Hansenula polymorpha as cell factory
Heide, Meis van der

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2002

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Chapter 1

General introduction

Introduction

Man have exploited microorganisms already for a very long time for food production, e.g. wine, beer and bread. These processes are generally regarded as traditional biotechnology. Modern biotechnology became feasible with the availability of the advanced recombinant DNA technology. Using these techniques, microorganisms nowadays are used for the production of foreign, biologically active, proteins for medical, scientific and commercial purposes.

To date a variety of host organisms (cell factories) have been explored to produce heterologous proteins, ranging from bacteria to mammalian cells. Mammalian cells, Eschericia coli and yeast are used e.g. for the production of therapeutic agents which are used for treatment of hemostatic and thrombotic disorders. In humans, deficiency of coagulation factor VIII causes the bleeding disorder hemophilia A. The current treatment of this disease is replacement therapy, using external recombinant factor VIII (rFVIII). Currently, four different rFVIII are available, which all are produced by recombinant mammalian cell lines [1]. One of the most widely used thrombolytic agents is the tissue-type plasminogen activator (t-PA). A recombinant form called reteplase (rPA), which showed an extended half-life in circulation is successfully produced by E. coli [2]. To date, yeast are used for the high-level production of the anticoagulant hirudin [3-5].

Yeast is often preferred for the synthesis of recombinant proteins from higher eukaryotes. Like bacteria, yeast cultivation is fast at low cost and a large variety of techniques exist for the production and manipulation of the proteins produced. Yeast, however, possess the specific advantage over bacterial production systems in that they provide an environment for post-translational modification and secretion (e.g. acylation, phosphorylation, glycosylation, formation of disulfide bonds). These modifications are often essential for the function and/or stability of the produced protein. For mainly historical reasons, Saccharomyces cerevisiae was the first yeast to be exploited for the production of recombinant proteins. Notwithstanding various successful applications, specific disadvantages of the use of this organism were also encountered; these include for instance low yields, undesired hyperglycosylation and instability of the engineered production strain.

Methylotrophic yeast have gained increasing interest for fundamental research and as attractive hosts for the production of biologically active proteins [6, 7]. The use of these organisms in fundamental research is mainly related to studies of peroxisome homeostasis and nitrate assimilation [8-10]. A characteristic feature of peroxisomes is that they are inducible. In Hansenula polymorpha, peroxisomes can be induced by methanol, ethanol, primary amines, D-amino acids, L-α hydroxy acids and purines [11]. For growth on these substrates peroxisome functions are essential. Highest induction levels are observed under methylotrophic growth conditions. Up to 80% of the cytoplasmic volume may be occupied by peroxisomes when H. polymorpha is grown in a methanol-limited chemostat [12]. These peroxisomes contain the key enzymes involved in methanol metabolism namely alcohol oxidase
(AO), dihydroxyacetone syntase (DHAS) and catalase (CAT). The two proteins AO and DHAS may constitute over 60% of total cellular protein under these conditions. This illustrates that the genes encoding these proteins are controlled by very strong promoters. This feature contributed to recognizing \textit{H. polymorpha} as an attractive host for the production of heterologous proteins.

\textbf{Tools for the production of heterologous proteins in \textit{H. polymorpha}}

Efficient and reliable transformation procedures for \textit{H. polymorpha} have been developed. Yeast cells can be transformed using whole-cell methods according to the lithium acetate-dimethylsulfoxide method described by Hill \textit{et al.} \cite{13}, by adding PEG \cite{14} or by electroporation \cite{15}.

An essential tool for the construction of recombinant \textit{H. polymorpha} strains are \textit{Eschericia coli} – \textit{H. polymorpha} shuttle vectors. Important feature of these vectors is the selectable marker genes, which can functionally complement various auxotrophic \textit{H. polymorpha} strains. Commonly used marker genes are \textit{H. polymorpha LEU2}, \textit{URA3}, \textit{TRP3} and \textit{ADE11}, the \textit{S. cerevisiae} genes \textit{LEU2}, \textit{URA3} as well as \textit{Candida albicans LEU2} \cite{16-19}. To date, also genes that confer resistance against G418, phleomycin or zeocin have been successfully used as dominant selection markers in \textit{H. polymorpha} \cite{20, 21}.

Propagation of plasmids in \textit{H. polymorpha} is established by autonomous replicating sequences (ARS’s). These DNA fragments are currently used in various shuttle vectors \cite{19, 22-25}. Generally, shuttle vectors that contain an ARS are not maintained under non-selective conditions. To obtain stable transformants, these vectors can be integrated into the genome of \textit{H. polymorpha} via legitimate or illegitimate recombination. Targeted integration is routinely achieved using linearized plasmids, which results in additive integration or gene replacement depending on a single or double cross-over recombination event, respectively. To date, various ways exist to select for single or multiple integrants. Taking advantage of the fact that the \textit{S. cerevisiae URA3} gene does not fully functionally complement the \textit{H. polymorpha ura3} strain, single-copy integrants will form small colonies whereas multicopy integrants will form large colonies on selective media. Recombinant strains could be generated containing up to 100 copies of the vector integrated into the genome \cite{26}. A similar strategy can be used when the \textit{H. polymorpha leu1.1} strain is transformed with a linearized vector containing the \textit{S. cerevisiae LEU2} gene. Again, single-copy integrants form small colonies whereas multicopy integrants will form large colonies on selective media. Multicopy integrants can also be obtained using the bacterial aminoglycoside 3-phosphotransferase (\textit{APH}) gene, which confers resistance to G418, under control of the \textit{H. polymorpha} glyceraldehyde-3-phosphate dehydrogenase (\textit{GAPDH}) promoter. Here, copy numbers could be controlled up to 50 copies integrated using different concentrations of G418 and several derivatives of the \textit{GAPDH} promoter \cite{20}.

Production systems based on \textit{H. polymorpha} rely on the use of various promoter elements, both inducible and constitutive. Commonly used promoter elements are those derived from genes of the methanol metabolism pathway, which are strongly inducible. The alcohol oxidase promoter (\textit{P}_{AOX}), dihydroxyacetone synthase promoter (\textit{P}_{DAS}) and the promoter of formate dehydrogenase (\textit{P}_{FMD}) gene are fully repressed by excess glucose and are strongly induced by methanol. Derepression
of these promoters is also possible at glucose- or glycerol-limiting conditions (e.g. in carbon-limited chemostat cultures). Thus, basically methanol is not necessary for protein production driven by these promoter elements. One example of a production process in the absence of methanol is the high yield production of the anticoagulant peptide hirudin [27]. Another inducible promoter, which can be used for production purposes, is the *H. polymorpha* amine oxidase promoter (P\textsubscript{AMO}). Amine oxidase (AMO) is the key enzyme in the metabolism of primary amines by *H. polymorpha*. The P\textsubscript{AMO} is repressed by ammonium and induced by primary amines to a level which is up to 20% compared to the P\textsubscript{AOX} induced by methanol. During recent years other inducible promoter elements have been identified in *H. polymorpha*. These include promoter elements derived from genes of the nitrate metabolism, YNT1, YNR1 and YNL1 [10, 28, 29] and the repressible acid phosphatase (PHO1) promoter [30]. These promoters could be used as new control elements for protein production in *H. polymorpha*. Also, it may be advantageous to constitutively express the desired gene product. Recently, such promoters have been identified in *H. polymorpha* e.g. the promoter of the plasma membrane H\textsuperscript+-ATPase (P\textsubscript{PMA1}) [31] and the promoters of the genes encoding the translation elongation factor-1\textalpha (TEF1/TEF2; J.A.K.W. Kiel, unpublished results).

**Production and sorting of heterologous proteins to specific subcellular locations**

The *H. polymorpha* toolbox, detailed above, is necessary for the introduction and expression of foreign genes in this organism. Being an eukaryote, the produced proteins can be sorted to specific subcellular compartments, which in some case is advantageous for optimal results. In order to mediate specific sorting, targeting signals for various cellular compartments have been identified. Signals, which are available for use in *H. polymorpha*, include signals for efficient sorting to peroxisomes (both peroxisomal matrix as well as the peroxisomal membrane), endoplasmic reticulum (ER), the secretory pathway and vacuoles. Production and sorting of heterologous proteins to peroxisomes is detailed below.

In contrast to mitochondria and chloroplasts, peroxisomes do not contain DNA and a protein synthesizing machinery. Hence, all peroxisomal proteins are encoded by nuclear genes. Precursors of these proteins are synthesized on free polysomes in the cytosol and imported post-translationally into the target organelle. Peroxisomal targeting signals (PTS) reside within the mature polypeptide. Firefly luciferase was the first protein in which a PTS was identified namely –SKL-COOH (PTS1) [32]. Later studies revealed that this PTS1 is a general and quite degenerate signal in nature [33-35]. In *H. polymorpha*, the enzymes involved in methanol metabolism, AO, DHAS and CAT, are sorted to peroxisomes via their respective PTS1 variants namely –ARF, -NKL and –SKI [36]. Mouse dihydrofolate reductase, bacterial β-lactamase, a fusion protein consisting of human insulin-like growth factor and a carrier protein and the Green Fluorescent protein (GFP) (Fig. 1) are examples of heterologous proteins which have been successfully produced to high levels and sorted to the peroxisomal matrix, using the PTS1 [36, 37].
Chapter 1

The SKL-motif is not the only PTS that has been identified. Rat thiolase was the first protein in which an alternative PTS was identified called PTS2. The PTS2 consensus sequence \((R/K)(L/V/I)X_5(H/Q)(L/A)\) [38] resides at the N-terminus of the protein. Subsequently, PTS2 sequence was identified in few other peroxisomal proteins like watermelon malate dehydrogenase, yeast thiolases and \(H. polymorpha\) amine oxidase [39, 40]. Also this signal is suitable to direct heterologous proteins to the peroxisomal matrix.

The advantage of the accumulation of proteins in peroxisomes is obvious in cases when the proteins are toxic for the host organism. In this case the peroxisomal membrane forms a barrier, preventing that the proteins can exert their toxic activity to the yeast cell. Another potential advantage of storage in peroxisomes is the absence of modifying enzymes inside the organelle. Undesired modifications like e.g. glycosylation are prevented in this way. Also, proteins, which are sensitive to proteolytic degradation, are protected from proteolysis inside the matrix of the organelle. In conclusion: in case of production of soluble heterologous proteins the peroxisomal matrix is an excellent storage compartment to establish a reliable cell factory.

Up to now a reliable system for the production of heterologous membrane proteins is not available. Such a system is of utmost importance for various reasons e.g. to facilitate rational drug design. The need for this became clear from the observation that many human diseases are caused by malfunctioning of membrane proteins. Also fundamental research, like resolving the structure and functional studies of important membrane proteins, relies on the availability of relative large amounts of biologically active membrane proteins. We showed that the \(H. polymorpha\) peroxisomal membrane might be the niche of choice for overproduced membrane proteins. In \(H. polymorpha\) excessive peroxisomal membranes are easily developed that hardly contain protein components and thus are an excellent storage place for heterologous membrane proteins (Fig. 2) [41]. Sorting to these membranes can be mediated by the targeting signal of the \(H. polymorpha\) peroxisomal membrane protein Pex3p [42]. Promising results for the overproduction of human CCR5 and CXCR4 (HIV-1 co-receptors) have already been obtained [43]. Thus, \(H. polymorpha\) is a precious tool for the production of membrane proteins.
Although peroxisomes can act successfully as a storage compartment for heterologously produced proteins, in many cases secretion is favorable. For instance when modification of the produced protein is necessary for activation and/or functioning (e.g. formation of di-sulphide bonds, glycosylation). Secretion also necessitates recovery of the produced protein from the culture fluid. Using *H. polymorpha*, purification of the secreted protein to homogeneity is relatively simple since conditions can be created that endogenous protein secretion is extremely low. In this way the protein that is produced constitutes >90% of the extracellular protein and can be readily recovered.

The secretory pathway is essential for the production and transport of e.g. plasma membrane proteins, cell wall proteins and extracellular proteins. It involves a complex transport pathway, which initiates with translocation of the protein from the cytosol into the lumen of the endoplasmic reticulum (ER). Below, an overview of the most important processes during secretion in yeast is presented.

**The secretory pathway**

Translocation of proteins into the ER is the first step committed in the secretory pathway. Most proteins which enter the ER contain an N-terminal extension known as signal peptide, signal sequence or leader sequence. Signal peptides contain a tripartite structure. The first part consists of one to five amino acids, which form a positively charged region. The central part is a hydrophobic core of seven to fifteen amino acids, which is finally followed by a polar region of three to seven amino acids. This last polar region is the site for recognition by the signal peptidase complex (SPC). The SPC consists of four different polypeptides namely Spc1p, Spc2p, Spc3p
and Sec11p, which catalyse the cleavage of the signal peptide from the preprotein [44-47]. Next to proteins containing a presequence, some secretory proteins contain a preprosequence e.g. the *S. cerevisiae* mating factor 1α (MF1α) [48]. In these cases, the presequence is cleaved by the ER-borne SPC whereas the prosequence is removed in the late Golgi by the endoprotease Kex2p. Kex2p cleaves the prosequence at the dibasic sequence (KR or RR) [49].

Various heterologous signal sequences are available for use in *H. polymorpha* these include the leader sequences of *S. cerevisiae* MF1α, *S. cerevisiae* invertase, *Schwanniomyces occidentalis* glucoamylase and the leader sequence of the hyperglycemic hormone from the shore crab *Carcinus maenas* [27, 50-53]. In *H. polymorpha*, heterologous proteins carrying their endogenous leader sequences also are effective, e.g. *Humicola insolens* cellulase II and xylanase I [54]. Only recently homologous *H. polymorpha* signal sequences have been identified, namely the leader sequences of acid phosphatase (encoded by *PHO1*), carboxypeptidase Y (CPY) and BiP [30, 53, 55]. Green Fluorescent Protein (GFP) is efficiently sorted to the ER using the N-terminal 30 amino acids of *H. polymorpha* BiP [56]; however, the functionality of these leader sequences to mediate secretion of foreign proteins remains to be established.

Targeting of secretory proteins to the ER is followed by translocation of these proteins into the ER lumen. In yeast, two modes of translocation into the ER exist namely co-translational and post-translational translocation. Co-translational translocation is dependent on the signal recognition particle (SRP-dependent) [57, 58]. As soon as the signal peptide emerges from the ribosome, it is recognised by the SRP. The subsequent association between the signal peptide and the SRP results in a translational arrest. Thereafter the complex is targeted to a receptor on the ER membrane (SRP-receptor). Following binding to this receptor, the elongation arrest is released and the protein will cross the membrane through a protein channel (detailed below) as it is being synthesized [59-61].

Post-translationally translocated proteins however, are synthesized to their mature size in the cytosol before being targeted to the ER. This process is SRP-independent and carries the potential danger that folding of the protein takes place in the cytosol, thus rendering it incompatible for translocation. To prevent premature folding, the function of cytosolic molecular chaperones of the Hsp70 family is required [62, 63]. Two cytosolic chaperones, Ssa1p and Ssa2p have been identified that are required for post-translational translocation [64, 65]. Like other Hsp70s, these chaperones interact with a co-chaperone, Ydj1p (homolog of the bacterial DnaJ). Ydj1p activates the ATPase activity of Ssa1p and suppresses protein aggregation [66, 67]. The signal sequence dictates whether a pre(pro)protein is targeted by the co-translational or the post-translational pathway. SRP preferentially recognises very hydrophobic signal sequences whereas less hydrophobic signal sequences could be targeted via either pathway [68].

After the pre(pro)protein is targeted to the ER, translocation of the polypeptide across the membrane into the ER lumen occurs through specific translocation sites. The translocon is formed by the Sec61p complex and the Sec62p-Sec63p complex. The Sec61p complex, which forms the actual translocation pore, is a heterotrimeric complex consisting of Sec61p, Sbh1p and Sss1p [69-71]. Recently a second trimeric complex was discovered related to the Sec61p complex. This complex designated as the Ssh1p (Sec sixty-one homolog 1) complex consists of Ssh1p, Sbh2p and Sss1p [72]. The Ssh1p complex is presumed to function
exclusively in the co-translational translocation pathway. The Sec62-Sec63 complex consists of four polypeptides: Sec62p, Sec63p, Sec71p and Sec72p [73, 74]. The translocon subunits Sec62p, Sec71p and Sec72p are proposed to function as the signal sequence recognition site and thus form a binding site for pre(pro)proteins before crossing the ER membrane [75]. Sec63p is an integral ER-membrane protein originally found in a subcomplex together with Sec71p, Sec72p and BiP [76].

The protein sequences of Sec63p and BiP indicate that they are members of the family of molecular chaperones. Sec63p contains a luminal segment, which is 43% identical to the bacterial DnaJ protein [77]. The luminal protein BiP, also known as Kar2p, is, like all other Hsp70p’s, an ATPase of which the activity is stimulated by its co-chaperone, Sec63p [78, 79]. The function of BiP in the early translocation process is probably the transition of docking of the pre(pro)protein to the Sec62p-Sec71p-Sec72p binding site to the insertion into the actual translocation pore [75]. The binding of the secretory precursor is ATP independent, whereas ATP hydrolysis is required to release the precursor from the binding site.

The interaction between BiP and Sec63p is also required for completion of the precursor translocation into the ER-lumen. Two models have been proposed for this second stage in translocation. According to one model, BiP pulls the translocating precursor into the ER-lumen, thereby ensuring that the process is unidirectional, by sequential ATP-binding, ATP hydrolysis and dissociation from the precursor [80]. The second model proposes that BiP acts as a molecular ratchet [81]. Recently evidence was presented that substantiated the latter model [82]. Using the S. cerevisiae prepro-α-factor as a substrate, it was shown that multiple BiP molecules bind to the translocating polypeptide in the ER lumen, thereby preventing the backward movement into the protein channel. The authors also showed that BiP function could be taken over by antibodies, which bind to the translocating polypeptide at the luminal side.

Protein folding, N-glycosylation and degradation

Next to its function in the translocation of pre(pro)proteins across the endoplasmic reticulum (ER) membrane, BiP is essential for folding of these proteins [83]. BiP appears to bind to hydrophobic regions of unfolded proteins [84], which prevents premature folding/aggregation of the translocated protein. Strong evidence for the critical role of BiP in promoting the correct folding of newly synthesized Carboxypeptidase Y (CPY) came from experiments using temperature-sensitive BiP mutants [85]. Without the functional chaperone, CPY aggregated and remained in the ER, whereas in wild type cells BiP binding to CPY was only transient and CPY was normally sorted to its target organelle, the vacuole. While the translocation activity of BiP is mediated through interaction with Sec63p, its chaperone function is probably mediated through Scj1p [86, 87].

Recently, a novel Hsp70-related protein, Lhs1p (Cer1p) was detected in the lumen of the ER [88, 89]. This non-essential protein seems to be involved in the translocation of only a subset of proteins into the ER, among these are the prepro-α-factor and the prepro-CPY [89]. In a lhs1 strain, the levels of chaperones, like BiP, are elevated, which indicates a role for Lhs1p in protein folding. However, the exact function of Lhs1p is still a matter of debate [90, 91].
The formation of native disulfide bonds is an important aspect in folding of a subset of secretory proteins. This process is catalysed by the essential gene encoding protein disulfide isomerase (PDI) [92, 93]. Pdi1p contains two domains, homologous to thioredoxin, each containing the CGHC red/ox active site motif [94, 95]. In S. cerevisiae four non-essential Pdi1p homologues have been found, Eug1p, Mpd1p, Mpd2p and Eps1p. All these PDI homologues are multicopy suppressors of a pdi1 deletion [96-99]. However, when EPS1 was expressed under control of the PDI1 promoter on a low copy plasmid, it was not able to rescue the Δpdi1 strain. In fact, Mpd1p was the only PDI homologue that could restore cell viability in a strain completely depleted of other Pdi1p like proteins [100].

A vast majority of secretory proteins undergo an essential protein modification step in the ER, namely asparagine-linked glycosylation (N-glycosylation). This process is one of the most common types of eukaryotic protein modifications and is important in the folding process of secretory proteins [101, 102]. Hence, N-glycosylation is important for the biological function and physico-chemical properties of many secretory proteins [103]. N-glycosylation starts with the formation of the oligosaccharide, GlcNac2Man9Glc3 on a specific polyisoprenol (dolichol) via a pyrophosphate bond. This lipid intermediate, oligosaccharide pyrophospho dolichol, is the substrate for the oligosaccharyltransferase (OST) which catalyses the en bloc transfer of the oligosaccharide onto selected asparagine residues in the consensus sequon, Asn-X-Ser/Thr of the secretory protein, where X can be any amino acid except proline [104, 105]. The OST complex is formed by eight proteins and it can only efficiently glycosylate acceptor asparagine residues located near the signal sequence when this signal sequence is cleaved by the SPC [106, 107]. After transfer of the tetradecasaccharide, three glucose molecules and a specific mannose molecule are removed, by glucosidase I and II and mannosidase respectively, giving rise to a core-glycosylated protein which is transported to the Golgi apparatus [104, 108]. In the Golgi, the glycoprotein is subsequently trimmed and elongated by several mannosidases and glycosyltransferases to obtain a fully maturated protein [109].

Secretory proteins that fail to fold properly are initially retained in the ER-lumen by interaction with ER-resident chaperones [110]. The accumulation of unfolded proteins in the ER induces transcriptional activation of genes encoding the ER chaperones, a process known as the unfolded protein response (UPR). The transcriptional activation occurs via the Hac1 protein. When the UPR is induced, Hac1p is produced, which in turn binds to a conserved promoter element called the UPRE [111]. The UPRE is present in the promoters of e.g. KAR2 (BiP), PDI1 and EUG1.

Many misfolded proteins are degraded via a process known as ER associated degradation (ERAD). By this mechanism, aberrant proteins are transported back to the cytosol where degradation via the proteasome occurs [112]. The majority of the ER-resident and secreted proteins are stable, hence the selectivity of ERAD for specific soluble and membrane proteins should be tightly regulated. The ER-resident molecular chaperones BiP and calnexin have been proposed to function in the selectivity of ERAD [113, 114]. Degradation of malfolded proteins by the cytosolic located proteasome requires retrograde translocation of the protein from the ER into the cytosol. This retrograde translocation is mediated by Sec61p and BiP [114, 115]. Recently, other essential components for retrograde translocation and subsequent degradation were identified, namely Der1p, Hrd1p/Der3p, Hrd3p, Hrd4p and Cue1p [116-120]. The ubiquitin-conjugating enzymes Ubc6p and Ubc7p, which catalyse the
specific covalent attachment of ubiquitin to proteins destined for degradation, also were identified as essential components for ERAD [121]. Ubc6p is an ER-membrane bound protein, whereas Cuelp recruits Ubc7p to the ER-membrane [120]. Very recently, Jarosch et al and Rabinovich et al showed that protein dislocation from the ER requires polyubiquitination and the triple A ATPase Cdc48 [122, 123]. Most likely, additional proteins that play a role in ERAD will be identified as this intriguing process is further analysed.

**Protein transport to the final cellular or extracellular destination**

When a secretory protein has been correctly folded, it is competent to leave the endoplasmic reticulum (ER) and will be sorted to its final destination e.g. the vacuole/lysosome or the extracellular space. Additional targeting signals dictate whether a protein resides within a subcellular compartment like the ER or the vacuole. If such signals are lacking, soluble proteins are transported directly to the plasma membrane and are secreted. Transport of proteins through the secretory pathway is mediated by membrane-bound vesicles [124]. Trafficking of these cargo vesicles involves the formation by budding from the donor membrane, transport, docking and fusion with the target membrane. This mode of trafficking is not unidirectional, both vesicle transport from the ER to the Golgi, transport from the Golgi to the cell surface (anterograde transport) and in the reverse direction (retrograde transport) occur. In order to maintain the membrane-bound compartmental organisation within the cell, proper targeting of the vesicles is necessary [125]. This is achieved by specific membrane-bound proteins called SNARE proteins (soluble NSF attachment protein (SNAP) receptors). SNAREs are present on vesicles (v-SNAREs) as well as on the target membrane (t-SNAREs). During vesicle docking, a SNARE complex is formed, when a v-SNARE binds to its cognate t-SNARE. This complex then recruits the soluble proteins SNAP and NSF and fusion of the vesicle with the target membrane is promoted by hydrolysis of ATP by NSF. This universal docking and fusion machinery called the SNARE hypothesis, proposed that individual v- and t-SNAREs provide the specificity in vesicular sorting [125]. Subsequent studies on the SNARE hypothesis revealed that some SNAREs are shared between different SNARE complexes, indicating that the specificity of vesicle targeting is not exclusively determined by the SNARE complexes [126].

Morphological studies have provided evidence that proteins that are leaving the ER are concentrated at specific domains of the ER [127]. These domains are known as the transitional ER (tER). At these sites, the cargo protein is packed in COPII vesicles. The formation of COPII vesicles has been studied extensively. These studies revealed the sequential steps in vesicle formation. The first step is the interaction of GTP-bound Sar1p with the ER membrane, where the exchange of nucleotide bound to Sar1p is catalysed by Sec12p. In the next step, Sar1p recruits the Sec23p/Sec24p complex, which results in a pre-budding complex. Finally, the Sec13p/Sec31p complex binds to the pre-budding complex, which results in the formation of the vesicle. Sec16p, which is a membrane-bound protein possibly serves as a scaffold for vesicle formation [128]. In reconstituted liposomes however, vesicle formation was observed in the absence of Sec12p and Sec16p [129].

During the formation of the vesicle, v-SNAREs are incorporated in the membrane of the vesicle. ER-derived vesicles are enriched in the v-SNARE proteins
Bet1p, Sec22p and Bos1p. Next, the cargo vesicle is transported to the Golgi apparatus, where it docks on the cis-Golgi membrane. Hereafter, the protein is transported from the cis-Golgi to the trans-Golgi network (TGN). For long it was thought that trafficking along the Golgi was driven by vesicle-mediated transport from one Golgi cisternae to another [130, 131]. More recent observations resulted in an alternative model for Golgi trafficking [132, 133]. This model, known as the cisternal maturation model, proposes that forward transport of proteins take place in cisternal membranes. These membranes assemble and mature in a cis to trans direction. As a consequence of this forward maturation, Golgi resident proteins need to move in the opposite direction. This transport is realized by the retrograde transport of COPI vesicles. The first Golgi cisternae are formed via fusion between COPI, which carry cis-Golgi proteins and COPII vesicles which are formed at the tER sites. During maturation, retrograde transport of COPI vesicles containing cis, medial or trans Golgi proteins are continuously fusing with, and recycling between, the maturing cisternae [134].

The final step in secretion is the delivery of secretory vesicles, which are formed in the TGN, to the plasma membrane. Transport of post-Golgi secretory vesicles requires Myo2p, a class V myosin, and an intact actin cytoskeleton [135]. Myo2p is probably the actin-based motor that brings the secretory vesicles to the sites of exocytosis [136]. Activation of Sec4p, mediated by Sec2p, which both localize on the vesicles, results in an interaction of Sec4p with the post-Golgi vesicle motor [135, 137]. The vesicles are delivered to a multi-subunit complex named the exocyst [138]. The exocyst is composed of Sec3p, Sec5p, Sec6p, Sec8p, Sec10p, Sec15p, Exo70p and Exo84p. Sec15p was shown to interact with activated Sec4p, and to associate with secretory vesicles. These interactions could serve as the connection between the vesicles and the specialized exocytic sites [139]. Hence, the exocyst seems to play a key role in vesicle docking and subsequent membrane fusion.

**Aim and outline of this thesis**

The research described in this thesis was part of a project entitled “Optimization of the secretion of active, oligomeric heterologous gene products by yeast”, which was supported by STW. The main goal of the work presented here was to identify molecular factors, involved in secretion of oligomeric flavo-proteins with the aim to enhance the secretion efficiency. An important protein within the secretory pathway is the endoplasmic reticulum (ER) resident Hsp70 chaperone KAR2/BiP. BiP function is important in both the import of proteins in the ER and for the proper folding of the imported proteins. The cloning of H. polymorpha BiP is described in chapter 2. Furthermore, the analysis of increased BiP levels on the secretion of Aspergillus niger glucose oxidase (GOX), which was used as a model protein, is presented.

The analysis of different aspects of GOX secretion by *H. polymorpha* is presented in chapter 3. These include, the effect of the copy numbers on the secretion level, a comparison of secretion levels using different ER leader sequences and the effect of *PEXI* and *PEX6*, two genes which are essential for peroxisome biogenesis in *H. polymorpha*, on the secretion of artificially produced GOX.
The second model protein which was used during this research was *H. polymorpha* alcohol oxidase (AO). AO is a homo-octameric flavo-protein of approximately 600 kDa, which normally is localized in the peroxisomal matrix.

Chapter 4 describes the efficient targeting of AO to the endoplasmic reticulum (ER) using the leader sequence of the *Saccharomyces cerevisiae* mating factor α (MFα). Secretion of mature, active AO however, was not observed under these conditions. Instead, the ER localized AO was monomeric and lacked its co-factor. Most likely, the presence of the leader sequence at the N-terminus of AO prevents normal FAD binding, which is a prerequisite for oligomerization and activation.

Finally, chapter 5 describes the secretion of active AO protein using the *S. cerevisiae* invertase secretion signal (ISS). This sequence is processed by the secretion machinery and thus can not affect the availability of the FAD-binding site. Also in this case AO exclusively entered the secretory pathway suggesting that as with the MFα, the secretion signal overrules the endogenous PTS1 signal which was still present on the protein. Using the ISS, significant amounts of AO were secreted. The secreted protein was normally processed and enzymatically active.
Chapter 1

References


Chapter 1


Chapter 1


