central chamber of the 20S proteasome particle houses three protease activities, the chymotrypsin-like activity is the most commonly used indicator of proteasome activity in purification methods. The activity of the proteasome can be quantified by measuring the amount of fluorescence produced by cleavage of the fluorogenic substrate LLVY-AMC. The peptide LLVY is bound to the fluorophore 7-Amino-4-Methylcoumarin (AMC), which is released by proteasomal cleavage, and the fluorescence emitted can be measured (Groll *et al*., 2008).

1.3 The secretory pathway in *Trichoderma reesei*

The filamentous fungus *T. reesei* is quite extraordinary in its capacity to secrete large amounts of proteins and metabolites into the growth medium. In addition, the Generally Regarded As Safe (GRAS) status for biotechnological processes makes it widely exploited for industrial enzyme production (Conesa *et al*., 2001). Knowledge of the fungal secretory pathway is still limited although considerable research has been carried in this field over the past two decades. The unicellular yeast *S. cerevisiae* has been one of the major models of eukaryotes for studies of intracellular protein transport. However, the secretory pathway in multicellular filamentous fungi differs from that in yeast and higher eukaryotes (reviewed in Archer and Peberdy, 1997). Fungi are composed of hyphal compartments and grow by elongating hyphae by extension at the tips. By branching, it creates a new growing axis. The hyphal area is divided by septa containing pores for inter-cellular communication purposes. The hyphal tip contains a highly polarised organelle distribution for cell growth (Maruyama and Kitamoto, 2006). This polarity is not found in either *S. cerevisiae* or in higher eukaryotes. This polarity may reflect the extraordinary high protein secretion capacity in filamentous fungi. For example, *T. reesei* has been reported to secrete up to 100 g L^{-1} of cellulases (Cherry and Fidantsef, 2003). Gram per litre yields of proteins have been reported for the unicellular yeast *P. pastoris*, which however does not secrete the protein effectively into the external cultivation medium (Kobayashi *et al*., 2000). A schematic view of the fungal secretory pathway is given in Fig. 9.

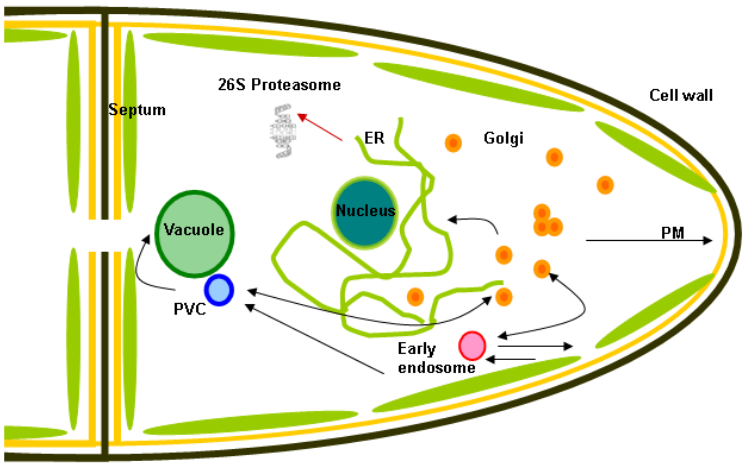

Figure 9. A schematic overview of secretory pathway in filamentous fungi. Newly -synthesised proteins enter the ER (■) and are folded and glycosylated there, then delivered to Golgi apparatus (■) by vesicular trafficking. After further glycosylation at the Golgi, proteins are secreted into the outside environment, mostly at hyphal tips. Unfolded proteins (red arrow) are identified and sorted in the ER and targeted to the proteasomes (grey) for degradation (modified from Kuratsu *et al*., 2007).

Newly-synthesised proteins enter the ER, where they are folded and undergo distinct modifications such as glycosylation, disulfide bridge formation, phosphorylation, and subunit assembly. Unless they contain an ER retention signal, the completely folded proteins proceed to the Golgi compartment for further glycosylation and processing by vesicular trafficking with a series of vesicle buddings and fusions. The vesicles bud from the donor membrane and fuse with the target membrane in a highly organised process. Proteins are then sorted for their final destinations in the *trans*-Golgi network (TGN). From TGN, proteins are directed to the endosomal / vacuolar route or to the plasma membrane for secretion at hyphal tips (Fig. 9). In some cases, the proteins do not reach the extracellular space but are targeted to intracellular compartments and undergo retrograde transport to the donor organelle and become resident proteins. Unfolded proteins are sorted in the ER and targeted to vacuoles or proteasomes for proteolytic degradation (Novick *et al*., 1995; Kuehn and Schekman, 1997; Conesa *et al*., 2001; Gordon *et al*., 2000).

Filamentous fungi, and *T. reesei* amongst them, are the most efficient eukaryotic cell factories available today because of their ability to correct post-translational modifications and their extraordinary protein secretion capacity. Nevertheless, the secretion of heterologous proteins remained 10-1000 times lower than the yields of endogenous

secreted proteins in spite of relatively efficient mRNA transcription. There is evidence that several foreign proteins expressed in filamentous fungi are lost in the secretory pathway, probably because of incorrect processing or misfolding, resulting in their elimination by proteases (Archer and Peberdy, 1997; Gouka *et al*., 1997). Because of the complexity of the molecular mechanisms of the secretory pathway in filamentous fungi, more research is required to understand the protein maturation, quality control, transport and selection in ER. In this work the major secreted protein, CBHI of *T. reesei,* was chosen as a model molecule to study protein quality control and the role of proteasome.

1.3.1 Cellobiohydrolase I, the main secreted protein in *Trichoderma reesei*

A wide variety of cellulolytic and hemicellulolytic enzymes are secreted efficiently by *T. reesei.* Genes encoding cellulases belonging to 22 families of the 115 glucoside hydrolase families have been found in the *T. reesei* genome based on amino acid sequence similarities (http://www.cazy.org/fam/acc_GH.html). Twenty two open reading frames (ORFs) corresponding to known cellulolytic enzymes have been identified, including two cellobiohydrolases, CBHI (Cel7A), CBHII (Cel6A) and six endoglucanases EGI (Cel7B), EGII (Cel5A), III (Cel12A), EGIV (CEL61A), EGV (Cel45A), EGVI (Cel74A), two β glucosidases (Cel1A,Cel3A) and one β-xylosidase (Bxl1). Besides the cellulases, three xylanase encoding gene ORFs (*xyn1, xyn2 and xyn3*) and one mannanase encoding gene ORF (*man5A*) were also found in the *T. reesei* genome (Divne *et al*., 1998; Bailey and Nevalainen, 1981; Bourne and Henrissat, 2001; Miettinen-Oinonen 2004; Ouyang *et al*., 2006). Cellobiohydrolases act synergistically with endoglucanases. CBHI contains 497 amino acid residues and has a molecular mass of 59-68 kDa depending on the level of glycosylation. CBHI is an exoglucanase which hydrolyses crystalline cellulose by cleavage from the reducing end of the polysaccharide chain. CBHI is capable of cleaving the 1,4-β D-glucosidic linkages that bind glucose residues together in the cellulose chain. CBHI makes up to 60 % of the total secreted protein detected in the culture medium of *T. reesei* (Penttilä *et al.*, 1998; Divne *et al*., 1994).

CBHI has a large catalytic core, which is connected to a small cellulose binding domain (CBD) by a heavily glycosylated linker region (Fig. 10). The CBD is responsible for most of the enzyme affinity for cellulose and its presence is essential for activity on crystalline

cellulose, but has no effect on hydrolysis of soluble substrates (Reinikainen *et al*., 1992). The CBD is formed from 63 amino acids (amino acids 435-497 in the CBHI molecule). Amino acids 1- 434 form the CBHI core region. The CBHI core contains a tunnel-shaped active site, formed by several long surface loops that connect tightly packed secondary structural elements, forming a β-sheet sandwich structure. The active site is at the far end of the tunnel and features three acidic residues, Glu212, Asp214 and Glu217, which are suggested to be catalytically important. All three residues are required for efficient catalysis (Ståhlberg *et al*., 1996; Divne *et al*., 1998; Koivula *et al*., 2002).

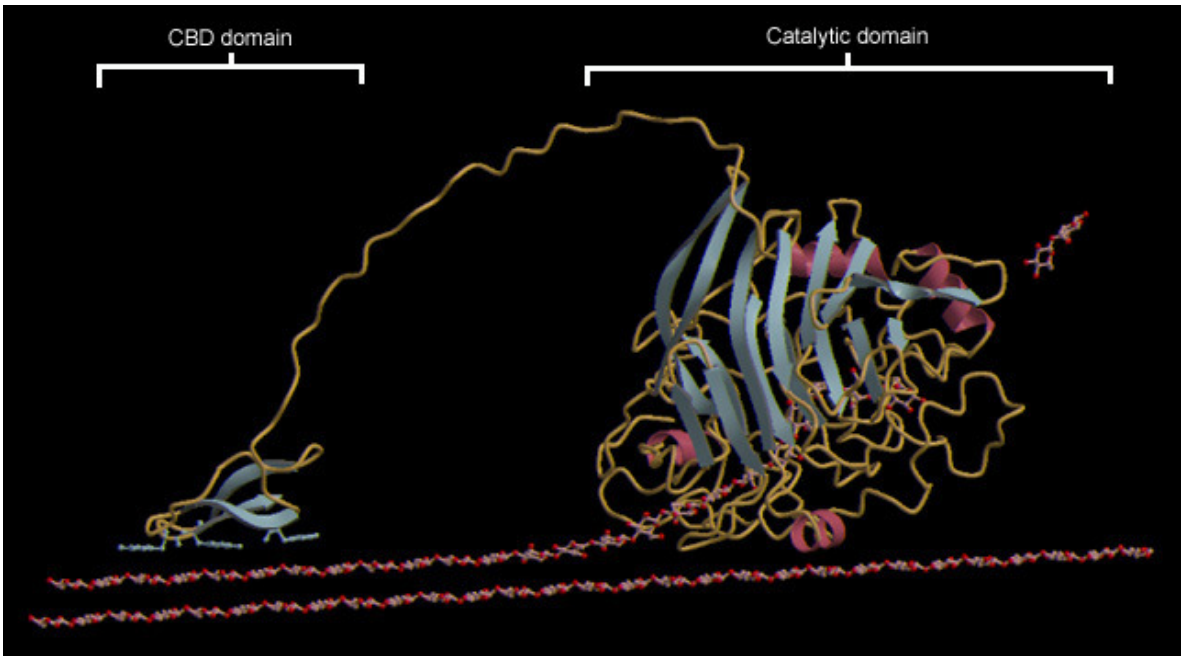

Figure 10. Model of *T. reesei* cellobiohydrolase, CBHI interacting with cellulose. The large module on the right hand side contains the catalytic centre. The catalytic domain is connected by a linker to the CBD on the left hand side. The CBD is required for binding to crystalline cellulose (by authorisation of Prof. C. Divne; http://mcdb.colorado.edu/courses/3280/chime/cel/cbh1x3.html)

The crystal structure of the CBHI catalytic domain has been determined (Divne *et al*., 1998). The CBHI catalytic core contains 20 cysteine residues which form disulfide bridges. Figure 11 shows the location of the active site and the positions of cysteine residues and the disulfide bridge-forming pairs in CBHI core.

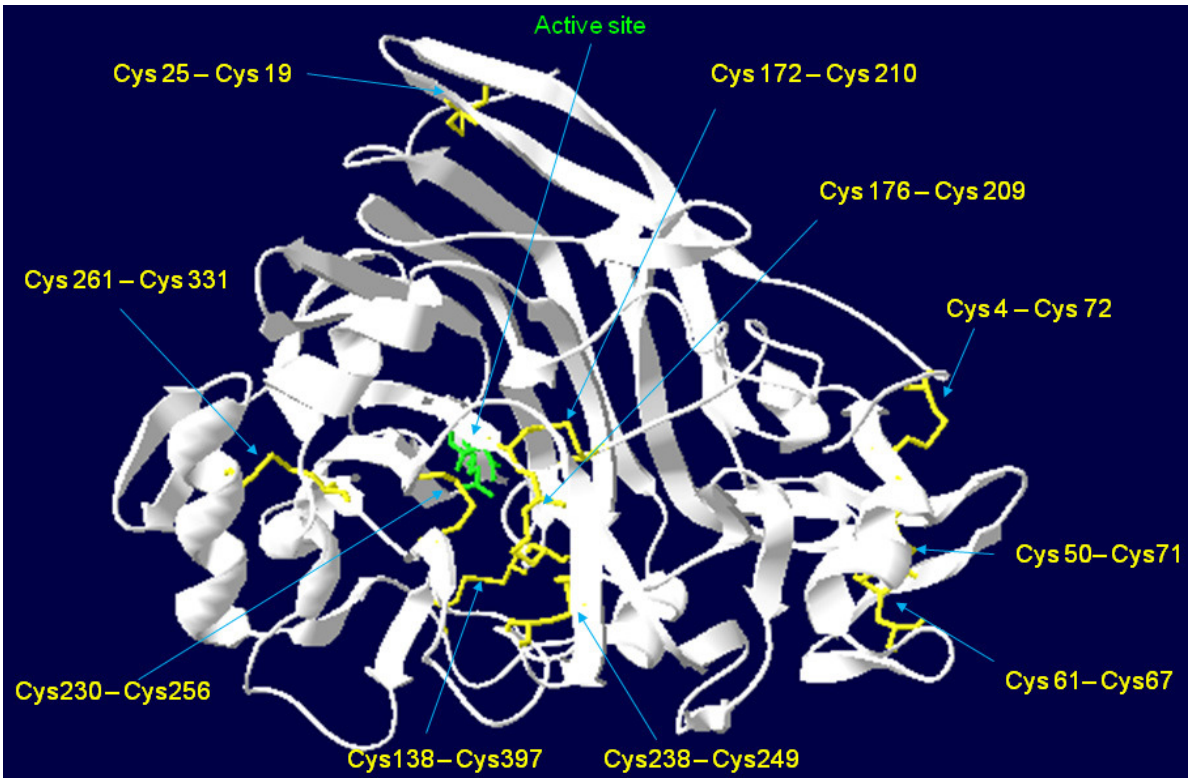

Figure 11. Positions and pairs of the 20 cysteine residues which form disulfide bridges, and location of the active site in the CBHI core (Divne *et al*., 1994).

For a cysteine residue, one sulphur-containing thiol group is capable of combining with the thiol group of another cysteine residue to form a disulphide bridge, either linking two peptide chains together or causing a single peptide chain to fold back on itself to make a loop, which is also important for structural proteins and enzymes to maintain their proper secondary structure. Generally, disulphide bridges stabilise protein structure by forming both intramolecular and intermolecular cross-links (Regeimbal and Bardwell, 2002).

The maintenance of a specific structure inside the cell is a requirement for efficient and rapid protein secretion. Katakura *et al*. (1999) showed that a single amino acid change can accelerate β-lactoglobulin production six times in *S. cerevisiae*. On the other hand a misfolded membrane glycoprotein G was shown to not only be retained in the ER alone, but also it was moved to the intermediate compartment and to the cis-Golgi network and later recycled back to the ER, due to its unfolded and unassembled conformation (Hammond and Helenius, 1994).

So far, the production of heterologous proteins of fungal origin has been successful in yielding grams per litre, but production of heterologous protein especially originating from more distant species has remained lower in *T. reesei.* Calf chymosin was the first mammalian protein successfully produced in *T. reesei* (Harkki *et al*., 1989). Since then, improvement of heterologous proteins secretion has been the centre of interest. Different strategies, such as applying strong gene promoters (Harkki *et al*., 1989, Uusitalo *et al*., 1991; Keränen and Penttilä 1995), gene fusion systems (Nyyssönen *et al*., 1993), increased gene copies (Veroes *et al*., 1993; Karhunen *et al*., 1993), modification of codon usage of genes encoding the expressed product (Moralejo *et al*., 1999; Te□o *et al*., 2000), over-expression of foldases and chaperones (van Gemeren *et al*., 1997; Punt *et al*., 1998, Moralejo *et al*., 2001; Conesa *et al*., 2002) used to improve the secretion of homologous protein production have not been very successful for heterologous protein expression.

Many studies showing the genes encoding quality control-related proteins are up-regulated in heterologous expression in *T. reesei,* such as BiP in case of antibody fragment expression (Saloheimo *et al*., 1999). In contrast, heterologous expression of a bacterial xylanase (XynB) in *T. reesei* did not induce UPR (Te□o *et al*., 2000; J. Hekelaar *et al*., unpublished work). Over-expression of the UPR pathway regulator, Hacp1 was found to increase expression of ER-resident foldases BiP and PDI in *T. reesei* (Valkonen *et al.*, 2004). Similarly over-expression of Hac1p in *T. reesei* increased production of *Bacillus* α-amylase. However, it did not occur with a heterologous laccase but its yield increased several fold when produced in *A. niger* (Valkonen *et al*., 2003b). In addition, the production of both *Bacillus* α-amylase and *T. reesei* endoglucanase EGI were reduced in *hac*-deletion strain of *T. reesei,* suggesting that there are some components in secretion pathway which are regulated by signals from the UPR pathway (Valkonen *et al*., 2004).

Because of the complexity of the mechanisms involved in the secretory pathway, which includes protein maturation, protein quality control and trafficking, further research is required for better understanding of this intriguing pathway. In the current study, several mutated forms the endogenous cellobiohydrolase, CBHI, were created to study the effects of their expression on the UPR and ERAD pathways.

1.4 Strategies to study protein quality control in filamentous fungi

The well-characterised and highly secreted CBHI protein of *T. reesei* was chosen as a model for the current study. Three transformant strains each expressing a different mutant CBHI molecule were constructed. Secretion of the mutant protein and changes in gene transcription levels using genome wide microarrays were studied to establish whether expression of a mutant CBHI resulted in induction of the UPR and/or ERAD pathways. Interaction between a mutant CBHI and the proteasome was also explored.

1.4.1 Generation of mutant proteins

In vitro mutagenesis can be applied to a specific site in a pre-determined way (site-directed mutagenesis, SDM), the strategy depending on the purpose of the mutagenesis approach. For example, Bardy *et al*. (2005) used the method of site-directed mutagenesis to replace the signal peptidase residues serine, lysine and aspartic acid by histidine to investigate the effect of a single amino acid replacement to the activity of the signal peptidase. The importance of C-terminal amino acid residues of a bovine pancreatic ribonuclease A in the formation of the three-dimensional structure of the protein was explored by site-directed mutagenesis by replacing a particular amino acid Vall24 with amino acid residues with different properties (Fujii *et al*., 2000). These are only a few examples of the use of this powerful technology.

In the current study, site-directed mutagenesis was used to engineer particular mutations in the *cbh1* catalytic core by replacing certain cysteine residues with proline. The replacement of cysteine would prohibit the formation of disulphide bridges in their original positions and thus destabilise the three-dimensional structure of the protein, resulting in a misfolded CBHI protein (Flinn *et al*., 1999; Knowles and Zahn, 2006). The changes were predicted to have an effect on secretion kinetics, product yields and binding of the substrate into the active site of CBHI due to conformational modifications to the protein.

There is some published literature reporting on the effect of mutations on protein yields. For example, Kumita *et al*. (2006) showed that secreted yields of mutant forms of human lysozyme produced in *Pichia pastoris* were 30-200 fold lower than those of a wild type

lysozyme. In another study, Katakura *et al*. (1999) changed one amino acid in β□ lactoglobulin which led into a six times decrease in the yield in *S. cerevisiae*.

It is a well-known fact that mutations in a protein can affect the kinetic behaviour of the protein (Kowalski *et al*., 1998). A small fraction of each secreted protein is retained and degraded by the quality control in the endoplasmic reticulum (ER) that restricts export to incorrectly folded proteins. Mutations in the protein increase the retention.

In our work, changing the cysteine residues to prolines in CBHI was expected to change the redox potential of the ER and thereby increase retention of the mutant protein by impairing oxidative folding of the proteins in the cells that secrete large quantities of protein (Merksamer *et al*., 2008).

1.4.2 Proteasome proteomics

As a broad definition, proteomics refers to the large-scale study of protein expression, structure and function. A proteomic study typically features separation of protein mixtures by one- or two-dimensional polyacrylamide gel electrophoresis (1D- or 2D-gel) and identification of individual protein spots or bands by mass spectrometric techniques e.g. matrix assisted laser desorption/ionisation mass spectrometry (MALDI-MS/ MS-MS) and electrospray ionisation liquid chromatography mass spectrometry (ESI-LC MS/MS), after digestion of the sample to generate peptide mixtures (Monti *et al*., 2005; Guerrera and Kleiner, 2005).

Current proteomic techniques offer the possibility of studying up- and down-regulation of protein expression, and also facilitate studies into functional proteomics featuring characterisation of protein activities, multiprotein complexes and signalling pathways. Identification of interacting partners is necessary to understand the function of a protein within the cell. Identification of interacting proteins in stable complexes in a cellular system is commonly achieved by affinity-based procedures (Monti *et al*., 2005). An example of such protein complex is the proteasome studied in this work.

Mass spectrometry-based proteomic approaches have been used recently for the identification of fungal proteasome subunits and proteasome-interacting proteins. In yeast, structural organisation of the 19S proteasome subunits and their interaction with each other was explored by combination of mass spectrometry and cross-linking of 19S subunits (Sharon *et al*., 2006). Verma *et al.* (2000) identified 52 proteasome-interacting proteins by combining affinity purification of the 19S and 20S proteasome particles with direct analysis of the composition of large protein complexes (DALPC) by mass spectrometry from the budding yeast. Multiubiquitylated proteins have also been captured by a two step affinity purification method that was combined with LC/LC-MS/MS analysis. This study resulted in identification of 120 potential ubiquitin target proteins, most of them functioning in translational and metabolic pathways (Mayor *et al*., 2005). In the current study, column chromatography and mass spectrometry analysis was applied for the purification, identification of the 20S proteasome particle and the entire 26S proteasome along with proteasome-interacting proteins from *T. reesei* (Grinyer *et al*., 2007; Kautto *et al*., 2009). These papers are the first to describe purification of proteasome particles from a filamentous fungus.

1.4.3 Approaches to study gene expression

Although more sensitive techniques have been developed, Northern blotting is still the most popular method used for detection and quantification of mRNA levels because it allows direct comparison of mRNA abundance between samples (Streit *et al*., 2009). However, the procedure is quite time consuming and requires preparation of a probe for each gene of interest. Therefore, it is not suitable for large scale screening purposes. On the other hand, Northern blot analysis allows the expression or lack of expression of specific DNA sequences to be related to the physiological and morphological properties of a living organism (Henderson, 1991).

In fungal studies, Northern blotting has been used to determine mRNA transcriptional levels of the *hac1* gene encoding a UPR transcription factor and mRNA levels of genes such as *pdi1* and *bip1* encoding ER-chaperones under different UPR-induced conditions in *T. reesei* (Saloheimo *et al*., 2003), *A. niger* (Mulder *et al*., 2004), and *A. fumigatus* (Richie *et al*., 2009). Northern blotting analysis clearly showed the effect of expression of the

heterologous protein tissue plasminogen activator in *A. niger*: ER stress was reflected by increased levels of mRNAs of *hacA*, *bipA* and *pdiA*. Decreases in gene expression levels of multiple cellulase genes was determined in a time course by Northern blotting, when UPR was induced with DTT, brefeldin A or A23187 (Ca^{2+} ionophore) in *T. reesei* (Pakula *et al*., 2003). Also, expression of secretion pathway related genes (*ypt1* and *nsf1*) of *T .reesei* was determined by Northern blotting in drug induced UPR conditions (Saloheimo *et al*., 2004).

Microarray technology is used widely for global studies into gene expression. In fungi, the first microarray studies came from the budding yeast *S. cerevisiae,* published 10 years ago (DeRisi *et al*., 1997 and Lashkari *et al*., 1997). Microarrays have since been applied to more than 20 species of filamentous fungi. Most have been carried out using PCR products amplified from expressed sequence tag (EST) libraries but more recently, predicted ORFs have been used in genome-wide expression studies. Currently, microarrays cover a wide spectrum of research areas addressing filamentous fungi, including studies into fungal metabolism, development, fungal symbiosis, pathogenesis of fungi and industrial applications (Breakspear and Momany, 2007). For example, Foreman *et al.* (2003) used microarrays to compare the expression levels of genes identified previously encoding biomass-degrading enzymes in *T. reesei.* More recently, Bonaccorsi *et al.* (2006) studied the transcriptional response of *T. reesei* to hypoxia and transient anoxia. Only a few microarray studies have been performed specifically to monitor the UPR and ERAD pathways and their regulation during environmental stress cultures applied either by different culture conditions or by drugs (Higgins *et al*., 2003). GeneChip® from Affymetrix (Lipshutz *et al.* 1999) is probably the most commonly used chip for microarray studies. The company's first product, HIV genotyping GeneChip®, was released in 1994 and since then Affymetrix have received significant benefits from its patent portfolio in this area.

While microarray technology has become used widely for initial profiling and primary screening of transcriptomes, further validation techniques such as Northern blotting or qRT-PCR are required to confirm the microarray results (Chuaqui, 2002). Real time PCR combined with expression quantification offers the method for analysing expression levels of many genes simultaneously. It can be also used as a quick screening method with a small amount of sample for the analysis of expression patterns in large gene families (Nebenführ *et al*., 2004; Bustin and Mueller, 2005; Nolan *et al*., 2009). Ye and Pan (2008) used qRT

PCR for comparison of transcripts of UPR-relevant marker genes in solid state and submerged cultures of *A. oryzae*.

Recently, a new method called TRAC (**tr**anscript analysis with the aid of **a**ffinity **c**apture; Rautio *et al*., 2006) has been developed for transcriptional analysis. TRAC facilitates a rapid, simultaneous analysis of 96 samples, allowing high throughput multiple target detection. Crude cell material can be used directly for hybridisation, thus bypassing RNA extraction and cDNA conversion (Rautio *et al*., 2008.) These advantages and the brief hands-on time involved make TRAC a versatile method for an on-line monitoring system. The method has been applied for monitoring biological processes eg. evaluation of gene expression stability in chemostat cultures of *T. reesei* and *S. cerevisiae* (Rautio *et al*., 2006; Wiebe *et al*., 2008) and studying expression of genes related to protein folding in strains producing foreign proteins in *S. cerevisiae* and *Pichia pastoris* (Gasser *et al*., 2007).

1.4.4 Visualisation of proteins by microscopy

The discovery of the green fluorescent protein (GFP) from the jellyfish *Aerquorea victoria* by Osamu Shimoura and Franck Johnson (1961) started a whole new era in cell biology enabling the monitoring of cellular processes in a living cell. Advanced technologies using optical microscopy and related methodologies coupled with widefield fluorescence and confocal microscopy, digital cameras and multitracking laser control systems are offering almost unlimited possibilities for simultaneous observation of multiple fluorochromes. The fluorescent properties of GFP have been enhanced by mutations to develop enhanced fluorophores such as blue (BFP), cyan (CPF) and yellow fluorescent proteins (YFP). DNA-based fluorescent proteins can be incorporated into proteins by genetic fusion. Fluorescent chimeric proteins can be expressed in living cells, tissues and entire organisms after transformation with engineered recombinant vectors with no requirement for exogenous substrates and cofactors (Feng *et al*., 2000; Zhang *et al*., 2002).

Fluorescent protein tags can be used in a variety of different imaging methods for establishing protein roles and interactions within subcellular compartments in living cells. More recently, fluorescent proteins from other species have been isolated and purified, expanding the color palette for a wide spectrum of applications using tagged biological

molecules. In addition, new techniques have made it possible to label proteins with quantum dots, metal-ligands, colloidal gold beads, metal clusters, metal binding domains, enzyme molecules, small organic fluorophores and semiconductor nanocrystals (He *et al*., 2001; Shen *et al*., 2008, reviewed in Malecki *et al*., 2001).

Venus, the yellow fluorescent protein, is the result of continued development of enhanced green fluorescent protein (EGFP). Many studies have confirmed that Venus is stable under low pH conditions and extremely resistant to photobleaching and brighter than EGFP (Nagai *et al*., 2002; Rekas *et al*., 2002). Venus has an absorption maximum at 514 nm and emission maximum at 528 nm. These unique properties and its fast and efficient maturation make it a versatile reporter in multiple labelling experiments. For example, it has been used as an acceptor in fluorescence resonance energy transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM) (Chaflie *et al*., 1994; Koushik *et al*., 2006). The unique characteristics of Venus make it attractive to be used as a reporter of gene expression. In this study, the Venus fluorescent protein was applied to monitor CBHI expression and secretion in *T. reesei*.

Fluorochromes have been widely used also in determination of the expression and subcellular localisation of various proteins in fixed cells. A wide range of antibodies conjugated to different fluorochromes is available commercially for direct or indirect immunolabelling. Affinity-purified Fab or F(ab)2 -conjugated fluorochromes have become widely used. The non-specific binding to cell Fc receptors can be avoided using Fab fragments lacking the Fc portion of IgG instead of using the entire IgG molecule. More recently, single chain fragments (scFv), consisting of the variable regions of light (V_L) and heavy (V_H) chains of an antibody molecule conjugated to the fluorochromes, have become available for biological applications (Uchanska-Ziegler *et al*., 2000). Similarly, gold-particles can be conjugated with IgG or IgM and have been used in immunohistochemistry (Kleymann *et al*., 1995; Fisher *et al.*, 2001). A wide range of different antibody fragments conjugated with different sized gold particles is available commercially for immunolabelling.

Confocal Laser Scanning Microscopy (CLSM) is a valuable tool for obtaining high resolution images and 3-dimensional reconstruction at the resolution of the light

microscope. The key feature is the production of blur-free images by thin optical sectioning. The use of 3-D reconstruction of a Z-series enables the visualisation of larger structures and compartments. Transmission electron microscopy (TEM) can provide high resolution images and ability to visualise particle ultrastructures, but it is a relatively laborious and time-consuming method when sample preparation, sectioning and staining are included. However, because of its high resolution, TEM can provide a versatile method to observe protein-protein interactions inside cells. (Zhao and Rodgers, 2006). Immunogold-labelling and transmission electron microscopy have been used widely to analyse intracellular protein and organelle distribution in eukaryotic cells. One example from fungi is the detection of the subcellular location of the endogenous CBHI protein and heterologous EPB protein in *T. reesei* (Nykänen *et al*., 1997). In the current study, two different sized gold labeled secondary antibodies were used to detect the primary antibodies against the 20S proteasome and the main secreted protein CBHI. Possible interactions between these two proteins were observed by TEM.

1.5 Aims of this study

The cellular mechanisms involved in protein production in *Trichoderma reesei* have been studied for a long time with the aim of improving the fungus as an industrial production host. While production of heterologous proteins has been successful in *T. reesei*, similar yields to endogenous proteins have not been obtained as yet. Some potential reasons for the low yield include misfolding and subsequent elimination of the foreign proteins from the secretory pathway. In this study, the highly secreted CBHI was chosen as a model protein to study the effects of expression and secretion of mutant forms of CBHI on protein quality control in *T. reesei*. The structure of CBHI is well known and the three dimensional structure of CBHI is available. The broad aim of this study was to analyse whether the mutant CBHI production in *T. reesei* leads to induction of the UPR and ERAD pathways.

More detailed aims of this study were:

1. Identification and characterisation of the 20S and 26S proteasomes from *T. reesei*.

2. Generation of unfolded mutant versions of the endogenous cellobiohydrolase I, CBHI.

3. Examination of the effect of expression and secretion of three mutant CBHI forms on protein quality control in the *T. reesei* hyphae by targeted and genome wide microarray analyses.

4. Exploration of the interaction between the proteasome and mutated CBHI in *T. reesei* by fluorescence microscopy and transmission electron microscopy.

The study presented in this thesis introduces a new strategy to explore the effects of secretion stress by expression of a native misfolded protein instead of treating the growing hyphae with drugs. This strategy proved successful and revealed effects that were different to earlier studies with drug-induced stress. The *T. reesei* proteasome was successfully

purified and co-localised with the mutant CBHI protein in the fungal hyphae providing indirect evidence for degradation of the faulty protein at the proteasome.

2 Materials and methods

2.1 Isolation and purification of the proteasomes from *Trichoderma reesei*

Fungal proteasomes were purified by two different methods. The protocols for preparing the mycelial extracts were the same except for lysis of the cells, which was performed either by sonication or using a French Press. The first purification was based on the method by Legget *et al*. (2005) with minor modifications. The second method was carried out with a protocol developed in-house involving a POROS® HQ column (anion exchange, Applied Biosystems, USA) and Sephadex S-500 HR column (GE Healthcare Life Sciences, Sweden). In addition, there were differences in the buffers used in the chromatography steps.

2.1.1 Fungal strain and cultivation conditions

The proteasome was isolated from *Trichoderma reesei* Rut-C30 (Eveleigh and Montenecourt, 1979). The strain was maintained and sporulated on potato dextrose agar (PDA, Oxoid, UK) plates. For liquid cultures, 2 L Erlenmeyer flasks containing 500 mL minimal medium (Penttilä *et al*., 1987) supplemented with 2 % (v/v) glycerol were inoculated with 10^7 mL^{-1} conidia and the cultures were grown at 28 ^oC for 48 h, at 250 rpm on a shaker.

2.1.2 Purification of the fungal proteasomes

Purification of the 20S proteasome

T. reesei mycelia were harvested by centrifugation at 4000 x g for 20 min, washed three times with distilled water and disrupted in two times the sample volume of proteasome extraction buffer (PEB; 50 mM Tris (pH 7.5), 5 mM $MgCl_2$, 10 % (v/v) glycerol, 5 mM ATP, 1 mM DTT, 1 mM PMSF) and 0.1 % (v/v) fungal and yeast protease inhibitor cocktail (P8215, Sigma, Australia). The mycelia were lysed by sonicating at 60 % intensity for five cycles of 15 s on a Branson Sonifier 250 (Branson Ultrasonic, CT, USA). The

samples were kept on ice throughout the sonication cycles. Cellular debris was pelleted by centrifuging at 4000 x g for 30 min at 4 °C followed by ultracentrifugation of the supernatant at 126 000 x g for 1 h at 4 °C. Finally, the extract was passed through Miracloth (# 475855, Calbiochem®, USA) to remove lipid material.

The cell extract prepared above was used as the starting sample for chromatographic purification steps which were performed at 4 ^{o}C to minimise dissociation of the 20S proteasome complex. The cell extract was applied onto a 100 mL DEAE-Affigel Blue resin (Biorad, Australia) column (affinity interaction) at a flow rate of 2 mL min^{-1}, washed with 100 mL PEB and then 200 mL PEB, 50 mM NaCl at 4 mL min^{-1} before eluting with 200 mL PEB, 150 mM NaCl at 4 mL min^{-1}. Each 10 mL fraction collected (Pharmacia FLPC System, Sweden) was assayed for chymotrypsin-like activity using the 20S Proteasome Assay Kit AK-740 (Biomol, USA) following the manufacturer s instructions. Fractions found positive for chymotrypsin-like activity were pooled and further purified on a 50 mL Source Q resin ion exchange column (GE Healthcare, USA) at 2 mL min^{-1}. The sample was eluted using a linear salt gradient of 100 mM NaCl to 500 mM NaCl in PEB buffer at 4 mL min^{-1}. Fractions of 6 mL were collected and assayed for chymotrypsin-like activity. All fractions showing chymotrypsin-like activity that eluted at approximately 300 mM NaCl were pooled and concentrated to 1 mL on a 30 kDa centrifugal concentrator (Amicon® Ultra-15, Millipore, USA). The concentrated sample was separated on a 100 mL size exclusion column containing Superose 6 (Biorad) at 1 mL min^{-1} and eluted with PEB, 100 mM NaCl at 1 mL min^{-1}. Fractions of 2 mL were collected and assayed for chymotrypsin-like activity as described above. Pooled fractions positive for chymotrypsin-like activity were concentrated by a 30 kDa centrifugal concentrator and the total protein concentration was determined by the method of Bradford (1976) using reagents purchased from Bio-Rad. Bovine serum albumin was used as a standard. The sample was stored at -80 ^{o}C until further use.

Purification of the 26S proteasome

Differences in the method applied for purification of the whole 26S proteasome complex compared to the previous protocol were: 1) cell lysis by a French Press, 2) a different buffer in the chromatography step (buffer B, see below) and 3) two columns instead of three.

After the mycelia were harvested by centrifugation, complete cell lysis was achieved by passing the fungal mycelia through a French Press (French Press®, Thermo Scientific, USA) at 15000 p.s.i. twice. After centrifugation and filtering through Miracloth (as described above), the fungal extract was concentrated down to a 5 mL volume, the total protein concentration measured as mentioned above and the sample stored at +4 ^{o}C until purification was continued with the chromatography step. The concentrated fungal extract was applied onto a 1 mL POROS® HQ column (anion exchange, Applied Biosystems, USA) at a flow rate of 5 mL min^{-1}, using an ÄKTA explorer 10S FPLC system (GE Healthcare Life Science, Sweden), washed with 25 mL buffer B (10 mM Tris - NaH_2PO_4 pH 7.5 - 4 mM ATP) at 5 mL min^{-1} before elution with a linear salt gradient of 0 mM NaCl to 1000 mM NaCl in buffer B at 5 mL min^{-1}. Each 1 mL fraction collected was assayed for chymotrypsin-like activity using the 20S Proteasome Assay Kit AK-740 (Biomol, USA) following the manufacturer's instructions. Fractions found positive for chymotrypsin-like activity were pooled and concentrated to 1 mL on an Ultrafree concentrator with a molecular weight cut-off of 5 kDa (Amicon® Ultra-15, Millipore, USA). The concentrated sample was separated by a 150 mL size exclusion column containing Sephadex S-500 HR (GE Healthcare Life Sciences, Sweden) at 0.5 mL min^{-1} with 1 M NaCl in buffer B. Fractions of 1 mL were collected and assayed for chymotrypsin-like activity as described above. Fractions found positive for chymotrypsin-like activity were pooled and a buffer exchange was performed to reduce the high salt concentration on an Ultra free concentrator with a molecular weight cut-off of 5 kDa (Amicon® Ultra-15, Millipore, USA). After concentration, a 1 mL sample was stored at -80 ^{o}C until required.

2.1.3 Denaturing polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting of the purified *Trichoderma reesei* proteasome

Approximately 45 µg of purified *T. reesei* proteasome was separated as described by Laemmli (1970) on a one-dimensional 4-12 % (w/v) SDS-PAGE gel and protein bands were visualised with Coomassie Colloidal Blue 250G. For Western blotting, proteins separated by SDS-PAGE were blotted onto PVDF (Immobilon-P, Millipore, USA) membranes at 30V for 1 h. Following blocking with 3 % (w/v) skim milk in the washing buffer (WB; PBS - 0.5 % Tween 20) for 1 h, the membranes were incubated for 1 h with the primary antibody mixture (rabbit polyclonal anti-20S; PW9355 and five monoclonal

anti-19S proteasome subunit IgG; PW8315, PW8160, PW8250, PW8220, PW8215; Biomol Int., USA) in 1:2000 dilution. The blot was washed three times with WB and then incubated with the secondary antibody (alkaline phosphatase conjugated goat anti-rabbit IgG or anti-mouse IgG, Vector Laboratories, USA) in 1:500 dilution for 1 h. After three washes with WB, the secondary antibody was detected with the alkaline phosphatase substrate (*FAST*□ BCIP/NBT; 5-Bromo-4-chloro-3-indolyl phosphate / Nitro blue tetrazolium; Sigma, Australia).

2.1.4 Two dimensional map of the fungal proteasome

Proteasome samples were prepared for 2D electrophoresis by purification with a 2-D Clean up Kit (GE Healthcare, USA) following the manufacturer□s instructions. Following precipitation, the cell pellet was resuspended in 200 µL of sample solution (7 M urea, 2 M thiourea, 1 % (w/v) C7Bz0, 40 mM Tris, 5 mM tributylphosphine, 10 mM acrylamide, 1 mM PMSF) and 0.1 % (v/v) protease inhibitor cocktail (P8215, Sigma, Australia) and incubated at room temperature for 90 min to allow complete reduction and alkylation of proteins (Herbert *et al*., 2003). The reactions were quenched with 10 mM dithiothreitol following centrifugation at 16 000 g for 10 min to remove insoluble material.

After conductivity measurement tests (< 300 µS cm^{-1}), 200 µL samples equivalent to 230□ 300 µg of protein were used directly to passively rehydrate 3-10 or 4□7, 11 cm IPG strips (Amersham Pharmacia, Sweden). IPG strips were focused to a total of 80,000 Volt hours (Vh) using a three-step focusing program. The focusing program included a rapid ramp to 300 V for 4 h, a linear ramp to 10,000 V over 8 h, and a 10,000 V step until 80,000 Vh were reached. IPG strips were equilibrated for 20 min in 6 M urea, 2 % (w/v) SDS, 50 mM Tris□HCl, 0.1 % (w/v) bromophenol blue, pH 8.8. The strips were then placed on top of Proteome Systems 6□15 % Gelchips (Proteome Systems, Australia) or BioRad Criterion Gels (Biorad, Australia) and run at 30 mA constant until the bromophenol blue dye reached the bottom of the gel. The gel was fixed in 10 % (v/v) methanol, 7 % (v/v) acetic acid solution for 30 min and then stained with Sypro Ruby solution (Molecular Probes, Australia) for 16 h. The gel was destained in the fixing solution before scanning on a fluorescence scanner (Alpha Innotech Corporation, USA). Restaining the gel for 16 h was carried out with Coomassie Colloidal Blue G250 (17 % (w/v) ammonium sulphate, 34 %

(v/v) methanol, 3.6 % (v/v) orthophosphoric acid, 0.1 % (w/v) Coomassie G-250) and destained with 1 % (v/v) acetic acid for further analysis as required.

Mass spectrometry and protein identifications

For MALDI-TOF MS analysis, SDS-PAGE gel pieces containing proteins were cut from 2D gels, destained, dried and incubated with trypsin at 37 °C overnight. Each peptide solution was then desalted and concentrated using ZipTips™ from Millipore (USA) and spotted onto the target plate with 1.0 μL matrix solution (4 mg mL^{-1} alpha-cyano-4□ hydroxy-cinnamic acid in 70 % (v/v) acetonitrile, 1 % (v/v) trifluoroacetic acid). Peptide mass fingerprints of tryptic peptides were generated by matrix assisted laser desorption / ionization-time of flight-mass spectrometry (MALDI-TOF-MS) using an Applied Biosystems 4700 Proteomics Analyser with TOF/TOF optics in the MS mode. A Nd:YAG laser (355nm) was used to irradiate the sample.

The spectra were acquired in reflectron mode in the mass range of 750 □ 3500 Da. The instrument was then switched to MS/MS (TOF/TOF) mode where eight peptides with significantly high signals from the MS scan were fragmented by collision-induced dissociation. A near point external calibration was applied to give a mass accuracy within 50 ppm.

The mass spectrometry data was searched against proteins from all fungal species using the Mascot Peptide Mass Fingerprint database where a modified MOWSE scoring algorithm was used to rank results (http://www.matrixscience.com/help/scoring_help.html) (Pappin *et al*., 1993), or using a custom protein database derived from the *T. reesei* sequencing project for peptide mass fingerprinting (PMF) for species specific protein identification as described in Grinyer *et al*. (2007).

2.1.5 Transmission electron microscopy of the *Trichoderma reesei* proteasome

To investigate further the ultrastructure of the proteasome, the purified and desalted proteasome sample was diluted to a final protein concentration of 0.2 μg mL^{-1} with phosphate buffer (50 mM, pH 7.5 containing 2 mM ATP) for transmission electron microscopy (TEM). Negative staining of samples was carried out on a carbon-coated

Formvar film which was glow-discharged prior to sample application. The sample solution was deposited on the film and excess solution was removed by blotting. The grid was washed with distilled water and 2 % (w/v) uranyl acetate solution was applied over the sample for 2 min. After removal of excess stain, the sample was dried and inspected using a Hitachi-17000 electron microscope. Image processing was carried out using ImageJ (Abramoff *et al*., 2004) and Photoshop© Elements programs.

2.2 Generation and analysis of cellobiohydrolase I mutants

Cellobiohydrolase I (CBHI) is the major secreted protein in *T. reesei* and was therefore chosen as a model protein for mutagenesis to examine the effects of the expression of a mutated (misfolded) protein on cellular functions, in particularly on the UPR-ERAD pathway.

CBHI catalytic core with 434 residues contains 10 disulfide bridges formed between 20 cysteine residues (Fig. 12). Table 2 shows the pairs of cysteines and their locations in the CBHI molecule. The following cysteine residues were chosen and replaced with proline amino acids: Cys4, Cys25, Cys172, Cys176 and Cys261 (shown in red in Table 2).

Table 2. Cysteine residues forming disulfide bridges in the native CBHI core protein with their location in the three dimensional structure. S, β-sheet; L, loop; H, α-helix. Cys residues substituted with Pro are labelled in red.

Disulfide bond	*First cysteine*	*Location*	*Second cysteine*	*Location*
1	Cys4	L (Pca1-His11)	Cys72	S (Asn70-Asp74)
2	Cys19	S (Pro12-Ser20)	Cys25	S (Thr24-Asp35)
3	Cys50	S (Thr48-Asp52)	Cys71	S (Asn70-Asp74)
4	Cys61	L (Ser58-Asp63)	Cys67	H (Asn64-Lys69)
5	Cys138	L (Val133-Leu140)	Cys397	L (Gly395-Val403)
6	Cys172	L (Met149-Leu180)	Cys210	Cys210
7	Cys176	L (Met149-Leu180)	Cys209	S (Gly205-Cys209)
8	Cys230	S (Ser222-Thr231)	Cys256	S (Thr255-Asp257)
9	Cys 238	S (Glu236-Gly240)	Cys243	L (Asp241-Gly254)
10	Cys261	S (Cys261-Trp263)	Cys331	H (Asp328-Phe338)

Three different CBHI mutant molecules were designed to induce misfolding of the CBHI core molecule. The location of all substituted Cys residues in the CBHI core are shown in Fig. 12 and labelled in yellow. Cys4 located at the N-terminal end of the CBHI molecule in a loop, where it forms a disulfide bridge with Cys72, and Cys25 is located in a β-sheet, where it forms a disulfide bridge with Cys19 located in a β-sheet (Pro12-Ser20). These two (Cys4 and Cys25) were substituted with proline (Pro) residues in the first CBHI mutant molecule called Δ2. Cys172 forms disulfide bridges with Cys210, and Cys176 with Cys209. These two cysteine residues are located close to the active site residues Glu212, Asp214 and Glu217 in the middle of the tunnel of the CBHI sandwich structure. Cys172 and 176 residues were substituted with a Pro residue in addition to previously changed Cys4 and Cys25 residues in the second CBHI mutant molecule called Δ4. The third mutant CBHI molecule contained one additional substitution of Cys261 with proline. Cys261 is located in a β-sheet (Cys261-Trp263), which forms a disulfide bridge with Cys 331, located in a helix.

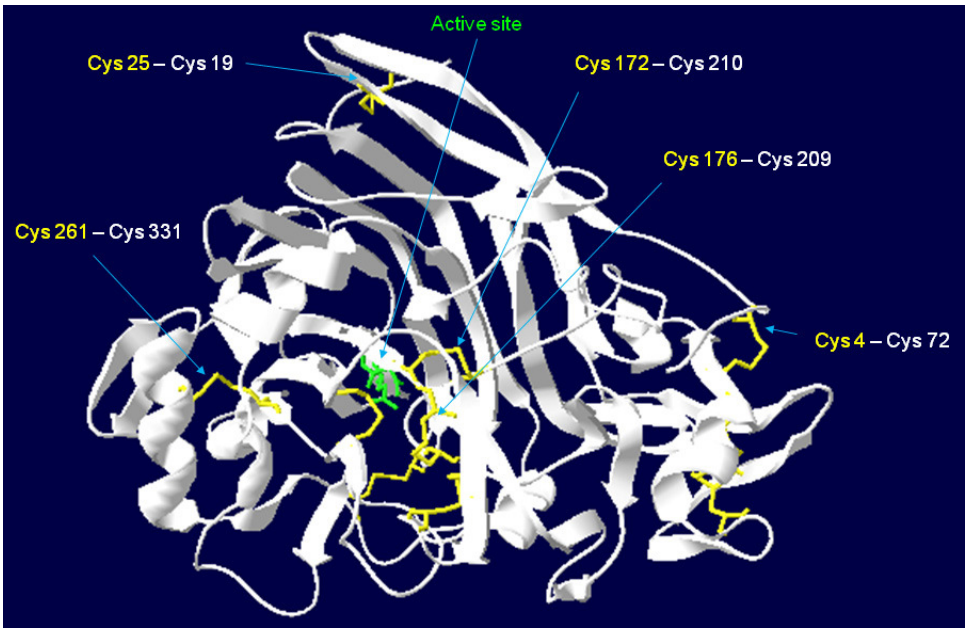

Figure 12. 3D structure of CBHI from *T. reesei* derived from crystallography (1Cel). Cysteine residues which were replaced with a proline in the mutant CBHI core molecules are shown in yellow, and the enzyme active site is highlighted in green (Divne *et al*., 1994).

Note that in the following discussion, "CBHI" refers to the CBHI core molecule. Similarly, "*cbh1*" gene may be used to refer to DNA encoding the the *cbh1* core.

2.2.1 Expression vector

The expression vector CVt (Fig. 13) was kindly provided by Dr. N. Curach (Macquarie University, Sydney). It contained the *cbh1* promoter and signal sequence, DNA encoding the CBHI core region, a gene encoding the fluorescent protein Venus (Nagai *et al*., 2002), *hph* selection marker (Gritz and Davies, 1983) under the control of the *pki* promoter (Schindler *et al*., 1993) that conferred resistance to hygromycin B, and a *cbh1* transcription termination region (Fig. 13).

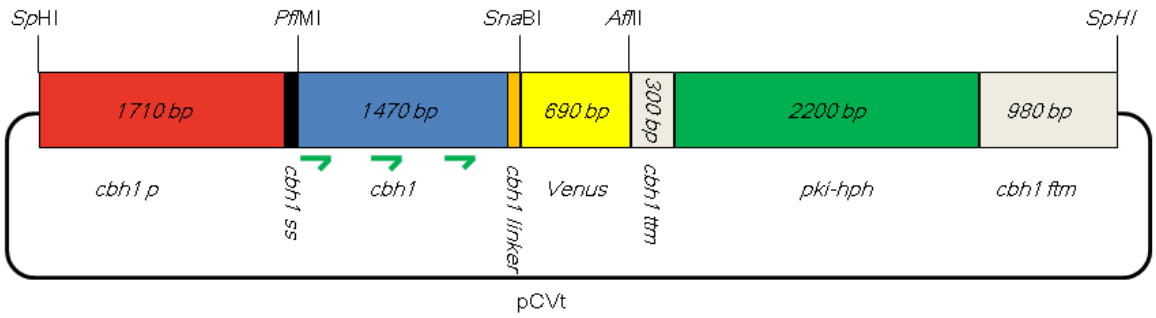

Figure 13. The expression vector pCVt in pUC19 contains the *T. reesei cbh1* promoter (red) and signal seguence (black). DNA encoding CBHI core region (blue), Venus gene (yellow), *cbh1* transcription termination region (ttm: truncated terminator; ftm: full terminator; grey) and hygromycin selection marker gene under the control of the *pki* promoter (green). Sequencing primers are shown in green. DNA featuring 15 bp of *cbh1* linker region is highlighted in orange.

The *cbh1* promoter and DNA encoding the *cbh1* core were sequenced using two forward primers located in the *cbh1* promoter region and three primers targeting the *cbh1* core region (Fig. 13) to confirm the introduced mutations. The CBHI core region composes of residues 1-434 of the 497 residues of the full-length CBHI molecule. More detailed discussion on the predicted effects of the mutations on the structure of the CBHI core can be found in 1.4.1.

2.2.2 Site-directed mutagenesis of the *cbh1* gene

Site-directed mutagenesis was carried out to generate a series of potentially misfolded CBHI protein molecules. The adopted strategy was to interfere with the formation of disulphide bridges between pairs of selected cysteine residues and thereby produce the series of mutated (misfolded) CBHI proteins described above. A pair of complementary primers of 25□35 bases was first designed for each mutagenesis round with sequences for

the amino substitution placed in the primers to change the selected cysteine codons to proline codons in the *cbh1* gene. All primers shown in Table 3 have one or two codons encoding cysteine (UGU or UGC) replaced with the CCG codon (shown red) encoding proline. The codon CCG is one of the most commonly used codons for proline in the *T. reesei* genome (www.kazusa.or.jp/codon/).

The mutated gene fragments were amplified using overlapping PCR. The first PCR round was carried out with the mismatching primers 2, 3, 4, 5, 6, 7 and primer 1 from the start of the *cbh1* gene and primer 10 hybridising to the end of *cbh1* gene (Table 3). The location of each primer is shown in Fig.14. In the second round, PCR products from the first round were used as templates for primers 1 and 10 to amplify the entire *cbh1* core region.

Table 3. Primers used in site-directed mutagenesis of the *cbh1* gene. The codon for proline is shown in red.

Primer number and name	Primer sequence with mismatching amino acid colored red
1. PflMfor.pr	5□-GTTCAGGGGCCACTGCATGGTTTCGAATAGAAAGAG-3□
2. Cys4for.pr	5□-TCGTGCTCAGTCGGCCCCGACTCTCCAATCGGAGACTC-3□
3. Cys4rev.pr	5□-GTCGGGGCCGACTGAGCACGAGCTGTGGCCAAGAAGGC-3□
4. Cys25for.pr	5□-GTCTGGTGGCACGCCGACTCAACAGACAGGCTCCGT-3□
5. Cys25rev.pr	5□-GAGTCGGCGTGCCACCAGACGAGCATTTCTGCCATGTC-3□
6. Cys172/176for.pr	5□-CGGGGTACCCGGACAGCCAGCCGCCCCGCGACC-3□
7. Cys172/176rev.pr	5□- GGGCGGCTGGCTGTCCGGGTACCCCGTGCCGTAC-3□
8. Cys261for.pr	5□-CTTGCGATCCCGATGGCCCGGACTGGAACCCATACCG-3□
9. Cys261rev.pr	5□- GTCCGGGCCATCGGGATCGCAAGTGCCGCCATATCTG-3□
10. SnaBIrev.pr	5□- GTAACCGTCGTGGCCGTTGGGATCGGGATCCATGCATCACTC-3□

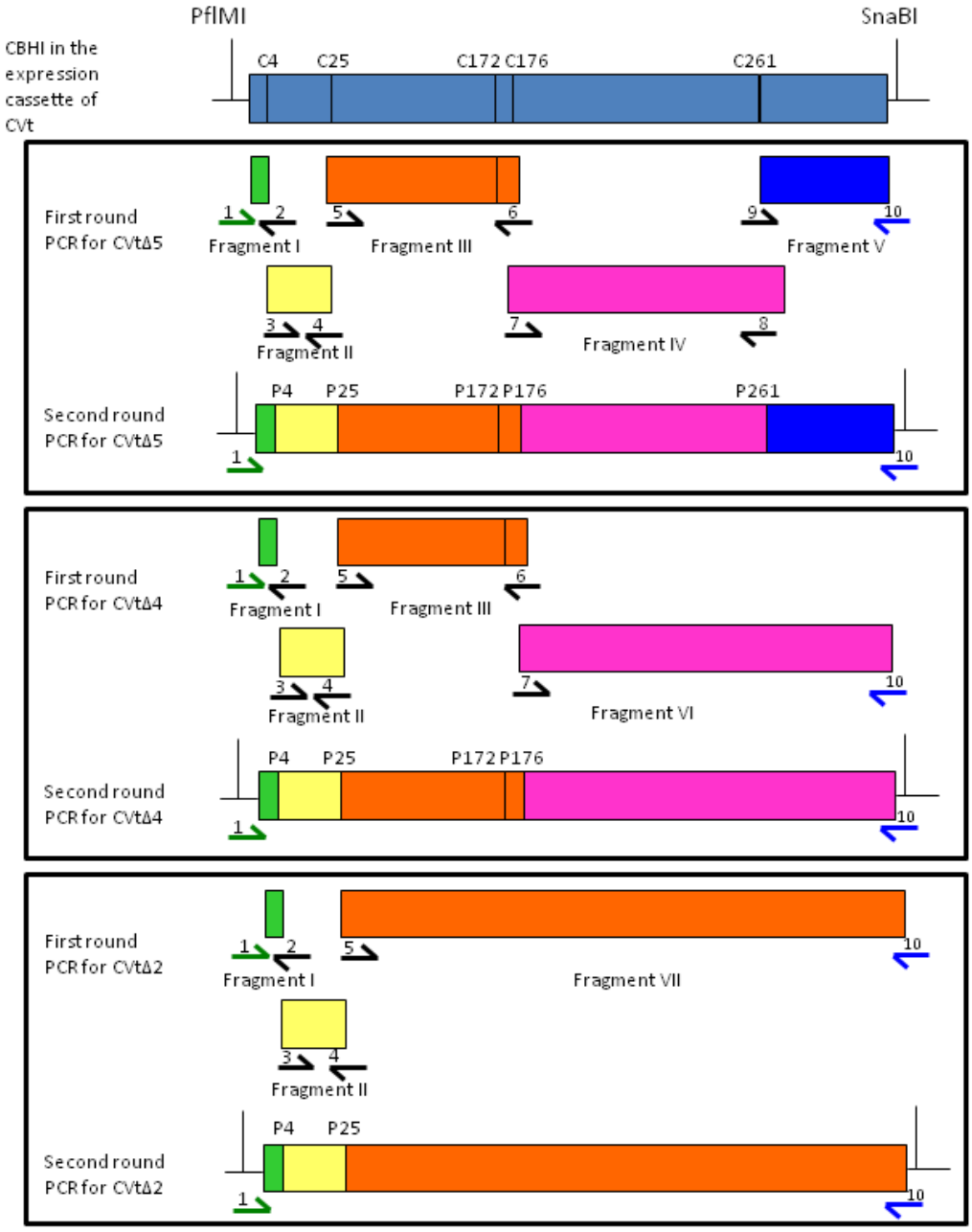

Figure 14. A schematic overview of two PCR rounds for the amplification of the mutant *cbh1* genes. For CVtΔ5, five fragments were amplified separately in the first round PCR using mismatching primers (black) containing five cysteine codons (C) changed to proline codons (P). In the second round PCR, the products from the first round were used as a template for amplifying the entire *cbh1* gene using the primers labelled 1 and 10. CVtΔ4 and CVtΔ2 were generated using a similar strategy as described for CVtΔ5.

In the first round PCR, each fragment, labelled I-VII (Fig. 14), were amplified separately. The template DNA (100 ng of pCVt), 25 µM dNTPs, 100 pmol of primers (forward and reverse: 1 and 3; 2 and 5; 4 and 7; 6 and 9), 1 x buffer, 1 unit of Triple Master polymerase (Eppendorf, Australia) and H_2O to 25 µL were mixed and subjected to the following PCR conditions: 1 x (94°C, 2 min) and 25 x (94°C, 30 s; 60°C, 30 s; 72°C, 1 min). The seven

amplified DNA fragments were then purified using the Qiagen PCR purification kit. To construct Δ5, all seven DNA fragments were mixed together in equal amounts and this mixture was then used as a template for the second round PCR together with 100 pmol of primers (1 and 10), 25 mM dNTP mix, 1 x buffer, 1 unit Triple Master polymerase and H_2O to 50 µL. The PCR reaction conditions were: 1 x (94°C, 2 min), 35 x (94°C, 30 s; 70°C, 20 s; 72°C, 2 min) and 1 x (72°C, 5 min). Similarly, to construct Δ4, DNA fragments I-IV were mixed in equal amounts and subjected to a second round PCR. For the third construct Δ2, a mixture of DNA fragments I-III were used as the template in the second round PCR. Following this, all three amplified full length *cbh1* core genes were checked on a 1 % (w/v) agarose gel and gel-purified with the Qiagen Gel Purification Kit. PCR fragments were then inserted into the PCR 2.1 TA cloning vector (Invitrogen, USA) and transformed into the *E. coli* DH5α using standard protocols (Sambrook and Russel, 2001). The recombinant plasmids containing mutated *cbh1* fragments were purified using the Qiagen Miniprep purification kit and stored at -20^{o}C until used.

2.2.2.1 Introduction of mutated fragments back to the expression vector pCVt

The three mutated *cbh1* fragments (Δ2, Δ4 and Δ5) were excised from the TA-cloning vector by digestion with *Pfl*MI and *Sna*BI and inserted into the expression vector pCVt to generate recombinant plasmids with the mutant *cbh1* genes.

After digestion with the *Pfl*MI and *Sna*BI enzymes, the expression vector pCVt was separated on a 1 % (w/v) agarose gel and purified using the Qiagen Gel Purification Kit following the manufacturer's instructions. Each of the three mutated *cbh1* fragments, Δ2, Δ4 and Δ5 were digested with the same restriction enzymes and inserted into the pCVt plasmid and transformed separately into *E. coli* DH5α cells. Recombinant DH5α cells were selected on LB plates containing 100 µg mL^{-1} ampicillin and grown at 37 ^{o}C overnight. Single colonies growing on the plates were picked and grown in 3 mL LB-medium with 100 µg mL^{-1} ampicillin overnight. Plasmid DNA was then extracted with the Qiagen Miniprep Kit and digested with *Pfl*MI and *Sna*BI to confirm the presence of a correct-size insert. Recombinant DH5α cells were then grown in 250 mL of LB - medium with 100 µg mL^{-1} ampicillin for 16 h and plasmid DNAs extracted by the Qiagen Maxi Prep Kit .

All the recombinant plasmids were sequenced with three sequencing primers featured in Table 4 to confirm the presence of the mutations in the *cbh1* core.

Table 4. Sequencing primers used to check the mutations introduced in the *cbh1* core sequence.

Sequencing primer 1	5□-CAAGTCCCCGGTGACGTACCAAAGCTTATCTTTCTC-3□
Sequencing primer 2	5□-GTCTGGTGGCACGCCGACTCAACAGACAGGCTCCGT-3□
Sequencing primer 3	5□-CGGGGTACCCGGACAGCCAGCCGCCCCGCGACC-3□

The ABI BigDye sequencing kit (Applied Biosystems, USA) was used for sequencing according to the manufacturer□s instructions. PCR was performed with 350 ng of vector plasmid DNA, 25 ng of each primer, and 4 µL of BigDye reaction mixture. The following PCR cycle was performed for sequencing: 1 x (96 ºC, 2 min), 35 x (96 ºC, 30 s; 50 ºC, 30s; 60 ºC, 40 s) and 1 x (60 ºC, 2 min). The PCR product was precipitated with 100 µL of 90 % (v/v) ethanol and 500 µL of 75 % (v/v) ethanol sequentially, centrifuged into pellets at 13000 rpm and air-dried.

DNA sequencing was carried out in the ABI Prism 377 (Applied Biosystems, USA) DNA sequencing facility at Macquarie University. Sequence analyses were performed using the software *Chromas* (Technelysium Pty Ltd, Tewantin, Australia) and BioManager interface provided by the Australian National Genomic Information Service (ANGIS).

Plasmid DNAs were linearised by *Sph*I restriction enzyme digestions (Fig. 13, p. 67), separated on 1 % (w/v) agarose gel and the expression DNA fragments purified with the Qiagen Gel Purification Kit following the manufacturer□s instructions. Expression vectors containing four different CBHI core regions (non-mutated CBHI as CVt, pCVtΔ2, pCVtΔ4 and pCVtΔ5) were stored -20 ^{o}C until used for the transformation of *T. reesei* protoplasts.

2.2.3 Transformation of *Trichoderma reesei*

Transformation of fungal protoplasts with the expression vectors constructed above was carried out with a method modified from Penttilä *et al*. (1987), provided by Dr. Nina Aro (VTT Biotechnology, Espoo, Finland). Protoplast transformation was carried out with the high cellulase-secreting strain of *T. reesei* Rut-C30 (Eveleigh and Montenecourt, 1979).

The *T. reesei* Rut-C30 strain was chosen as the transformation host because it is currently widely used in the studies addressing UPR (Pakula *et al*., 2000; Pakula *et al*., 2003; Saloheimo *et al*., 2004). Therefore, results obtained in this study would be comparable to the information available in the published papers.

2.2.3.1 Transformations with pCVtΔ2, pCVtΔ4, pCVtΔ5 and the non-mutated pCVt DNA

Freshly harvested *T. reesei* conidia were plated onto cellophane discs placed on 20 PDA (PDA, Oxoid, UK) plates and incubated at 28 ^{o}C for 18 h. Young *T. reesei* hyphae were scraped off from the plates and resuspended in a 40 mL lysing buffer 1.2 M $MgSO_4$ - Na_2PO_4 buffer containing 10 mg mL^{-1} of the *T. harzianum* lysing enzymes (L1412, Sigma) and the mixture was incubated at 30 ^{o}C with shaking at 80 rpm for 90 min. Protoplast formation was monitored under the light microscope at 30 min intervals. Protoplasts released from mycelia were harvested and filtered through a sintered glass filter (no. 1) and rinsed with fresh 1.2 M $MgSO_4$ - Na_2PO_4 buffer. An equal volume of 0.6 M sorbitol - 0.1 M Tris-HCl pH 7.5 buffer was added and the mixture centrifuged at 4000 x g at RT for 15 min. Most of the supernatant was removed and the pellet washed three times in a 1.2 M sorbitol-10 mM Tris-HCl pH 7.5 buffer. Purified protoplasts were then resuspended in 300 µL of a 1.2 M sorbitol - 10 mM Tris-HCl - 10 mM $CaCl_2$ pH 7.5 buffer. A sample of protoplasts was removed and a dilution series (10^{-1} to 10^{-7}) was performed before 100 µL samples were plated onto MM plates containing minimal salts, trace elements and 1 M sorbitol as an osmotic stabiliser at pH 5.5 (Penttilä *et al*., 1987) as a regeneration control.

Approximately 1 µg of linearised DNA was added to 200 µL of solution containing 4 x 10^6 protoplasts followed by slow addition of 50 µL of a 40 % PEG 6000-50mM $CaCl_2$-10 mM Tris-HCl pH 7.5 buffer with gentle mixing. After incubation on ice for 20 min, an additional 1 mL of 40 % PEG 6000 - 50 mM $CaCl_2$ - 10 mM Tris-HCl pH 7.5 buffer was added. One hundred µL aliquots of the transformed protoplast suspension were mixed into 10 mL of molten 2.8 % (w/v) MM agar containing 60U mL^{-1} hygromycin B (Calbiochem, USA) and poured onto MM plates. The plates were incubated at 28 ^{o}C for 5-7 d. After incubation, colonies growing on the selection plates were picked and plated onto PDA plates containing 60 U mL^{-1} hygromycin B for a second round selection.

For the rest of this study, the transformant with two mutations in the *cbh1* core gene will be called CVtΔ2, and the transformants carrying four and five mutations in the *cbh1* core gene are labelled as CVtΔ4 and CVtΔ5, respectively. The transformant carrying an intact *cbh1* core gene is called CVt.

2.2.4 Confirmation of DNA integration into the *cbh1* locus by PCR and Southern blotting

A PCR strategy was used initially to confirm homologous integration of the expressed DNAs. Amplification of a 2 kb PCR product with primers hph3.for targeting the 3□end of the hygromycin resistance gene and pkihphint.rev DNA flanking 300 bases downstream of the 3□end of the *cbh1* terminator (Fig.15) would indicate that the gene cassette had integrated into the *T. reesei* genome at the *cbh1* locus. Southern blotting analysis was then employed to confirm the copy numbers and homologous integration of the expressed DNA. For the Southern blotting the genomic DNA was digested with either *Eco*RV or *Bam*HI restriction enzymes (Fig. 15).

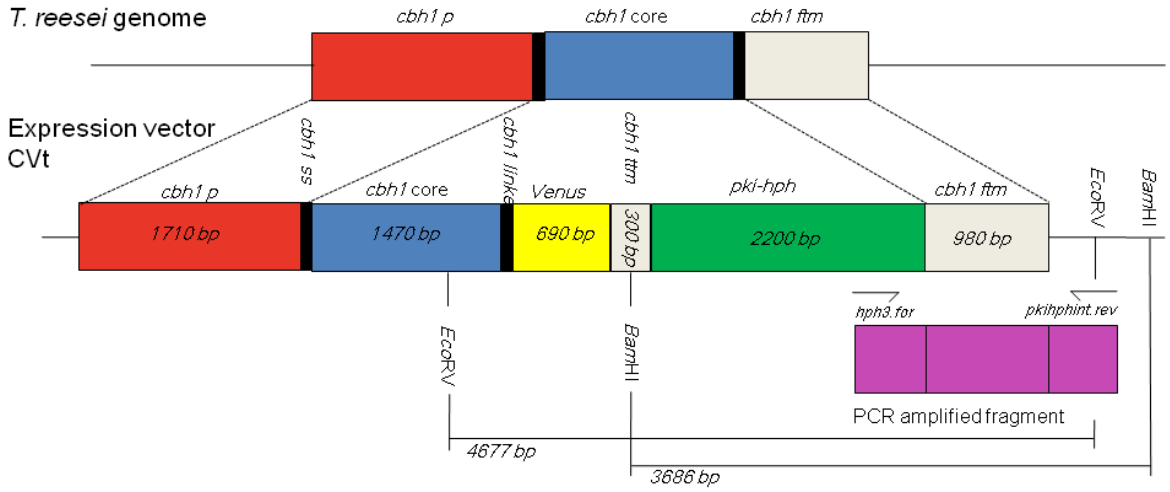

Figure 15. Schematic of homologous integration of an expression cassette into the *T. reesei* genome. The forward primer (hph3.for) was designed to anneal about 500 bp upstream of the stop codon of the *hph* gene and the reverse primer (pkihphint.rev) was designed to anneal at about 300 bp downstream of the end of the terminator of the *cbh1* gene to amplify a DNA fragment of about 2 kb. Genomic DNA was digested with either *Eco*RVor *Bam*HI for Southern blotting. Integration of the transforming DNA into the endogenous *cbh1* locus would produce fragments of 4677bp and 3686 bp in size, respectively.

2.2.4.1 Extraction of genomic DNA

Genomic DNA was extracted from the fungal mycelia following a method adapted from Lee *et al*. (1988). *T. reesei* transformants CVt, CVtΔ2, CVtΔ4 and CVtΔ5 and the non-transformant strain Rut-C30 were first grown for five days on cellophane discs placed on PDA plates. Mycelia were then harvested and ground to a fine powder under liquid nitrogen using a mortar and pestle. After incubation in the lysis buffer (50 mM (w/v) Tris-HCl - 50 mM EDTA (w/v) - 3 % SDS (w/v) - 1 % (v/v) 2-mercaptoethanol) at 65 ^{o}C for 1 h, an extraction step using buffer saturated phenol (Invitrogen, USA) was executed. Further purification of the genomic DNA involved two extraction rounds with chloroform: phenol (1:1) and one extraction with chloroform: isoamylalcohol (24:1). All extractions were carried out using 2 mL Eppendorf Phase Lock gel tubes (Light, Eppendorf, Germany). The concentration and purity of DNA was determined by measuring the absorbance of the samples at 260 nm on a BioPhotometer (Eppendorf) and the 260/280 nm ratio was calculated.

2.2.4.2 Confirmation of homologous integration into the cbh1 locus by PCR

A standard PCR reaction was used for DNA amplification from the genomic DNA to confirm homologous integration of the different *cbh1* genes into the native *cbh1* gene locus. The forward primer hph3.for (5'-TTGGGGAATTCAGCGAGAG-3') was designed to anneal approximately 500 bp upstream from the end of the hygromycin resistance gene. The reverse primer, pkihphint.rev (5'-GACTTTCGACTCGCACGCGAGAAGGGA-3') was designed to anneal approximately 300 bp downstream from the end of the expression cassette (Fig.15). Fifty µL of a PCR reaction included 10 pmol of both primers, 1x High Fidelity buffer with Mg^{2+}, 200 µM dNTPs, 2 U Triple Master Polymerase mix (Eppendorf, USA) and approximately 100 ng of template DNA. Reaction conditions were 1 x 95 ^{o}C, 15 min; 30 x (95 ^{o}C 30 sec; 65 ^{o}C 30 sec; 72 ^{o}C 1 min); 72 ^{o}C 5 min. Amplified DNA products were checked by running on an agarose gel containing 0.5 µg mL^{-1} ethidium bromide at 100 V for 30 min. Using primers (hph3.for and pkihphint.rev) with their locations shown by thin arrows in Fig.15, a 2.0 kb DNA fragment was expected to be amplified confirming homologous integration of the different mutant *cbh1* genes into the native *cbh1* gene locus.

2.2.4.3 Analysis of the copy number of the mutant genes by Southern blotting

Good quality genomic DNA (5 µg) was digested overnight with *Eco*RV or *Bam*HI restriction enzymes in their appropriate buffers. The digested DNA was electrophoresed on a 0.8 % (w/v) agarose gel at 80 V for 90 min with 0.5 µg mL^{-1} ethidium bromide. The double-stranded DNA was denatured into single strands by incubating the agarose gel in a denaturing buffer (0.5 M NaOH - 1.5 M NaCl) for 30 min. After neutralisation in 0.5 M Tris-HCl pH 7.5 - 3M NaCl -buffer for 30 min, the DNA was transferred onto a nylon membrane (Roche, Germany) with a vacuum blotter (BioRad, USA) using 10 x SSC buffer (20 x stock: 3.0 M NaCl, 0.3 M Sodium citrate, pH 7.0) as recommended by the manufacturer. Nucleic acids were fixed onto the membrane by baking at 120 ^{o}C for 20 min. Hybridisation and detection of the *cbh1* gene copy number was then carried out using the DIG Luminescent Detection Kit for Nucleic Acids (Roche, Germany) following the manufacturer's instructions. Membranes were incubated in 20 mL of reconstituted DIG-Easy Hyb Granules (Roche, Germany) in a hybridisation oven at 50 ^{o}C for 1 hour as a pre-hybridisation procedure. Thereafter, the membrane was incubated overnight with a 790 bp *hph*-probe. The *hph*-probe was previously amplified by PCR using primers hph3.for (5'-TTGGGGAATTCAGCGAGAGC-3'), hygroSpeI.rev (5'-CGGGGATCCACTAGTCATGCATCTATTC-3') and PCR DIG-labelling Mix (Roche, Germany) in 100 µL dH_2O and was purified before use at a working concentration of 20 ng mL^{-1}.

After an overnight incubation with the *hph* probe, the blot was washed 2 x for 5 min with 2 x SSC-0.1 % SDS at RT and 2 x 15 min with 0.1 x SSC-0.1 % SDS followed by rinsing with the washing buffer. The membrane was blocked in a 1 % (w/v) blocking solution (Blocking Reagent; Roche, Germany) at RT for 1 h and then incubated for 30 min with the Anti-Digoxigenin-AP antibody (Roche, Germany) diluted 1:10000 in a fresh blocking solution. After washing twice with the washing buffer at RT, the membrane was placed in a plastic bag with approximately 150 µL of ready-to-use CDP-star (Roche, Germany).

The membrane was then exposed to a CL-Xposure film (Thermo Fisher Scientific Inc., USA) and the film was developed with Kodak developing solutions. In case one copy of the expression cassette was integrated into the endogenous *cbh1* gene locus, digestion with the restriction enzyme *Eco*RV will result in a 3.7 kb DNA fragment which can be detected

with a Dig-labelled hygromycin probe. Similarly, a 4.7 kb fragment will be detected by the Dig-labelled hygromycin probe probe when DNA digestion is performed by the *BamH*I enzyme (Fig.15).

2.2.5 Detection of protein secretion

T. reesei Rut-C30 (non-transformant) and transformants CVt, CVtΔ2, CVtΔ4 and CVtΔ5 were grown in a liquid culture inducing the production and secretion of the CBHI protein. Culture supernatants were analysed for secreted proteins and pH change over seven days of growth.

2.2.5.1 Cultivation and total protein secretion

Approximately 10^7 mL^{-1} conidia from each transformed strain and a non-transformant strain were incubated in 50 mL *cbh1* inducing CLS medium containing minimal salts, trace elements, 1 % (w/v) lactose, 1 % (w/v) cellobiose and 3 % (w/v) soybean flour extract, pH 6.5 (Lim *et al*., 2001) was incubated at 28 ^oC with vigorous shaking at 250 rpm for 7 d. Samples were removed after 6, 9, 12, 24 and 36 h, 2, 3, 4, 5, 6 and 7-days incubation. Culture supernatant was separated from the mycelia by centrifugation at 4000 x g at RT for 15 min and stored at +4^oC until used. Total protein concentration of the secreted proteins was determined from the supernatant using a modified Bradford method (Bradford, 1976) with bovine serum albumin used as a standard.

The secreted protein profile of each *T. reesei* transformant and Rut-C30 was analysed using the culture supernatants from samples cultured for 12, 24 and 36 h, 2, 3 and 5 days (from chapter 3.2.4.1) using 12 % (w/v) Criterion Bis-Tris Gels (Biorad, USA) as described by Laemmli (1970). Forty µL of each supernatant was loaded onto the gel and run at 200 V for 45 min, after which the proteins were stained with Coomassie Colloidal Blue G250 and destained with 1 % (v/v) acetic acid.

2.2.5.2 Detection of secreted CBHI and Venus proteins

Culture supernatants were resolved by SDS-PAGE as described above and then electro transferred onto a PVDF membrane (Immobilon□ , Millipore) for Western blotting using a semi-wet transfer system (Invitrogen, Australia). Detection of secreted CBHI proteins from

2 d, 3 d and 5 d cultures was performed with a monoclonal anti-CBHI antibody CI-261 (dilution 1:1000, kindly provided by Roal Ltd, Finland). The primary antibody was visualised with an alkaline phosphatase-conjugated rabbit anti-mouse IgG immunoglobulin (Calbiochem®, USA) and alkaline phosphatase substrate (*FAST*□ BCIP/NBT; 5-Bromo-4□ chloro-3-indolyl phosphate / Nitro blue tetrazolium; Sigma, Australia). Detection of the secreted Venus protein was performed as described above from 1 d, 3 d and 5 d samples, but using the polyclonal anti-GFP (ab290 dilution 1:500, Abcam®, USA) as primary antibody and an alkaline phosphatase-labeled goat anti-rabbit IgG immunoglobulin (Calbiochem®, USA) as secondary antibody.

2.2.6 Analysis of hyphal morphology under the light microscope

Approximately 200 µL of samples containing fungal hyphae from 12 h, 24 h, 36 h and 48 h cultures were washed once with 0.9 % (w/v) NaCl and 10 µL of each sample was observed under the Olympus FV300 laser scanning system equipped with an inverted microscope (Olympus, IX70) and a 100x oil immersion objective. Observation of transmitted light with DIC was carried out using an Argon Laser (488). The images were captured using Fluoview software (Olympus) and analysed using the ImageJ software (Abramoff *et al.*, 2004).

2.2.7 Northern analysis of *cbh1* expression

Northern analysis was performed to explore the abundance and size of specific mRNA transcripts from the non-transformant *T. reesei* Rut-C30, the transformant CVt carrying a non-mutant *cbh1* gene, and the CBHI mutants CVtΔ2, CVtΔ4, CvtΔ5.

Mycelia were collected after 12, 24 and 48 h cultivation in the CLS medium by centrifugation (4000 x g for 10 min at +4 ^{o}C) and then washed in cold diethyl pyrocarbonite (0.1 % (v/v), DEPC)-treated 0.9 % (w/v) NaCl. Samples were then ground under liquid nitrogen using a mortar and pestle anf total RNA was extracted using the Trizol® reagent (Invitrogen, Australia) according to the manufacturer□s instructions. Isolated mRNA was then analysed by Northern blotting.

Digoxigenin (DIG) labelled DNA probes were generated by a standard PCR reaction using a PCR DIG labelling mix (Roche, Germany) as instructed by the manufacturer. Probes were prepared from the Rut-C30 DNA as follows. A 528 bp *gapdh* (jgi Tri2 accession number 119735) DNA fragment was amplified using the primers 5□□ CATCAAGGTCGGCATCAACGGC-3□ (forward) and 5□-GACGGTGGTCATGAGACC CTCA -3□(reverse). A 506 bp probe for *cbh1* (jgi Tri2 accession number 123989) was amplified using primers 5□- GTCTGGTGGCACGCCGACTCAACAGACAGGCTCCGT □ 3□(forward) and 5□-GGGCGGCTGGCTGTCCGGGTACCCCGTGCCGTAC -3□(reverse). Following PCR and separation by electrophoresis, the DIG-labelled probes were excised from the agarose gel and purified using the Qiagen QIAquick Gel Extraction Kit. Concentrations of the DIG-labelled DNA probes were measured and the probes were denaturated at 100 ^{o}C for 10 min before use.

For Northern analysis 20 µg of RNA was diluted with three volumes of RNA loading buffer and denatured at 65 ^{o}C for 10 min. Samples were separated on a 1 % (w/v) agarose/ formaldehyde gel as described in Sambrook and Russell (2001) at 80 V for 90 min. The gel was stained in 0.5 µg mL^{-1} Acridine orange in DEPC-treated H_2O for 10 min and destained with DEPC-treated H_2O for 15 min before imaging under UV illumination for confirmation of equal amount of RNA on each gel. The RNA was then transferred onto a nylon membrane (Roche, Germany) using a vacuum blotter (BioRad, USA) with 10 x SSC buffer (20 x stock: 3.0 M NaCl, 0.3 M sodium citrate in DEPC-treated H_2O, pH 7.0) as recommended by the manufacturer. Nucleic acids were fixed onto the membrane by baking at 120 ^{o}C for 20 min. Hybridisation and detection of mRNAs were carried out using the DIG Luminescent Detection Kit for Nucleic Acids (Roche, Germany) following the manufacturer□s instructions. Membranes were then incubated in 20 mL of DEPC-treated H_2O reconstituted hybridisation solution (DIG-Easy Hyb Granules; Roche, Germany) in a hybridisation oven at 50 ^{o}C for 1 h as a pre-hybridisation procedure. Thereafter, the membrane was incubated overnight with either the *gapdh* or *cbh1* probe at a concentration of 20 ng mL^{-1} in fresh hybridisation solution. After the overnight incubation, the blots were washed and prepared for detection as described in section 2.2.4.3.

2.2.8 Cellobiohydrolase enzyme activity assay

Cellobiohydrolase enzyme activity in the culture supernatants of the different transformants was determined as described by Henriksson *et al*. (2000). Supernatants from the transformant and non-transformant strains that had been cultured for five days were collected, together with mycelia to be used in the microarray experiment (chapter 2.3.2). Fifty µL of each culture supernatant was incubated with 400 µL of the substrate 4□ nitrophenyl β-D-lactopyranoside (1 mg mL^{-1}; N1752, Sigma, Australia) in 50 mM (w/v) citric acid buffer (pH 5.0) in an Eppendorf tube at 37 °C for 1 h. The reaction was stopped by adding 500 µL of 1 M (w/v) Na_2CO_3. The absorbance of the p-nitrophenol released during the reaction was measured on a Fluostar Galaxy plate reader (BMG Labtech, Germany) at a wavelength of 405 nm.

2.3 Gene expression analysis using microarrays

Gene expression studies were carried out with the three *T. reesei* transformant strains carrying mutant CBHI molecules (CVtΔ2, CVtΔ4, CVtΔ5), the non-mutant strain (CVt) and non-transformant strain (Rut-C30) using microarrays. The work flow is shown in Fig. 16. Gene expression profiles of the fungal strains cultured in the CLS medium were determined at 12 h, 24 h and 48 h. The Custom Array™ 12K cDNA slides from CombiMatrix (USA) were used for hybridisation of the cDNAs labelled with the fluorescent dyes Cy5 and Cy3. The intensity of the two dyes on the slides was measured with a Genepix 4000B laser scanner and data were further analysed with the Genepix 6 software. The genes which showed over 1.5 fold expression changes were BLAST-searched against all known proteins in all databases available at http://blast.ncbi.nlm.nih.gov/Blast.cgi and the functions of genes encoding proteins were searched with the Blast2GO program (http://www.blast2go.de). Northern blotting analysis of a group of selected genes was carried out to confirm the microarray results.

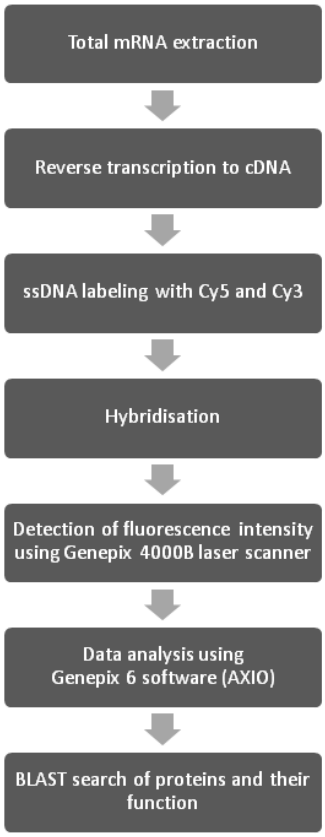

Figure 16. The work flow schematic for microarray studies. The extracted total mRNA from different samples including control was reverse transcribed to cDNA and then labeled with Cy5 or Cy3 fluorescent labels. The intensity of the fluorescence of each spot representing the sample and the control hybridised with a probe immobilised on a slide was measured using the laser scanner. The data were analysed with the Genepix 6 software, and BLAST searches were performed to find corresponding proteins and their functions.

2.3.1 CombiMatrix Custom Array slides

The *T. reesei* microarray experiments were carried out using Custom Array™ 12K slides provided by CombiMatrix (USA). To construct a genome wide gene-specific microarray, 30-45 mer oligonucleotide probes representing the 9129 gene open reading frames (ORFs) derived from the *Trichoderma reesei* sequencing project (http://genome.jgi-psf.org/Trire2/Trire2.home.html) were immobilised on the slides.

Two types of gene arrays were set up. The first, called Set1 contained four to eight replicates of the selected 66 gene probes coding for proteins known to function in the UPR and ERAD pathways (Table 5). These genes were chosen based on a literature search. In cases where there were eight replicates of the same gene spotted on the slide, four of the probes originated from different parts of that gene. The second set, Set2 featuring 9063

gene probes (9129 minus 66 from Set1) was spotted on the same slide as the UPR and ERAD- specific genes but were present in one copy only.

Table 5. Set1 gene probes, their gene number from the second round sequencing of the *T. reesei* genome (jgi Trire) and the relation to UPR (U) or ERAD (E) pathways or housekeeping genes (HK).

jgi Trire	*Gene name*	*U/E*	*jgi Trire*	*Gene name*	*U/E*	*jgi Trire*	*Gene name*	*U/E*
120153	19S RPN1	E	120650	20S ALPHA6	E	122396	NPL4	E
78423	19S RPN2	E	76010	20S ALPHA7	E	131033	PEX4	E
77591	19S RPN3	E	78882	20S BETA1	E	50647	HRD1	U
82512	19S RPN4	E	53446	20S BETA2	E	64023	HRD3	U
68304	19S RPN5	E	58125	20S BETA3	E	121977	PPI	U
54454	19S RPN6	E	78925	20S BETA4	E	52050	CPR3(CYPA)	U
49923	19S RPN7	E	121009	20S BETA5	E	122920	BIP	U
80843	19S RPN8	E	66707	20S BETA6	E	73678	CAL	U
77330	19S RPN9	E	105189	20S BETA7	E	119903	DER1	U
66591	19S RPN10	E	22994	CDC48	E	35465	LHS1	U
12189	19S RPN11	E	121397	SEC61	E	122415	PDI	U
48366	19S RPN12	E	80400	UCH1	E	28928	PRPA	U
73574	19S RPT1	E	21246	UFD1	E	119890	TIGA	U
77587	19S RPT3	E	72606	UBA1	E	64285	EDEM	U
63751	19S RPT4	E	123753	RUB1	E	119664	DOA4	U
23206	19S RPT5	E	47635	SKN7	E	46902	HAC	U
78817	19S RPT6	E	123493	SSM4(DOA4)	E	45242	IRE	U
121343	20S ALPHA1	E	123773	UBC1	E	81164	PTC	U
79825	20S ALPHA2	E	123559	UBC12	E	55362	HSP70	U
73564	20S ALPHA3	E	77732	UBC6	E	119731	HSP60	U
124031	20S ALPHA4	E	59987	UBC7	E	44504	ACT1	HK
55644	20S ALPHA5	E	55788	UBC7homolog	E	119735	GAPDH	HK

2.3.2 Experimental design for expression of the mutant CBHI molecules

The non-mutant transformant strain CVt was used as a universal control for the entire experiment. The relative expression levels of the cDNAs from CVt were compared to Rut□C30, CVtΔ2, CVtΔ4 and CVtΔ5 cDNAs. Dual labelling of the cDNA was carried out in all experiments so that the cDNA of the control strain CVt was labelled with Cy-3 (green fluorescence) and the cDNAs from Rut-C30, CVtΔ2, CVtΔ4 and CVtΔ5 were labelled with Cy-5 (red fluorescence). In addition, a dye swap between the fluorescent labels Cy-3 and Cy-5 was performed for each cDNA at each time point. All strains were analysed at three time points, after 12 h, 24 h and 48 h cultivation in the CLS medium. The experimental design is shown in Fig.17.

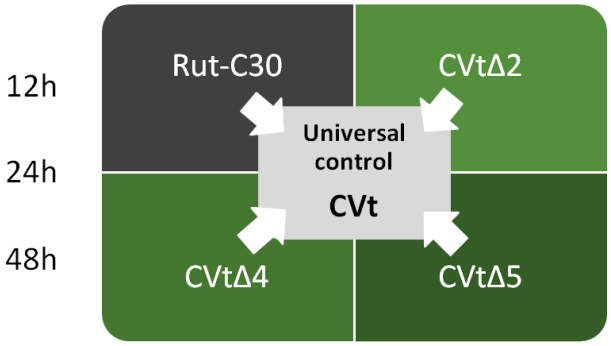

Figure 17. Experimental design for the analysis of the effects of the expression of three mutant CBHI molecules on the UPR and ERAD pathway genes (Set1, Table 5) and on global gene expression (Set2, section 2.3.1). The three time points used for collecting samples were 12 h, 24 h and 48 h. Gene expression in Rut-C30 CVtΔ2, CVtΔ4 and CVtΔ5 was compared to the CVt strain used as a universal control. A dye swap was performed for each sample.

2.3.3 Experimental design for the proteasome inhibition analysis

The non-transformant *T. reesei* Rut-C30, non-mutant transformant CVt and the CVtΔ4 strain expressing a mutant CBHI molecule were chosen for proteasome inhibition analysis. The relative expression of cDNAs was compared between the hyphae treated with the proteasome inhibitor MG132 (Fig.18) and hyphae not treated with the compound. A dye swap was performed for cDNAs from each strain.

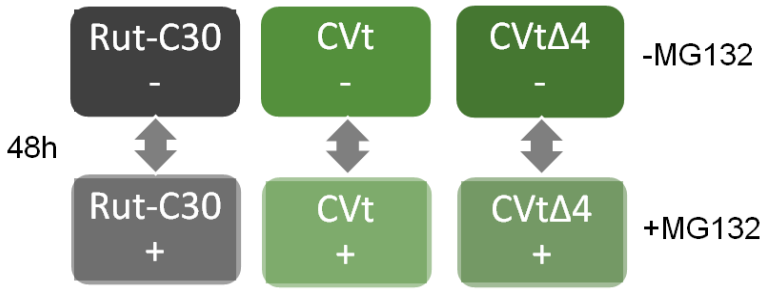

Figure 18. Experimental design for the analysis of gene expression changes in *T. reesei* strains with or without the proteasome inhibitor MG132. Each sample was analysed after 48 h cultivation which featured 46h cultivation without the inhibitor plus a 2h further incubation with (+) or without (-) MG132. A dye swap was performed for each sample.

2.3.4 Cultivation of transformants for total RNA extraction for microarray analysis

Fungal transformants CVt, CVtΔ2, CVtΔ4, CVtΔ5 (from section 2.2.3.1) and the non-transformant Rut-C30 were sporulated and maintained on PDA. Conidia were harvested in 0.9 % (w/v) NaCl-0.01 % (v/v) Tween 80 and 10^8 mL^{-1} of conidia were inoculated into 600

mL of CLS medium in 2 L conical flasks. Cultures were grown at 28 ^{o}C on a shaker at 250 rpm for up to 2 days. A 50 mL sample was collected from each culture after 12, 24 and 48 h of incubation.

For the proteasome inhibition experiment, cultures were grown at 28 ^{o}C on a shaker at 250 rpm for 46 h, after which MG132 (final concentration 60 μM; C2211 Sigma, Australia) in phenylmethylsulfonyl fluoride (PMSF, Sigma, Australia) was added to the cultures. PMSF alone was added to the control cultures. MG132 is a peptide aldehyde that inhibits ubiquitin-mediated proteolysis by binding to and thereby inactivating the 20S proteasome (Lee and Goldberg, 1998). Following collection, mycelia were washed with 0.9 % (w/v) NaCl in DEPC-treated water, quick-frozen in liquid nitrogen and stored at -80 ^{o}C until RNA extraction.

2.3.5 Total RNA extraction, cDNA synthesis, labelling and hybridisation

Preparation of RNA was undertaken by grinding mycelia in liquid nitrogen using a mortar and pestle, and extracting total RNA with Trizol® (Invitrogen, Australia) according to the manufacturer□s instructions. Briefly, after sample lysis with Trizol®, chloroform was added and the samples were subjected to phase separation by spinning at 12000 g for 10 min at +4 ^{o}C. The upper aqueous phase was transferred into a new tube, precipitated with isopropyl alcohol and washed with 75 % ethanol, air dried at RT and eluted in DEPC-treated water. The concentration and purity of the RNA was determined by measuring the absorbance of the samples at 260 nm and 280 nm by a BioPhotometer (Eppendorf, Germany) and ratios 260/280 nm and 260/320 nm were calculated. RNA with a ratio of at least 1.8 indicated that the RNA quality was good. The extracted total RNA was stored at -80 ^{o}C.

The synthesis and labelling of cDNA was performed using the Kreatech Synthesis and Labelling Kit (Kreatech Diagnostics, Netherlands). The cDNAs produced from each sample and from the control were prepared in parallel. Approximately 60 μg of total RNA was used to synthesise the first-strand cDNA with reverse transcriptase. After denaturation with 1 M NaOH, 0.5-1 μg of control ssDNA was labelled with Cy3-ULS and sample ssDNA with Cy5-ULS labels. Labelled ssDNAs were purified with KREApure column (Kreatech

Diagnostics, Netherlands) following the manufacturer's instructions. The concentrations of labelled ssDNAs were determined on a Biophotometer at 260 nm (Eppendorf).

Hybridisation conditions were as determined by the manufacturer (Kreatech Diagnostics, Netherlands). After rehydration with 100 µL of nuclease-free water at 65 ^{o}C for 10 min, the Custom Array® 12K microarray slides (section 2.3.1) were incubated in a prehybridisation solution (PS; 6 x SSPE, 0.05 % (v/v) Tween-20, 20mM (w/v) EDTA, 5 x Denhardt's solution, 5 x salmon sperm DNA 100 ng µL^{-1}, 0.05 % (v/v) SDS) at 65 ^{o}C for 30 min. Equal amounts of labelled ssDNA from the sample and control were combined together and resuspended in the hybridisation solution (HS), which contained 6 x SSPE, 0.05 % (v/v) SDS, 5 x Denhardt's solution, 100 ng µL^{-1}of salmon sperm DNA. The samples in HS were heated to 95 °C for 3 min, vortexed and collected by a brief spin using a microcentrifuge. The labelled samples were applied onto the microarrays slides containing a particular set of probes of interest and the slides were incubated in a hybridisation chamber at 50 °C for 19h. After incubation, the hybridised slides were washed once with each of three washing solutions: solution 1 (6 x SSPE - 0.05 % (v/v) Tween-20), solution 2 (3 x SSPE - 0.05 % (v/v) Tween-20) and solution 3 (0.5 x SSPE - 0.05 % (v/v) Tween-20) for 5 min and then further washed for 1 min in 2 x PBS - 0.1 % (v/v) Tween-20 and 2 x PBS. The slides were then scanned at Cy3 (532 nm) and Cy5 (632 nm) channels with a Genepix 4000B microarray scanner (Molecular Devices, USA) at 5-µm resolution and exported as 16-bit TIFF images for analysis.

Relative fluorescent intensities of the Cy5 and Cy3-labeled samples bound to each probe of the microarray were determined and the data normalised and analysed with Genepix 6 software (AXIO). The log ratio of the median of the two fluorescent dyes in each spot reflects the relative expression level of a particular gene in a sample. Genes were considered to have an expression difference if the expression values showed at least 1.5 x up-regulation or 1.5 x down-regulation across the replicate slides (from the dye swap).

2.3.6 Confirmation of the microarray results by Northern blotting

Northern blotting was performed for several key gene transcripts to confirm the expression differences indicated by the microarray experiments. Twenty µg of the total RNA extracted

from the *T. reesei* transformants CVt, CVtΔ2, CVtΔ4, CVtΔ5 and the non-transformant Rut-C30 (section 2.2.3) was diluted with three volumes of RNA loading buffer and denatured at 65 ^{o}C for 10 min. Samples were separated on a 1 % (v/v) agarose/formaldehyde gel. RNA blotting, hybridisation and detection procedures were as described in section 2.2.4.3.

Specific probes were designed to hybridise to the *cbh1, bip, beta-1, beta-5, alpha-5, alpha-7, rpn-8* gene regions and the *gapdh* gene was used as an internal control. Each gene probe was amplified as described in chapter 2.2.7. Sequences of the primers used in the PCR reactions for amplifying the probes for Northern analysis are shown in Table 6.

Table 6. The sequences of primers used in the PCR amplification of the selected probes for Northern analysis.

Name	***Sequence***
cbh1.for	5□ GTCTGGTGGCACGCCGACTCAACAGACAGGCTCCGT -3□
cbh1.rev	5□ GGGCGGCTGGCTGTCCGGGTACCCCGTGCCGTAC -3□
Bip.for	5□ GGTGGTGGTACCTTTGATGTTTCTC -3□
Bip.rev	5□ ACTGAACCTTGGGGATACGGGTG -3□
Alpha-5.for	5□ AGACCGGGGAATCAACACCTTC -3□
Alpha-5.rev	5□ TTGTGGTATTCGTTTTGCAGCT -□3□
Alpha-7.for	5□ GAACATGCGCAGTTGGCGTCAA -3□
Alpha-7.rev	5□ AAGCTGAGTAGATGCCAGAAGG -□3□
beta-1.for	5□ CTACTCCGACAATGGCGGCTCT -3□
beta-1.rev	5□ CTCAACCTCTGGCCTTGTCAAG -□3□
beta-5.for	5□ CCCTCTGCCTGGCTCCGCGCCG -3□
beta-5.rev	5□ GCCAGGATGCTCCTCCTCTTGCCCA-□3□
rpn8.for	5□ ATGGAACGACGACTTCCCGCAC -3□
rpn8.rev	5□ CGTCCTTCTTCTCGTCCTTGTC -□3□
gapdh.for	5□ CATCAAGGTCGGCATCAACGGC -3□
gapdh.rev	5□ GACGGTGGTCATGAGACCCTCA -3□

Relative differences in the amount of RNA were determined by measuring the intensity of bands using the ImageJ program (Abramoff *et al*., 2004) and normalised against the *gapdh* band intensity.

2.4 Visualisation of the 20S proteasome and interaction between the 20S proteasome and mutated forms of CBHI in *Trichoderma reesei*

Interactions between the 26S proteasome and the main secreted protein CBHI of *T. reesei* were analysed by fluorescence and transmission electron microscopies using dual immunolabelling of sections cut from the hyphae of the non-transformant strain Rut-C30, non-mutant strain CVt and the CBHI mutant strains CVtΔ2, CVtΔ4 and CVtΔ5.

2.4.1 Visualisation of the 20S proteasome in *Trichoderma reesei*

For visualisation of the fungal proteasome inside the hyphae, 10^7 mL^{-1} of conidia from the *T. reesei* Rut-C30 was cultivated in 5 mL of CLS medium for 19h and 2 d. In order to visualise the 20S proteasome inside the hyphae, *T. reesei* hyphae were fixed in 2 % (v/v) paraformaldehyde in 1x PBS for 1 h. After partial degradation of the cell wall by incubation in the lysing buffer 1.2 M $MgSO_4$ - Na_2PO_4 containing 10 mg mL^{-1} of the *T. harzianum* lysing enzyme (L1412, Sigma, Australia), hyphae were washed with 1 x PBS. Samples were then incubated in ice cold methanol for 8 min at □20 ^oC and washed with 1 x PBS twice. Samples were incubated in a blocking solution (10 % (v/v) FBS in 1 x PBS) for 30 min and thereafter with the primary antibody, polyclonal yeast anti-20S-ab (PW9355, Biomol Int., USA) for 1 h. After further washing of five times with 1 x PBS, the samples were incubated with the secondary antibody Alexa Fluor® (Molecular probes, Australia) for 1 h in the dark. After washing five times with 1 x PBS, samples were incubated with the nuclear stain To-Pro3 (Molecular probes, Australia) for 10 min. After washing with 1 x PBS the hyphae were placed on a microscopy slide, allowed to settle down for 10 min, and mounted with Gelmount (Proscitech Australia). Samples were then inspected at 488 nm excitation and 515 nm emission wavelenghts using an Olympus FV300 laser scanning system equipped with an inverted microscope (Olympus, IX70) and a 100 x oil objective.

2.4.2 Sample preparation for sectioning

Conidia (10^8) of the *T. reesei* transformants CVt and CVtΔ2, CVtΔ4and CVtΔ5 (from chapter 3.2.2), and the non-transformant Rut-C30 were inoculated into 5 mL of CLS medium and cultured at 28 ^oC on a shaker at 250 rpm. Samples were then collected after 48

h incubation by centrifugation at 4000 x g at RT for 15 min and 200 µL of fungal hyphae was washed three times with PBS. The hyphae were then fixed with 4 % (w/v) paraformaldehyde -1 % (v/v) glutaraldehyde at +4 °C overnight. After 3 x washes with PBS, the hyphae were dehydrated using an increasing series of ethanol (50 % to 100 %). Following the 100 % ethanol dehydration step, the hyphae were infiltrated using an increasing series (1:1 to 3:1) of LR White Resin and finally left at +4 °C overnight with 100 % resin. The hyphae embedded in White resin were trimmed and cut to 1 µm thin sections for fluorescence microscopy and 0.7 µm ultrathin sections for TEM.

2.4.3 Fluorescent labelling and co-localisation of the 20S proteasome and the mutated CBHI

The thin resin sections (1 µm) of CVt, CVt Δ4 and Rut -C30 hyphae, placed on the microscopy slides were incubated for 30 min in the blocking solution (10 % (v/v) FBS in PBS) and washed in incubation buffer (0.1 % (v/v) fetal bovine serum, FBS in PBS) for 3 x 5 min. The sections were then incubated in the primary antibodies, mouse mAB against CBHI CI-261 (1:50, final concentration 185 mg mL^{-1}, provided by Roal Ltd, Finland) and rabbit polyclonal antibody against the yeast 20S (1:50; PW9355, Biomol Int., USA) for 1 h. The sections were then washed with the incubation buffer for 6 x 5 min and incubated further with the secondary antibodies, anti-mouse Alexa Fluor® 488 and anti-rabbit Alexa Fluor® 546 (Molecular probes, Australia) for 1 h, respectively. Both secondary antibodies were used at a final dilution of 1:400 in the incubation buffer. After washing for 6 x 5 min with the incubation buffer and with PBS for 3 x 5 min, the sections were mounted in Gelmount (Proscitech, Australia) and inspected on an Olympus FV300 laser scanning system equipped with an inverted microscope (Olympus, IX70, Japan), using a 100 x oil objective. Alexa Fluor® 488 was detected at 488 nm excitation and 515 nm emission wavelengths, and the detection of anti-rabbit Alexa Fluor® 546 was carried out with excitation at 543 nm and emission at ≥ 565 nm.

2.4.4 Immunolabelling of the target proteins for transmission electron microscopy

In order to localise the CBHI molecules and the 20S proteasome in the fungal hyphae and to visualise potential interaction between them, 70 nm ultrathin sections of CVt, CVt Δ4 and

Rut-C30 were mounted onto 300 mesh nickel grids. The sections were first incubated in 0.05 M (w/v) glycine (Sigma) in PBS for 15 min, followed by incubation in the blocking solution at RT for 30 min. After washing for 3 x 5 min with the incubation buffer, the sections were incubated with the primary antibodies, mouse mAB against CBHI CI-261 (1:50, final concentration 185 mg mL^{-1}, provided by Roal Ltd, Finland) and rabbit polyclonal antibody against the yeast 20S (1:50; PW9355, Biomol Int., USA) for 1 h. This step was followed by 3 x 5 min washes in the incubation buffer. Control sections were run without the incubation step in the primary antibodies. All sections were then incubated in the secondary antibodies, 5-nm gold goat anti-rabbit conjugate (BioCell Intl., UK), diluted 1:150 in the incubation buffer and 10-nm gold goat anti-mouse conjugate (BioCell Intl., UK), diluted 1:100, for 1 h. The sections were then washed in the incubation buffer for 6 x 5 min and post-fixed in 2 % (w/v) glutaraldehyde in PBS for 5 min followed by 3 x washes in the PBS buffer. The sections were stained with 2 % (w/v) aqueous uranyl acetate (pH 7.4) for 30 min and Reynold□s lead citrate for 4 min and were then examined using a Philips CM10 TEM (Philips, Eindhoven) at an operating voltage of 100 kV. Images were captured on a plate film (Kodak Electron Microscope Film 4489, 8.3 x 10.2 cm) and scanned at 600 dpi (Microteck) to obtain digital images. Quantification of different sized gold particles from the scanned films was carried out using the Cell^P software (Olympus). The size of gold-particles was adjusted based on the information of the manufacturer. The data were statistically evaluated for the mean, standard deviation, Gaussian distribution and p value using InStat 3.1(GraphPad Software Inc., CA, USA). The tests performed using this program were One-way Analysis of Variance (ANOVA) and Tukey-Kramer multiple comparison test; $p<0.05$ was considered significant and $p<0.01$ highly significant.