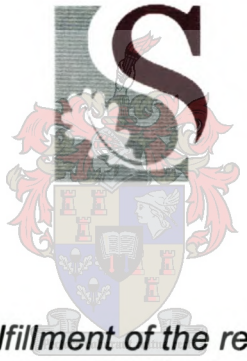
Development of synthetic signal sequences for heterologous protein secretion from *Saccharomyces cerevisiae*

by

JOHAN H. KRIEL



Thesis presented in partial fulfillment of the requirements for the degree of Master of Science at Stellenbosch University

December 2003

Supervisor:
Dr P van Rensburg

Co-supervisors:
Dr RR Cordero Otero
Prof IS Pretorius

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

26 · 11 · 2003

JOHAN H. KRIEL **Date**

SUMMARY

Protein secretion and intracellular transport are highly regulated processes and involve the interplay of a multitude of proteins. A unique collection of thermosensitive secretory mutants allowed scientists to demonstrate that the secretory pathway of the yeast *Saccharomyces cerevisiae* is very similar to that of the higher eukaryotes. All proteins commence their journey in the endoplasmic reticulum, where they undergo amino-linked core glycosyl modification. After passage through the Golgi apparatus, where the remodelling of the glycosyl chains is completed, proteins are transported to their final destinations, which are either the cell surface, periplasmic space or the vacuole.

Proteins destined for secretion are usually synthesised with a transient amino-terminal secretion leader of varying length and hydrophobicity, which plays a crucial role in the targeting and translocation of their protein cargo. Considerable effort has been made to elucidate the molecular mechanisms involved in these processes, especially due to their relevance in a rapidly expanding biotech industry.

The advantages of *S. cerevisiae* as a host for the expression of recombinant proteins are well documented. Unfortunately, *S. cerevisiae* is also subject to a number of drawbacks, with a relative low product yield being one of the major disadvantages.

Bearing this in mind, different secretion leaders were compared with the aim of improving the secretion of the *LKA1* and *LKA2* α-amylase enzymes from the *S. cerevisiae* secretion system. The yeast *Lipomyces kononenkoae* is well known for its ability to degrade raw starch and an improved secretion of its amylase enzymes from *S. cerevisiae* paves the way for a potential one-step starch utilisation process. Three sets of constructs were prepared containing the *LKA1* and *LKA2* genes separately under secretory direction of either their native secretion leader, the *S. cerevisiae* mating pheromone α-factor (MFα1) secretion leader, or the MFα1 secretion leader containing a synthetic C-terminal spacer peptide (EEGEPK). The inclusion of a spacer peptide in the latter set of constructs ensured improved Kex2p proteolytic processing of the leader/protein fusion. Strains expressing the amylase genes under their native secretion leaders resulted in the highest saccharolytic activity in the culture medium. In contrast to this, strains utilising the synthetic secretion leader produced the highest fermentation yield, but had a lower than expected extracellular activity. We hypothesise that the native amylase leaders may function as intramolecular chaperones in the folding and processing of their passenger proteins, thereby increasing processing efficiency and concomitant enzyme activity.

OPSOMMING

Proteïensekresie en intrasellulêre transport is hoogs gereguleerde prosesse en betrek die onderlinge wisselwerking van 'n verskeidenheid proteïene. 'n Unieke versameling van temperatuur-sensitiewe sekresiemutante het wetenskaplikes in staat gestel om die ooreenkoms tussen die sekresiepad van die gis *Saccharomyces cerevisiae* en dié van komplekser eukariote aan te toon. Alle proteïene begin hul reis in die endoplasmiese retikulum, waartydens hulle ook amino-gekoppelde kernglikosielveranderings ondergaan. Nadat die proteïene deur die Golgi-apparaat beweeg het, waar die laaste veranderings aan die glikosielkettings plaasvind, word hulle na hul finale bestemmings, waaronder die seloppervlak, die periplasmiese ruimte of die vakuool, vervoer.

Proteïene wat vir sekresie bestem is, word gewoonlik met 'n tydelike, amino-eindpuntsekresiesein, wat 'n kritiese rol in die teiken en translokasie van hul proteïenvrag speel, gesintetiseer. Heelwat pogings is in hierdie studie aangewend om die molekulêre meganismes betrokke by hierdie prosesse te ontrafel, veral as gevolg van hul toepaslikheid in 'n vinnig groeiende biotegnologiebedryf.

Die voordele van *S. cerevisiae* as 'n gasheer vir die uitdruk van rekombinante proteïene is alombekend. *S. cerevisiae* het egter ook verskeie nadele, waaronder die relatiewe lae produkopbrengs die belangrikste is.

Teen hierdie agtergrond, is verskillende sekresieseine met mekaar vergelyk met die doel om die sekresie van die *LKA1* en *LKA2* α-amilasegene vanuit die *S. cerevisiae*-uitdrukkingsisteem te verbeter. Die gis *Lipomyces kononenkoae* is bekend vir sy vermoeë om rou stysel af te breek en 'n verbeterde sekresie van sy amilasegene vanuit *S. cerevisiae* baan die weg vir 'n moontlike een-stap styselgebruiksproses. Drie stelle konstrukte is gemaak wat die *LKA1*- en *LKA2*- gene onafhanklik onder sekresiebeheer van onderskeidelik hul inheemse sekresiesein, die *S. cerevisiae* paringsferomoonsekresiesein (MFα1) of die MFα1-sekresiesein met 'n sintetiese koppelingspeptied aan die C-eindpunt (EEGEPK), plaas. Die insluiting van 'n koppelingspeptied in die laasgenoemde stel konstrukte verseker verbeterde Kex2p proteolitiese prosessering van die sein/proteïenfusie. Rasse wat die amilasegene onder beheer van hul inheemse sekresieseine uitdruk, het die beste saccharolitiese aktiwiteit in die kultuurmedia getoon. In teenstelling hiermee, het rasse wat van die sintetiese sekresiesein gebruik maak, die beste fermentasie-opbrengs getoon, maar met 'n laer as verwagte ekstrasellulêre aktiwiteit. Ons vermoed dat die inheemse amilaseseine as intramolekulêre begeleiers optree in die vou en prosessering van hul proteïenpassasiers, wat lei tot verbeterde prosessering en ensiemaktiwiteit.

Hierdie tesis word aan my ouers opgedra.
This thesis is dedicated to my parents.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

PROF IS PRETORIUS, Australian Wine Research Institute, Adelaide, for his optimism and devotion throughout my postgraduate studies;

DRS P VAN RENSBURG AND RR CORDERO OTERO, Institute for Wine Biotechnology, Stellenbosch University, for their guidance, patience and encouragement;

NIVI RAMACHANDRAN AND JACO FRANKEN, for critical input;

THE STAFF of the Institute for Wine Biotechnology;

THE NATIONAL RESEARCH FOUNDATION AND WINETECH, for financial support of this study.

PREFACE

This thesis is presented as a compilation of four chapters. Each chapter is introduced separately. Chapter 3 is written according to the style of the journal *Applied Microbiology and Biotechnology*, to which it will be submitted for publication.

Chapter 1 **GENERAL INTRODUCTION AND PROJECT AIMS**

Chapter 2 **LITERATURE REVIEW**
Secretory pathway function in *Saccharomyces cerevisiae*

Chapter 3 **RESEARCH RESULTS**
Optimisation of the secretion of two *Lipomyces kononenkoae* α-amylases in *Saccharomyces cerevisiae*
ADDENDUM
Development of synthetic signal sequences for heterologous protein secretion from *Saccharomyces cerevisiae*

Chapter 4 **GENERAL DISCUSSION AND CONCLUSIONS**

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION AND PROJECT AIMS

1. GENERAL INTRODUCTION AND PROJECT AIMS

1.1 INTRODUCTION

All eukaryotes, including the unicellular microorganism *Saccharomyces cerevisiae*, have secretory machinery composed of membrane-enclosed organelles: the endoplasmic reticulum (ER), the Golgi apparatus and vacuoles (Cleves and Bankaitis 1992; Larriba 1993; Cao and Saier 2003). All proteins commence their journey in the ER, and it is here where they are folded and also undergo distinct modifications, such as glycosylation, phosphorylation and disulphide bridge formation. The translocation of proteins into the ER lumen is perhaps the most important step in the overall secretion process. The molecular mechanism underlying translocation has baffled many researchers for the last 15 years. It was only recently, through the pioneering experimental work done by Simon and Blobel (1991), that the existence of a protein-conducting channel was identified that is involved in the translocation process.

The commercial use of the secretory machinery of yeast is well documented (Bitter et al. 1984; Brake et al. 1984; Zsebo et al. 1986; Eckart and Bussineau 1996; Sudberry 1996). In some ways, *S. cerevisiae* is regarded as an optimal host for the production of heterologous proteins. It offers the ease of microbial growth and gene manipulation found in bacteria, along with the eukaryotic environment and ability to perform various eukaryotic posttranslational protein modifications (Gellissen and Hollenberg 1997; Cereghino and Cregg 1999). *S. cerevisiae* does not produce toxic compounds and the recent development of a range of stable expression vectors has made it an attractive choice for heterologous protein production. The heterologous production of approximately half of the insulin used worldwide underlines the importance of *S. cerevisiae* in biotechnology (Kjeldsen et al. 1998).

Nevertheless, there are disadvantages to using the *S. cerevisiae* secretion system. The expression levels of heterologous proteins are usually low, representing less than 5% of total cellular protein (Peberdy 1994). Many of the heterologous proteins expressed in *S. cerevisiae* are not released into the extracellular medium, but end up in the periplasmic space. This is especially true in the case of heterologous proteins larger than 30 kDa.

In view of its secretion potential, the low product yields obtained from the *S. cerevisiae* secretion system pose an intriguing question (Conesa et al. 2001; Punt et al. 2002). Recent attempts to improve secretion focused mainly on the translocation event and, more specifically, on the components involved in the process. It has been argued that, in a heterologous expression system, proteins may encounter translocation difficulties and that this problem can possibly be overcome by increasing the levels of chaperones (Andrews and Johnson 1996; Brodsky 1998). Unfortunately, overexpression of ER-resident chaperones led to inconclusive and

sometimes even contradictory results, with some proteins showing an improvement in secretion and others not (Tuite and Freedman 1994; Shusta et al. 1998).

The most satisfying results were obtained with the use of synthetic secretion leaders (Clements et al. 1991). The synthetic leaders are based on the *S. cerevisiae* α-factor prepro-peptide (Kurjan and Herskowitz 1982; Gierasch 1989) and were engineered by a combination of semi-random mutations and stepwise optimisation. Proteolytic processing of the leader-heterologous protein fusion is improved by the introduction of a removable spacer peptide (EEAEAEAEPK) between the dibasic Kex2p site and the N-terminal of the heterologous protein (Kjeldsen et al. 1996; 1999). The ability of the synthetic leaders to facilitate secretion was analysed and the results indicated a significant increase in fermentation yield for most of the leaders (Kjeldsen 2000).

The success of the leaders appears to be based on the quality of the interaction with existing chaperones in the ER, rather than on the quantity. Several of the leader peptides have heptapeptides with chaperone (Kar2p)-binding scores of +13 (Blond-Elguindi et al. 1993), indicating a very high probability of Kar2p-binding. An improved interaction of the leaders with Kar2p ensures a quicker entrance into the ER and prolongs the retention of the heterologous protein in the ER lumen (Holkeri et al. 1998), thereby providing additional time for correct folding and secretion (Shinde and Inouye 1993; Eder and Fersht 1995; Kjeldsen et al. 1997).

As we enter the so-called 'Biotech Century', scientists are continuing to engineer yeast cells with increased secretion ability. The efficient secretion of eukaryotic proteins from yeast cells requires more than just targeting the proteins to the ER and directing their translocation into the lumen. There remains a lot to be learned about the yield-limiting steps responsible for the fermentation yields obtained with the current heterologous expression systems. An appropriate place to start would seem to be the at the focus point of this study, namely the secretion leader.

1.2 PROJECT AIMS

The present study forms part of a research project, funded by the NRF and WINETECH, to improve the secretion of heterologous proteins from the microbial host *S. cerevisiae*. α-Amylases from the indigenous species *Lipomyces* have previously been cloned and expressed in *S. cerevisiae* under the direction of its native secretion signals. Here we report on the expression of these genes under different secretion signals. We also compare and quantify the differences in fermentation yield obtained from these expression cassettes.

The specific aims of the study were:

(i) the construction of *S. cerevisiae* expression cassettes consisting of the yeast phosphoglycerate kinase gene promoter ($PGK1_P$) and terminator ($PGK1_T$) and

different secretion signals (native, MFα1 and MFα1 with a C-terminus spacer peptide);

(ii) the cloning of the *L. kononenkoae LKA1* and *LKA2* α-amylase genes under direction of these secretion signals;

(iii) the insertion of the expression cassettes into a set of yeast-integrating plasmids;

(iv) the transformation of a laboratory strain of *S. cerevisiae* with the full set of yeast-integrating plasmids and the concomitant selection of amylolytic *S. cerevisiae* transformants;

(v) quantification of the total extracellular protein and amylolytic activities of the *S. cerevisiae* transformants.

1.3 LITERATURE CITED

Andrews DW, Johnson AE (1996) The translocon: more than a hole in the ER membrane? TIBS 21:365-369

Bitter GA, Chen KK, Banks AR, Lai P (1984) Secretion of foreign proteins from *Saccharomyces cerevisiae* directed by α-factor fusions. Proc Natl Acad Sci USA 81:5330-5334

Blond-Elguindi S, Cwirla SE, Dower WJ, Lipshutz RJ, Sprang SR, Sambrook JF, Gething MJ (1993) Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. Cell 75:717-728

Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, Urdea MS, Valenzuela P, Barr PJ (1984) α-Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 81:4642-4646

Brodsky JL (1998) Translocation of proteins across the endoplasmic reticulum. Int Rev Cytol 178:277-328

Cao TB, Saier MH (2003) The general protein secretory pathway: phylogenetic analyses leading to evolutionary conclusions. Biochim Biophys Acta 1609:115-125

Cereghino GPL, Cregg JM (1999) Applications of yeast in biotechnology: protein production and genetic analysis. Curr Opin Biotechnol 10:422-427

Clements JM, Catlin GH, Price MJ, Edwards RM (1991) Secretion of human epidermal growth factor from *Saccharomyces cerevisiae* using synthetic leader sequences. Gene 106:267-272

Cleves AE, Bankaitis VA (1992) Secretory pathway function in *Saccharomyces cerevisiae*. Adv Microbial Physiol 33:73-144

Conesa A, Punt PJ, Van Luijk N, Van den Hondel CAMJJ (2001) The secretion pathway in filamentous fungi: a biotechnological view. Fungal Genet Biol 33:155-171

Eckart MR, Bussineau C (1996) Quality and authenticity of heterologous proteins synthesized in yeast. Curr Opin Biotechnol 7:525-530

Eder J, Fersht AR (1995) Pro-sequence-assisted protein folding. Mol Microbiol 16:609-614

Gellissen G, Hollenberg CP (1997) Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. Gene 190:87-97

Gierasch LM (1989) Signal sequences. Biochemistry 28:923-930

Holkeri H, Paunola E, Jämsä E, Makarow M (1998) Dissection of the translocation and chaperoning functions of yeast BiP/Kar2p *in vivo*. J Cell Sci 111:749-757

Kjeldsen T (2000) Yeast secretory expression of insulin precursors. Appl Microbiol Biotechnol 54:277-286

Kjeldsen T, Brandt J, Andersen AS, Egel-Mitani M, Hach M, Petterson AF, Vad K (1996) A removable spacer peptide in an α-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in *Saccharomyces cerevisiae*. Gene 170:107-112

Kjeldsen T, Petterson AF, Hach M, Diers I, Havelund S, Hansen PH, Andersen AS (1997) Synthetic leaders with potential BiP binding mediate high-yield secretion of correctly folded insulin precursors from *Saccharomyces cerevisiae*. Protein Expression Purification 9:331-336

Kjeldsen T, Andersen AS, Hach M, Diers I, Nikolajsen J, Markussen J (1998) α-Factor pro-peptide N-linked oligosaccharides facilitate secretion of the insulin precursor in *Saccharomyces cerevisiae*. Biotechnol Appl Biochem 27:109-115

Kjeldsen T, Pettersson AF, Hach M (1999) The role of leaders in intracellular transport and secretion of the insulin precursor in the yeast *Saccharomyces cerevisiae*. J Biotechnol 75:195-208

Kurjan J, Herskowitz I (1982) Structure of a yeast pheromone gene (MFα): a putative α-factor precursor contains four tandem copies of mature α-factor. Cell 30:933-943

Larriba G (1993) Translocation of proteins across the membrane of the endoplasmic reticulum: a place for *Saccharomyces cerevisiae*. Yeast 9:441-463

Peberdy JF (1994) Protein secretion in filamentous fungi – trying to understand a highly productive black box. Trends Biotechnol 12:50-57

Punt PJ, Van Biezen N, Conesa A, Albers A, Mangus J, Van den Hondel, C (2002) Filamentous fungi as cell factories for heterologous protein production. Trends Biotechnol 20:200-206

Shinde U, Inouye M (1993) Intramolecular chaperones and protein folding. TIBS 18:442-446

Shusta EV, Raines RT, Pluckthun A, Wittrup KD (1998) Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments. Nat Biotechnol 16:773-777

Simon S, Blobel G (1991) A protein-conducting channel in the endoplasmic reticulum. Cell 65:371-380

Sudberry PE (1996) The expression of recombinant proteins in yeasts. Curr Opin Biotechnol 7:517-524

Tuite M, Freedman RB (1994) Improving secretion of recombinant proteins from yeast and mammalian cells: rational or empirical design? Trends Biotechnol 12:432-434

Zsebo KM, Lu H, Fieschko JC, Goldstein L, Davis J, Duk K, Suggs SV, Lai P, Bitter GA (1986) Protein secretion from *Saccharomyces cerevisiae* directed by the prepro-α-factor leader region. J Biol Chem 261:5858-5865

CHAPTER 2

LITERATURE REVIEW

Secretory pathway function in *Saccharomyces cerevisiae*

2. LITERATURE REVIEW

2.1 INTRODUCTION

The eukaryotic secretory pathway plays an important role in maintaining the cellular requirements for biochemical compartmentalisation and represents a major aspect of intracellular protein traffic within the eukaryotic cell. It is defined by a set of morphologically different membrane-enclosed organelles that house activities associated with the catalysis of protein transport and the sorting of proteins from the cytoplasm to various intracellular organelles and the cell surface (Cleves and Bankaitis 1992). It is therefore not surprising that the study of protein transport commands a great deal of scientific effort and represents a major discipline in molecular biology.

The basic form of the eukaryotic secretory pathway was elucidated in a series of studies conducted by Palade and his collaborators (Palade 1975). Proteins are synthesised on cytosolic ribosomes, inserted into the lumen of the endoplasmic reticulum (ER), and subsequently transported to the Golgi complex (**Figure 2.1**). Delivery of fully processed and, in some cases, glycosylated proteins to the cell surface or periplasmic space, is the end result of the fusion of Golgi-derived vesicles to the plasma membrane (Zsebo et al. 1986; Peberdy 1994).

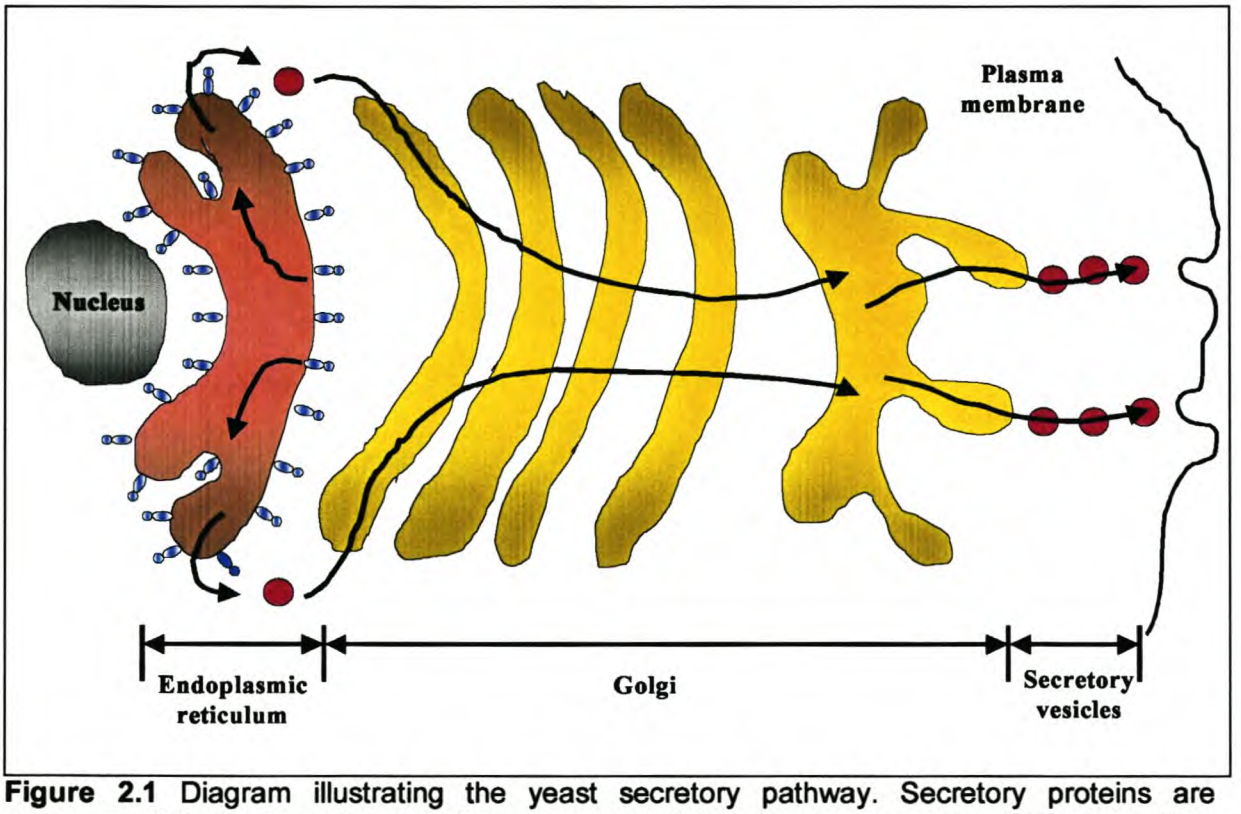


Figure 2.1 Diagram illustrating the yeast secretory pathway. Secretory proteins are synthesised in the cytoplasm and are transported into the lumen of the ER, where they undergo N-linked core glycosylation. After passage through the Golgi complex, where the remodelling of glycosyl chains is completed, glycoproteins are delivered to their final destinations, which are either the cell surface or the vacuole (Tuite and Freedman 1994).

Pioneering work done by Schekman (1985) demonstrated the direct analogy of the yeast secretory pathway to that of the mammalian pathway. *Saccharomyces cerevisiae*, not surprisingly, became an attractive experimental system for the study of protein transport and secretion. Of particular importance was the isolation of a large collection of temperature-sensitive mutants with defects in proteins constituting/controlling the secretory machinery (Novick et al. 1980). The isolation of such mutants was facilitated by the density-enrichment strategy that took advantage of the fact that these yeast mutants became dense, with respect to wild-type cells, upon imposition of the secretory block. These proteins are collectively known as the *sec* mutants and through their cloning and characterization we are beginning to understand the underlying molecular processes of protein secretion.

The main focus of this review is on what is perhaps the most crucial step in the entire secretory process, namely the translocation of newly synthesised proteins into the ER lumen. The elucidation of the molecular mechanisms responsible for translocation has addressed the efforts of many researchers during the past 15 years. Proposed mechanisms have ranged from the idea that the transport of a polypeptide chain occurs directly through the phospholipid bilayer without the participation of membrane proteins, to models in which polypeptides are transported through a hydrophilic or amphiphilic channel formed from trans-membrane proteins (Rapoport 1991). It now seems that a protein-conducting channel does indeed exist. The evidence comes from electrophysiological data and from the identification of membrane proteins as putative constituents of such a channel (Rapoport et al. 1996).

The following sections of this review focus on the various components involved in and factors influencing the translocation of proteins across the membrane of the ER, with special attention to signal sequences and protein targeting, components of the translocation apparatus, as well as the mechanistic aspects of protein translocation. It concludes with a brief overview of the biotechnological advances made in and applications of protein secretion.

2.2 THE YEAST SECRETORY PATHWAY

2.2.1 THE SIGNAL PEPTIDE

Most proteins in eukaryotic cells are synthesised in the cytosol, but many are eventually found at different sites in the cell, or are secreted from the cell. They find their destinations by specific targeting sequences. These targeting sequences have been termed transit, pre- or signal peptides (Gierasch 1989; Schatz and Dobberstein 1996; Zheng and Gierasch 1996); here they will be referred to as signal peptides. In addition to their targeting function, signal peptides (SPs) also facilitate the translocation of the passenger protein. In contrast to these precise functions, SPs show a high degree of diversity in both primary sequence and amino acid composition (Von Heijne 1981). However, a general design exists that appears to be common from prokaryote to eukaryote (Von Heijne 1985; Kaiser et al. 1987). A signal

peptide contains an extremely variable amino-terminal stretch (n region), followed by a larger, highly hydrophobic core region (h region), which in turn is flanked at its carboxyl terminal side by a slightly hydrophilic chain of five to seven amino acids (c region). The n region usually has a net positive charge, which in some cases appears to be important for the correct orientation of the SP into the ER membrane (Larriba 1993). Statistical analysis of a wide spectrum of eukaryotic SPs indicated that (i) the h region adopts an α-helix conformation, and (ii) in most cases, it accounts for about one half of the total length of the SP (Von Heijne 1985). Finally, the carboxyl terminus of the c region contains the cleavage site for the signal peptidase, an endoprotease that removes the SP from proteins entering the ER membrane (VanValkenburgh et al. 1999).

Several studies performed on bacteria have shown that the overall length and hydrophobicity of the h region are crucial factors for SP function (Allison and Young 1988). However, it has been reported that individual amino acids also influence this function, with certain positions in the core region being more important than others (Allison and Young 1989; Hikita and Mizushima 1992). Thus, in addition to length and hydrophobicity, structural properties of the core also seem to be important for SP function (Larriba 1993). Following this general rule, SPs from yeast secretory proteins vary in composition and structure. The most widely studied are those from prepro-α-factor, acid phosphatase and invertase.

2.2.1.1 Prepro-α-factor

Most studies on the secretion of yeast proteins have used the *S. cerevisiae* mating pheromone α-factor. This 13 residue peptide is secreted into the culture medium by *MATα* cells as a prelude to efficient conjugation with cells of the opposite *a* mating type (Brake et al. 1984; Zsebo et al. 1986). Nucleotide sequence analysis of the α-factor-encoding gene, *MFα1* (Kurjan and Herskowitz 1982), indicates that the peptide hormone is synthesised as a 165 amino acid precursor protein (prepro-α-factor) that consists of a 19 amino acid signal peptide (pre region), a 64 amino acid pro region containing three potential N-linked glycosylation sites, and four repeats of the mature α-factor sequence preceded by a spacer peptide (**Figure 2.2**). Processing and modification of the precursor requires the activities of four different proteolytic enzymes.

Prepro-α-factor (ppαF) is translocated into the ER, where the pre region is cleaved by signal peptidase to produce the pro-factor, and N-linked carbohydrate chains are added to the three sites within the pro region (Zsebo et al. 1986; Waters et al. 1988; Caplan et al. 1991). Outer chain carbohydrate modification occurs in the Golgi before proteolytic cleavage by the *KEX2*-encoded endoproteinase (Jiang and Rogers 1999). Kex2p cleaves on the carboxyl side of the Lys-Arg residues at the N-terminus of each spacer, thereby liberating each of the α-factor repeats (Caplan et al. 1991; Bevan et al. 1998; Suzuki et al. 2000). Within the secretory vesicles, processing of the α-factor continues through the action of at least two more

proteolytic enzymes. A *KEX1*-encoded carboxypeptidase removes the Lys-Arg residues at the carboxyl terminus of the first three α-factor repeats, and dipeptidyl aminopeptidase A, the product of the *STE13* gene, removes the spacer sequences at the amino terminus from each repeat to generate the mature α-factor peptides (Zsebo et al. 1986; Caplan et al. 1991).

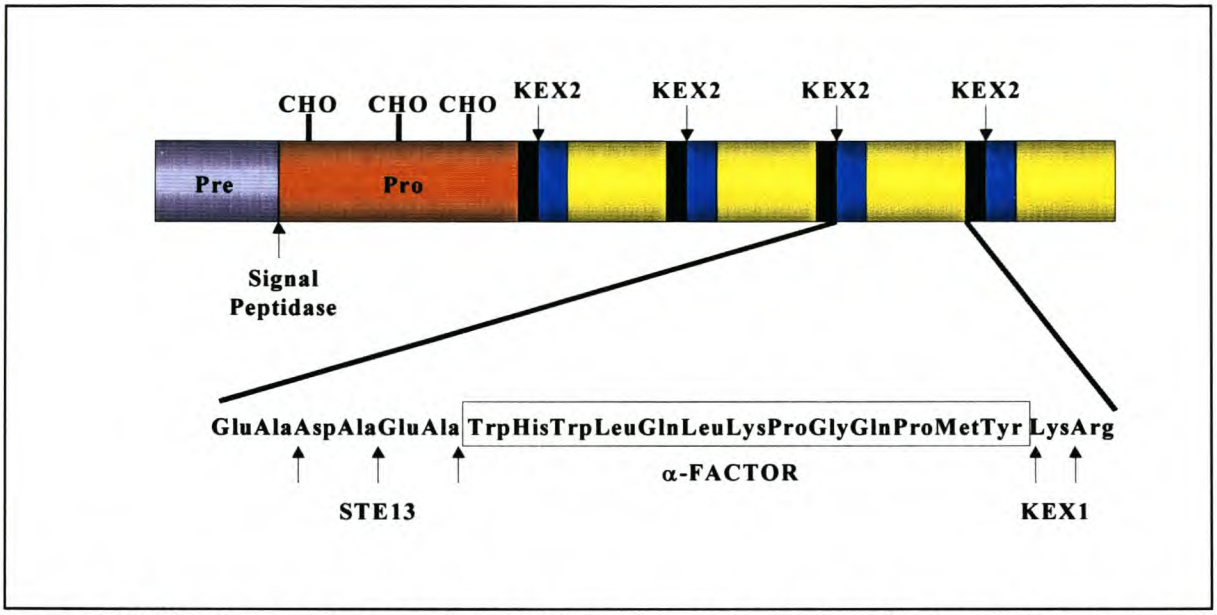


Figure 2.2 Structure and processing pathway of prepro-α-factor. The translation product of the *MFα1* gene has three sites for Asn-linked oligosaccharide addition and sites for proteolytic cleavage by signal peptidase and Kex2p protease. An expanded view (below) of the peptide released by Kex2 cleavage shows the sites for exoproteolytic processing by the products of the *STE13* and *KEX1* genes (Kurjan and Herskowitz 1982).

2.2.1.2 Acid phosphatase

Acid phosphatase (APase) is synthesised with a typical hydrophilic 17 amino acid SP. In the case of APase, deletion of the entire SP as well as four adjacent amino acids of the mature sequence slowed down, but did not prevent, its translocation by yeast microsomes, both *in vivo* and *in vitro* (Haguenauer-Tsapis 1992; Larriba 1993). Apparently, APase is able to translocate in the absence of an SP (Sidhu and Bollon 1987; Monod et al. 1989).

2.2.1.3 Invertase

The invertase SP is rather atypical in comparison to most other SPs; it is extremely hydrophobic, with 16 of its 19 amino acids being hydrophobic, and it lacks the positively charged amino acid that characterises the N-terminal sequence of SPs (Taussig and Carlson 1983). These features are probably central to the capacity of the molecule to overcome mutations in the translocation machinery of the yeast (Ngsee et al. 1989; Larriba 1993; Rothe and Lehle 1998).

2.2.2 THE SIGNAL RECOGNITION PARTICLE

The first step in the targeting of nascent secretory proteins to the ER membrane involves an interaction of the SP with the signal recognition particle (Sanz and Meyer 1988; Rapoport 1992; Arnold et al. 1998). The signal recognition particle (SRP) is a ribonucleoprotein particle consisting of a 7S RNA molecule and six polypeptide units of 9, 14, 19, 54, 68 and 72 kDa (Rapoport 1992; Lütcke 1995). *In vitro* experiments with the mammalian SRP have suggested a function similar to the scheme in **Figure 2.3**. As soon as the SP of a growing polypeptide chain has emerged from the ribosome, it is bound by the SRP. This binding occurs through the methionine-rich M-domain of the 54 kDa SRP subunit (SRP54), which also contains a second domain, the G-domain, with a GTP-binding site (High and Dobberstein 1991; Rapoport et al. 1996). The methionines in the M-domain are found on the outside of three α-helixes and contribute to the formation of a hydrophobic pocket in which the hydrophobic core of the SP is buried. While bound to the SP, SRP also interacts with the ribosome, and this interaction increases the affinity of SRP54 for GTP. The ribosome thus serves as a "GTP-loading factor" (Rapoport et al. 1996). Next, the entire complex, consisting of the ribosome, the growing polypeptide, and SRP containing bound GTP, binds to the ER membrane. This binding involves two important interactions, one between the SRP and its heterodimeric membrane receptor (docking protein), and one between the ribosome and ER membrane proteins (Gilmore et al. 1982). The α subunit of the SRP receptor (SRα) interacts in its GTP-bound form with SRP, which is released from both the ribosome and the SP. The elongating polypeptide is transferred into the ER membrane. SRP, however, remains bound to its receptor until GTP hydrolysis releases it into the cytosol, from where it can begin a new targeting cycle (Connoly et al. 1991; Rapoport et al. 1996). The ribosome-polypeptide-SRP complex shows direct interaction with ER membrane proteins. The question that arises is what prevents SRP-absent complexes from being targeted to the translocation site. An answer to this question comes from observations made by Wiedman et al. (1994). It is suggested that a protein complex, called "nascent polypeptide-associated complex" (NAC), serves as an inhibitor of SRP-absent interaction of the ribosome with the ER membrane. NAC can be removed from ribosomes by high-salt washing; its re-addition restores SRP dependence. It was also proposed that NAC prevented the interaction of SRP with polypeptides lacking an SP; when NAC was removed from the ribosomes, SRP was able to interact with such polypeptides (Wiedman et al. 1994).

SRP has been classified as a polypeptide binding protein, as one of its functions is to keep the bound SP from the rest of the polypeptide chain and thereby prevent premature folding (Rapoport 1992). Another, equally important, function is reducing the rate of translation. SP-SRP association stops further elongation of the growing polypeptide (Mason et al. 2000). After binding to the SRP receptor, elongation arrest is released at the time that the growing chain is transferred into the ER membrane. This assures that the protein crosses the membrane as it is being made, thus

preventing its folding in the cytosol in a translocation-incompatible form (Larriba 1993). However, another SRP-independent, protein-targeting pathway exists. This is demonstrated by the fact that *S. cerevisiae* mutants lacking SRP and an SRP receptor are still viable, although they grow poorly (Rapoport et al. 1996).

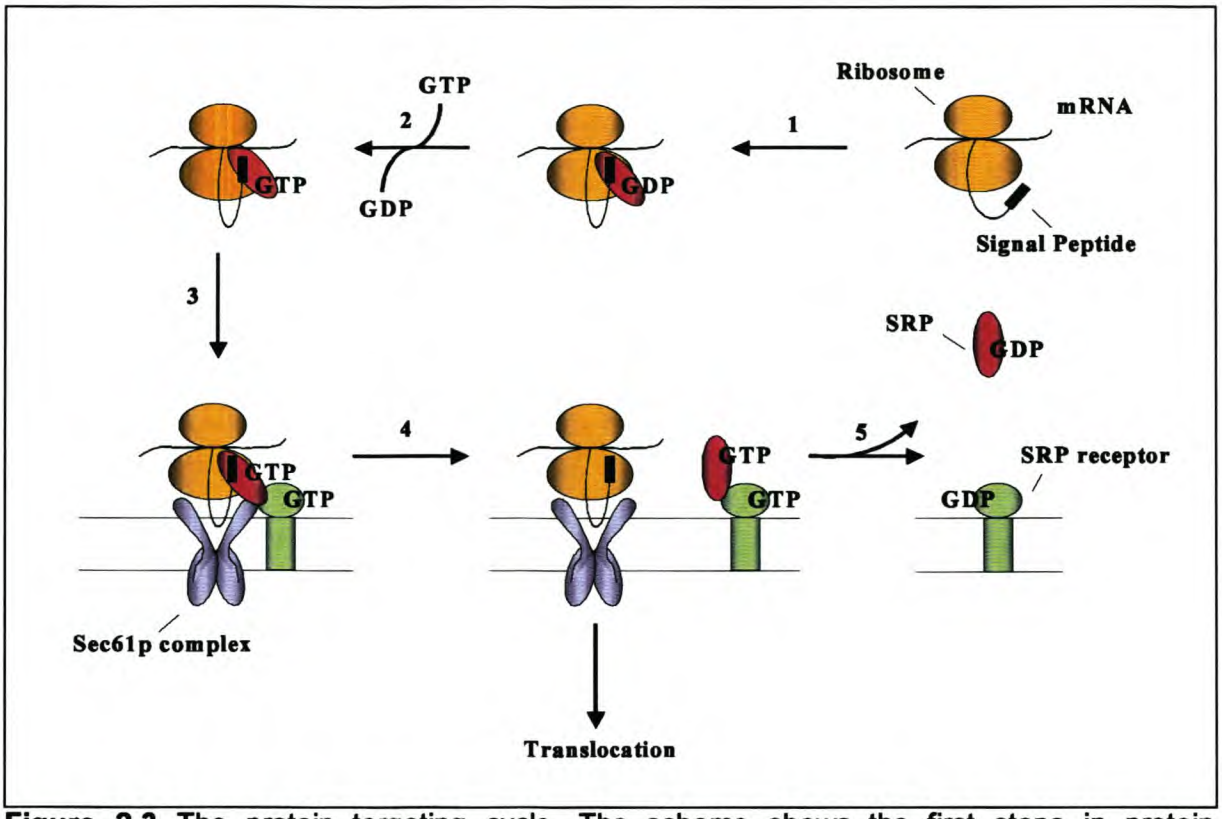


Figure 2.3 The protein targeting cycle. The scheme shows the first steps in protein translocation across the ER membrane. When the signal peptide of a growing polypeptide chain has emerged from the ribosome, SRP is bound to both the nascent chain and the ribosome (step 1). In step 2, GTP replaces GDP bound to SRP54. In step 3, the SRP-ribosome-nascent chain complex binds to the ER membrane by interactions with both the SRP receptor (in the GTP-bound form) and the Sec61p complex. In step 4, SRP is released from both the signal peptide and the ribosome, and the membrane-bound ribosome-nascent chain complex can be translocated. In step 5, GTP is hydrolysed in both SRP and its receptor. SRP is released and can begin a new cycle (Rapoport et al. 1996).

2.3 COMPONENTS OF THE TRANSLOCATION APPARATUS

After targeting, polypeptides are transported through the ER membrane at specific sites called translocons (Andrews and Johnson 1996; Cao and Saier 2003). Genes encoding potential components of the translocon were first identified by the isolation of mutants defective in translocation (Deshaies and Schekman 1987; Deshaies et al. 1989). This method, termed *HOL*$^+$ selection, is based on the use of the normally cytoplasmic enzyme His4p, which converts histidinol into histidine. Modified by fusion to an N-terminal SP, this chimera is targeted to the lumen of the ER and, accordingly, strains carrying this chimeric protein will not be able to grow on minimal media supplemented with histidinol. Alternatively, growing in this medium would constitute a mutation(s) in the translocon apparatus.

In this section, the known components of the translocon will be listed and their features will be discussed.

2.3.1 THE SEC61p/SECYEGp COMPLEX

Every living organism examined so far, including all the organisms with a completely sequenced genome, contain a heterotrimeric membrane protein complex, called Sec61p in eukaryotes and SecYEGp in prokaryotes (Deshaies et al. 1991; Rapoport et al. 1996; Cao and Saier 2003). Sec61p was first discovered in genetic screens for translocation defects in *S. cerevisiae*. A mammalian homolog of Sec61p, Sec61α, was identified. Its amino acid sequence is 56% identical to that of the yeast protein. In addition, both proteins have significant sequence similarity with the SecYp of bacteria. All three proteins have identical topologies, containing 10 membrane-spanning regions (**Figure 2.4**).

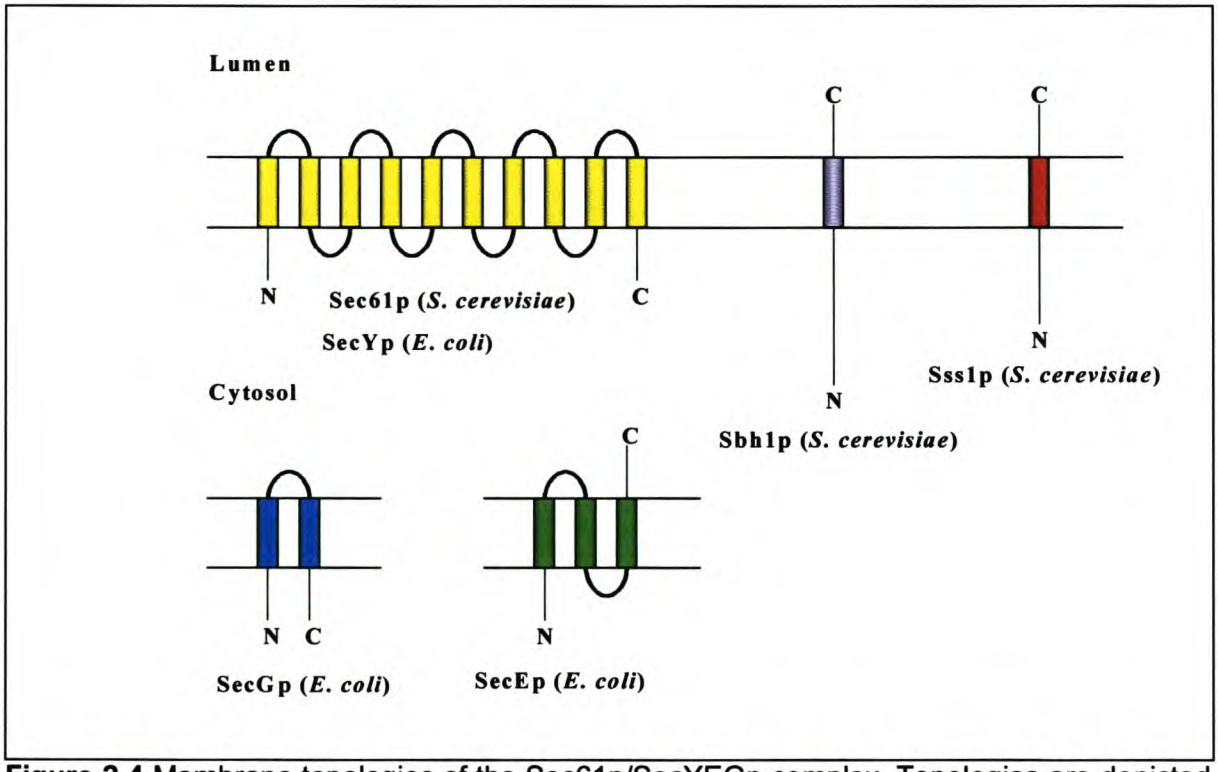


Figure 2.4 Membrane topologies of the Sec61p/SecYEGp complex. Topologies are depicted in a schematic manner and the relative sizes of the proteins are not drawn to scale (Rapoport et al. 1996).

Several hydrophilic amino acids within these membrane-spanning regions are conserved, suggesting that they may be essential for a hydrophilic environment within the membrane (Rapoport 1992).

The smallest subunit of the trimeric complex is called Sss1p in *S. cerevisiae* and SecEp in bacteria. Sss1p was originally found as a suppressor of temperature-sensitive mutations in Sec61p. Both proteins span the membrane once via a segment close to the C-terminus (see **Figure 2.4**). Sss1p and SecEp are both encoded by essential genes (Rapoport 1992; Mothes et al. 1998).

The intermediate-sized subunit of the complex is called Sbh1p in *S. cerevisiae* and SecGp in bacteria. The bacterial subunit shows no obvious similarity to the eukaryote homologue. Sbh1p spans the ER membrane once via a C-terminal segment, while a comparison of SecGp sequences of various bacteria suggests that it spans the membrane twice (see **Figure 2.4**). Neither Sbh1p nor SecGp is essential for cell viability (Rapoport 1992).

2.3.2 THE SEC62/SEC63p COMPLEX

The Sec62/Sec63p complex contains four subunits: Sec62p, Sec63p, Sec71p and Sec72p. Experimental evidence points to an association of the complex with the trimeric Sec61p complex to form a heptameric complex. Both Sec62p and Sec63p are encoded by essential genes and span the ER membrane bilayer two and three times respectively (Feldheim et al. 1992; Rapoport et al. 1996). There is more Sec63p than Sec62p in yeast cells and a dynamic interaction is proposed. Sec63p has a lumenal segment homologous to a portion of DnaJ, the *E. coli* heat-shock protein (**Figure 2.5**). It is via this segment that DnaJ interacts with its partner DnaK. Similarly, Sec63p interacts trough its J-domain with the lumenal chaperone

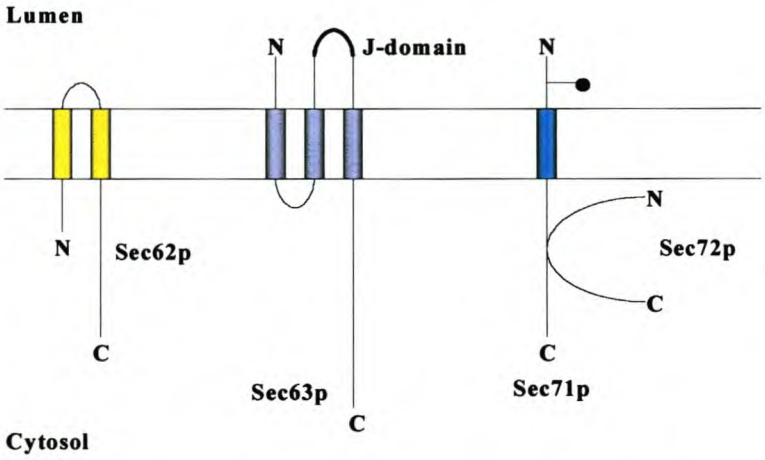


Figure 2.5 Membrane topologies of the Sec62/Sec63p complex. Topologies are depicted in a schematic manner and the relative sizes of the proteins are not drawn to scale. The black dot indicates N-linked carbohydrate chains. The thick line in Sec63p indicates the J-domain, a region of homology to DnaJ, which interacts with the chaperone Kar2p (Rapoport et al. 1996).

Kar2p (yeast BiP), another member of the Hsp70 family. Evidence for such an association comes from genetic data, demonstrating the lethality in a *sec63, kar2* haploid mutant, as well as from the isolation of a complex containing both these proteins (Rapoport 1992).

Sec71p and Sec72p were found in genetic screens as proteins associated with Sec62p and Sec63p. Sec71p is a glycoprotein that spans the membrane once; Sec72p, on the other hand, is located on the cytoplasmic side of the ER membrane

(**Figure 2.5**). Deletion of the Sec71p-encoding gene leads to the additional absence of Sec72p, suggesting a possible interaction with each other. None of the two gene products is essential for cell viability.

2.3.3 THE TRANSLOCATING CHAIN-ASSOCIATED MEMBRANE PROTEIN

Another component of the translocon is the *t*ranslocating chain-*a*ssociated *m*embrane (TRAM) protein. It was identified through chemical crosslinking experiments that showed growing chains of secretory proteins to be crosslinked to TRAM, immediately after their transfer from SRP into the ER membrane (Rapoport 1992; Voigt et al. 1996). Cloning the corresponding genes from organisms as divergent as *Caenorhabditis elegans* and mammals, suggest that TRAM spans the ER membrane eight times and that it has a cytoplasmic tail of about 60 amino acids (**Figure 2.6**). Most of the amino acids in the membrane-spanning region are either hydrophilic or charged. The TRAM protein is about as abundant as ER-bound ribosomes, suggesting its presence at each translocon (Rapoport et al. 1996).

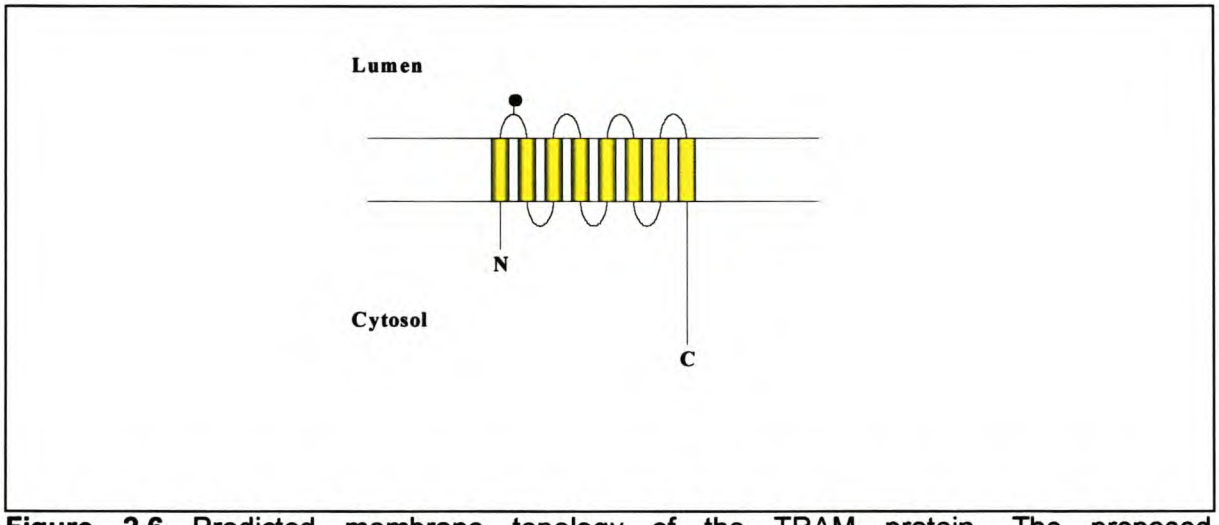


Figure 2.6 Predicted membrane topology of the TRAM protein. The proposed topology is based on various prediction algorithms. The black dot indicates N-linked carbohydrate chains (Görlich et al. 1992).

The effect of TRAM on the translocation of secretory proteins was tested in reconstituted proteoliposomes and it was shown to be required for the translocation of most polypeptides (ppαF and pre-β-lactamase), but not for all (preprolactin). One explanation offered for this differential effect of TRAM on secretory proteins is that it is only required for proteins with weak SPs, as those with strong SPs are transferred into the translocon through direct interaction with Sec61p (Görlich et al. 1992; Rapoport 1992).

2.3.4 ENZYMES IN THE TRANSLOCON

The eukaryotic translocon contains two enzymes that catalyse co-translational modifications of the polypeptide chain: the signal peptidase and the

oligosaccharyltransferase. Both are unique enzymes in that they are as abundant as their substrates (Larriba 1993; Rapoport et al. 1996).

Signal peptidase (SPase) generally cleaves the SP at a site that has small aliphatic amino acids at positions -1 and -3. It is an important reaction, as the rate of protein export from the ER is considerably hastened by SP cleavage, even though translocation into the ER lumen does not depend on SPase (Cleves and Bankaitis 1992). The specificity of the proteolytic reaction is highly conserved (Larriba 1993), since signal peptidases from both prokaryotes and eukaryotes have been seen to cleave *in vitro* substrates of either origin at the correct site. Böhni et al. (1988) have cloned the SPase-encoding gene, *SEC11*. It predicts an 18.8 kDa protein containing an N-terminal SP and one potential glycosylation site.

The oligosaccharyltransferase is responsible for the transfer of an oligosaccharyl moiety from a dolichol intermediate to Asn-residues located in the sequence context Asn-X-Ser or Thr (with X being any amino acid other than Pro). N-linked glycosylation occurs when this site has reached a distance of at least 12 amino acids from the plane of the lumenal side of the ER membrane and entails the addition of an oligosaccharide unit comprising two N-acetylglucosamine (GlcNac), nine mannose (Man) and three glucose residues. The glucose residues, as well as one mannose residue, are subsequently removed. These steps are common to all eukaryotes.

The oligosaccharyltransferase enzyme has been characterised in both mammals and yeast; it does not exist in prokaryotes. Three subunits have been identified in mammals, namely ribophorin I, ribophorin II and Ost48 (Rapoport et al. 1996). In *S. cerevisiae*, the enzyme was purified as a complex of six subunits, with several of the subunits playing a role in glycosylation *in vivo* (Rapoport et al. 1996). The largest α subunit, Ost1, is a homolog of ribophorin I; the β subunit, Wbp1, is a homolog of Ost48, and the δ subunit, Swp1, is a homolog of ribophorin II (Rapoport et al. 1996). All these proteins span the ER membrane once. The γ subunit, Ost3, and the ε subunit, Ost2, span the membrane four and three times respectively.

2.4 MECHANISTIC ASPECTS OF PROTEIN TRANSLOCATION

The idea of a protein-conducting channel that transiently opens so that polypeptides can move across the ER membrane was postulated a long time ago (Blobel and Dobberstein 1975). Simon and Blobel (1991) supported their postulation by direct evidence provided by electrophysiological experiments. They fused rough microsomes into planar lipids and demonstrated that a large number of ion-conducting channels appear when the growing polypeptides are removed from the membrane-bound ribosomes with the chemical puromycin. These channels were previously plugged by the polypeptides passing through the membrane.

The protein-conducting channel is different from a channel that transports molecules because it must open in two dimensions: perpendicular to the ER bilayer, to allow hydrophilic polypeptides across, and within the plane of the bilayer, to allow

hydrophobic anchors of membrane proteins into the phospholipid membrane (Simon and Blobel 1991; Görlich and Rapoport 1993).

The following section deals with the two distinct pathways involved in protein targeting to the ER membrane: co-translational and posttranslational translocation. Both these pathways make use of the protein-conducting channel, although the precise mechanism of translocation differs (De Nobel and Barnett 1991; Johnson 1993; Brodsky 1998).

2.4.1 CO-TRANSLATIONAL TRANSLOCATION

In yeast, the targeting of secretory proteins to the translocon can occur in two distinct pathways that are distinguished by their dependence on the SRP. The SRP-dependent pathway, or co-translational translocation pathway, as it is more commonly known, requires SRP and its membrane receptor, whereas the SRP-independent pathway requires a separate membrane-bound receptor consisting of the tetrameric Sec62/63p complex (Brodsky et al. 1995; Lyman and Schekman 1996; Young et al. 2001).

The translocation apparatus involved in co-translational translocation has been characterised best in mammalian cells. Reconstituted systems consisting only of highly purified membrane proteins were developed and, surprisingly, proteoliposomes containing only three components, namely the SRP receptor, the Sec61p complex and the TRAM protein, were competent to translocate the proteins tested. The molecular mechanism of co-translational protein transport is only now being unravelled (Rapoport et al. 1996; Mothes et al. 1998). Most mechanistic studies make use of defined translocation substrates, which are produced by translating mRNAs *in vitro* in the presence of SRP and microsomal membranes; this yields ribosome-bound growing chains of varying lengths that are caught at various stages of their transfer through the membrane.

The first step following the targeting by SRP is the binding of the ribosome-nascent chain complex to the translocon of the ER membrane (step 1 to step 2 in **Figure 2.7**). At this stage, step 2, the ribosome-nascent chain complex can be removed from the ER by high salt concentrations, which is indicative of weak electrostatic interactions. The nascent polypeptide is also susceptible to externally added protease. The initial weak membrane contact of the ribosome-nascent chain complex is caused by a direct interaction of the ribosome with Sec61p (Rapoport et al. 1996; Mothes et al. 1998).

Clearly distinct from this initial membrane interaction is the much stronger membrane association that occurs when the polypeptide chain becomes longer (see step 3 in **Figure 2.7**). This form of binding is no longer only due to electrostatic interactions, because the nascent polypeptide cannot be removed from the membrane even at high salt concentrations. A tight seal between the ribosome and the membrane is also suggested by the fact that the nascent chain is now resistant to proteolysis.

The transition from the weak to the tight association requires a longer nascent polypeptide chain with a functional SP (Bird et al. 1987). The nascent chain needs to be about 15 amino acids longer for signal recognition in the membrane than for interaction with the SRP (Rapoport et al. 1996; Young et al. 2001) The polypeptide chain is inserted into the membrane in a loop structure and the SP is able to reach its binding site within the membrane.

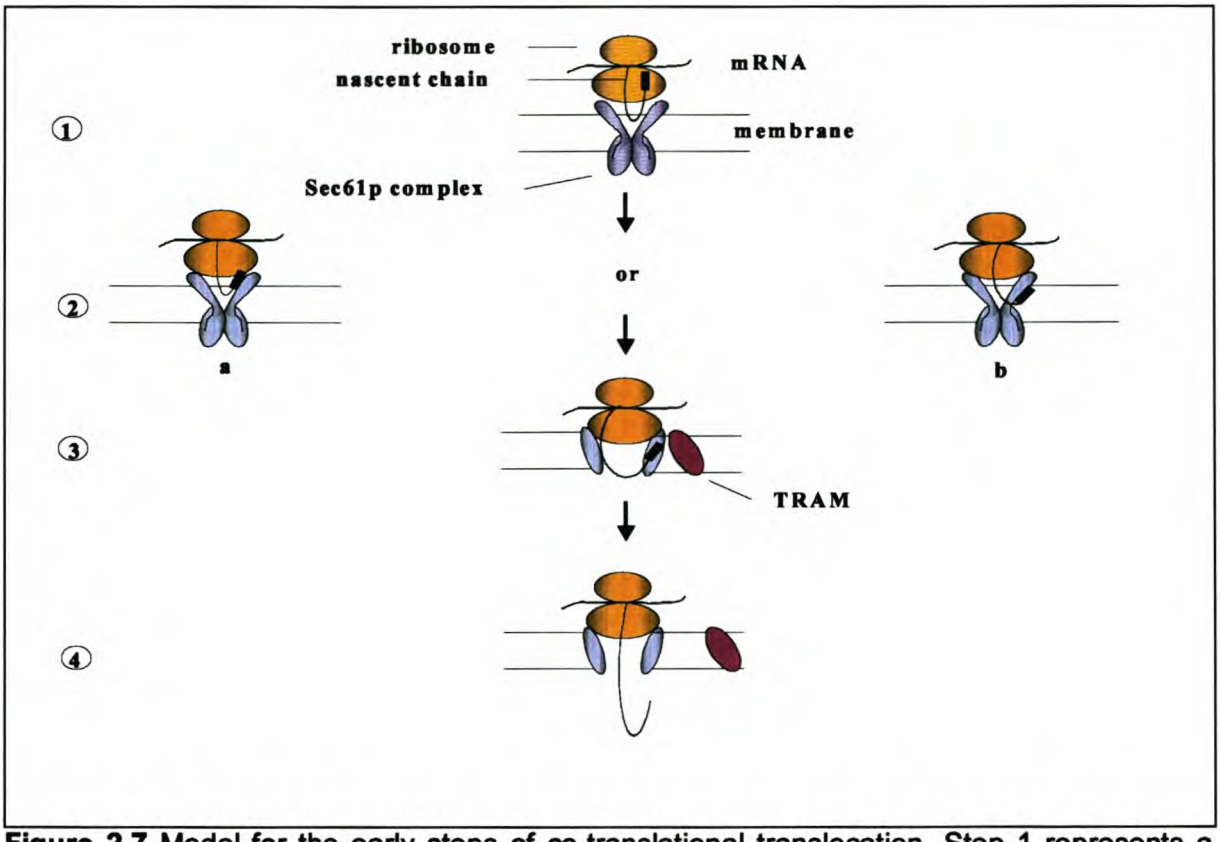


Figure 2.7 Model for the early steps of co-translational translocation. Step 1 represents a hypothetical state shortly after the disengagement of SRP. At step 2, the nascent chain is transferred into the membrane and the ribosome is bound loosely to the Sec61p complex. At this point, the signal peptide may contact either only the Sec61p complex (model a), or both the protein and lipids (model b). Following further chain elongation, the nascent polypeptide adopts a loop structure and its signal peptide is recognised in a process that involves the Sec61p complex and the TRAM protein (in the case of TRAM-dependent proteins). At this stage, step 3, the nascent chain is inserted into the translocon, its signal peptide contacts the TRAM protein, the ribosome is firmly bound to the Sec61p complex, and the protein-conducting channel is open towards the ER lumen. In step 4, the signal peptide has been cleaved off, the TRAM protein is no longer in close proximity to the nascent chain and the nascent chain has adopted a transmembrane orientation. The elongating polypeptide chain is transferred across the protein-conducting channel extending from the ribosome through the tightly linked Sec61p complex (adapted from Rapoport et al. 1996).

It is still unknown whether the SP enters the translocon laterally through the lipid phase, or directly from the cytoplasm through a proteinaceous environment (step 2a or 2b in **Figure 2.7**), although recent experiments by Mothes et al. (1998) point to the latter option as a possible mode of action.

After insertion into the translocon, the hydrophobic core of the SP is in contact with both Sec61p and the TRAM protein (step 3 in **Figure 2.7**). The existence of two

consecutive SP recognition events during co-translational translocation, one in the cytosol by the SRP and the other in the membrane, increases the fidelity of the process (Jungnickel and Rapoport 1995). Another reason for the second SP recognition event in the membrane is that it serves as the point at which co- and posttranslational translocation pathways merge.

Once the initiation phase is completed, SPase cleaves off the SP and the polypeptide adopts a transmembrane orientation. At this point, TRAM is no longer in the immediate proximity of the polypeptide (step 4 in **Figure 2.7**). The elongating chain is transferred directly from a channel in the ribosome to the protein-conducting channel of the membrane. No pushing or pulling of the growing chain is necessary; the polypeptide has only one way out of the extended channel (Lyman and Schekman 1996; Rapoport et al. 1996). Another benefit of the tight coupling between the ribosome channel and the protein-conducting channel is that it prevents both the premature folding of the polypeptide into a translocation-incompetent conformation and the passage of other molecules through the membrane.

The final step in the process of co-translational transport is coupled to the termination of translation. Upon arrival at the mRNA stop codon, the C-terminal residues that are still in the ribosome are transferred across the membrane, the channel is closed and the ribosome is detached from the membrane (Mothes et al. 1998; Young et al. 2001).

2.4.2 POSTTRANSLATIONAL TRANSLOCATION

The mechanism of posttranslational protein transport differs in fundamental aspects from that of co-translational transport. The ribosome obviously plays no role in posttranslational translocation, and vectorial synthesis of the protein chain can no longer be the mechanism by which the directionality of the transport across the membrane is ensured (Cao and Saier 2003; Willer et al. 2003).

Posttranslational translocation in *S. cerevisiae* can be reproduced in reconstituted proteoliposomes and requires the association of the trimeric Sec61p (Sec61p, Sbh1p and Sss1p) with an additional membrane protein complex, the tetrameric Sec62/Sec63p complex, to form the seven-component Sec complex (Matlack et al. 1999). These studies indicate that the Sec61p complex is the major factor responsible for the formation of the protein-conducting channel, since, in the presence of the Sec62/63p complex, it adopts an oligomeric ring structure in the plane of the bilayer (Plath et al. 1998; Wilkinson et al. 2000). Using ppαF as a translocation substrate, about 20% of the molecules are translocated into the proteoliposomes, as demonstrated by their resistance to externally added protease. Neither of the two sub-complexes shows significant translocation activity when dissociated, but the activity is restored to its original level when recombined (Rapoport et al. 1996; Matlack et al. 1999). It was also observed that the addition of ATP and Kar2p to the proteoliposomes increased the translocation rate by a factor of five.

These and other results form the basis of a model for the mechanism of posttranslational translocation (**Figure 2.8**). According to this model, the pathway consists of two stages: the first involves a targeting reaction in which SPs bind to a membrane receptor; the second involves the ATP-dependent translocation of the polypeptide across the ER membrane (Willer et al. 2003).

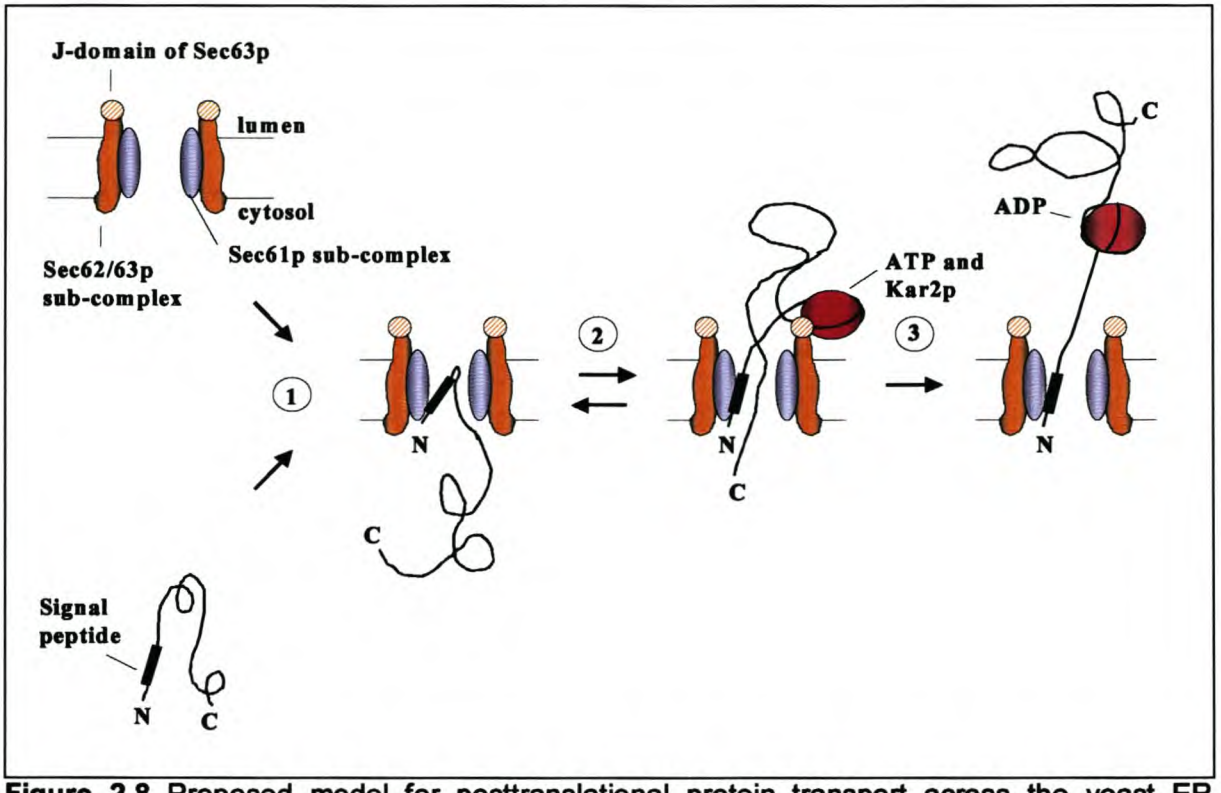


Figure 2.8 Proposed model for posttranslational protein transport across the yeast ER membrane. In step 1, the completed protein chain is inserted in a loop structure into the translocon formed from the Sec complex. In step 2, the C-terminal portion of the hairpin formed by the polypeptide chain can slide back and forth through the protein-conducting channel, but once on the ER lumenal side, Kar2p can bind to it. This process involves an ATP-dependent interaction of Kar2p with the J-domain of the Sec63p subunit of the Sec complex. In step 3, after ATP-hydrolysis, Kar2p binds to the polypeptide chain and prevents its retrograde movement through the protein-conducting channel (adapted from Rapoport et al. 1996; Matlack et al. 1999).

Reconstitution studies with the ppαF as substrate suggest that recognition of the SP and insertion into the protein-conducting channel are in fact the same process (Plath et al. 1998). After binding to the Sec complex, the SP interacts primarily with the Sec61p trans-membrane (TM) domains. The core region of the SP, position 9 to 17 for ppαF, adopts a helical structure that is in contact on different sides with the TM domains 2 and 7 of Sec61p (**Figure 2.9**). The helix contains the hydrophobic part of the SP and forms two to three turns. Other signal sequences could interact with the TM domains in a similar way, but because of the wide variation in composition and length of the SPs, the specific orientation would differ among the SPs (Plath et al. 1998). The bound SP is in a perpendicular orientation to the plane of the bilayer due to its contact with the two TM domains over a considerable distance in the ER.

Insertion of the ppαF into the protein-conducting channel can be explained by means of a hairpin model (see **Figure 2.8**). The N-terminal of the polypeptide faces the cyto-

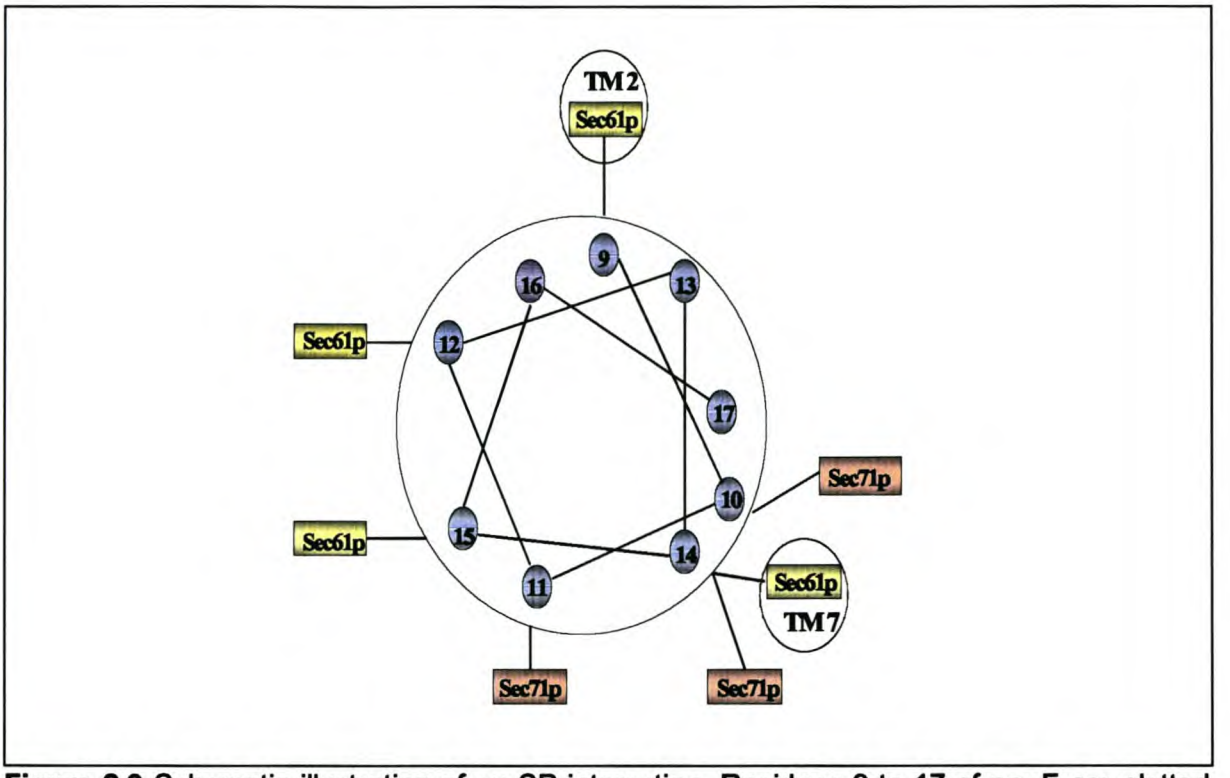


Figure 2.9 Schematic illustration of an SP interaction. Residues 9 to 17 of ppαF are plotted as an α-helix (top view shown), although different helical structures are possible (adapted from Plath et al. 1998).

sol; the C-terminal is in close contact with the lumenal end of the channel, and the next segment of the polypeptide would be dragged through the channel (Rapoport et al. 1996; Plath et al. 1998; Matlack et al. 1999).

Kar2p provides the driving force for translocation by functioning as a molecular ratchet. The ratchet mechanism is activated by the interaction of Kar2p, in its ATP-bound form, with the J-domain of the Sec63p subunit. The J-domain induces hydrolysis of the ATP, converting the open peptide-binding pocket of Kar2p-ATP into the closed peptide-binding pocket of Kar2p-ADP, which results in the subsequent binding of the ADP form to the polypeptide (Brodsky et al. 1995; Young et al. 2001). Activated Kar2p binds peptides with a low specificity, allowing it to bind to any segment of the peptide close to the lumenal end of the channel. The attached Kar2p molecules prevent the segment of the polypeptide from re-entering the channel. The polypeptide may slide back and forth, but, once enough has moved into the lumen, the next activated Kar2p would bind by the same mechanism until the entire polypeptide chain is eventually translocated. The presence of multiple Kar2p molecules on a peptide increases the efficiency of the ratchet mechanism. Kar2p dissociates from the peptide through nucleotide exchange and the binding of ATP reopens the peptide-binding pocket and releases the peptide (Holkeri et al. 1998; Matlack et al. 1999).

This model is quite similar to the one suggested for the import of proteins into the mitochondrial matrix. In the latter case, mt-Hsp70, the homolog of Kar2p, cooperates with the inner membrane protein MIM44/Isp45 in an ATP-dependent manner (Rapoport et al. 1996).

Theoretically, any binding partner would drive translocation, but, as previously indicated, Kar2p has several properties that make it more suitable than other potential ratcheting molecules: (i) it binds immediately at the lumenal end of the channel, where it most effectively can prevent retrograde movement of the polypeptide; (ii) it has a low binding specificity, which allows it to transport a wide variety of peptides; and (iii) its interaction with the peptide is transient, as Kar2p is bound long enough to allow complete translocation, but dissociates quickly enough so that it does not to interfere with folding or modification reactions (Matlack et al. 1999). Kar2p is probably also the only molecule functioning as a ratchet during translocation; it coats the polypeptide and prevents other molecules from participating (Rothman 1989).

2.5 THE UNFOLDED PROTEIN RESPONSE

Considerable evidence points to the translocon having an additional function, namely as a "retrotranslocon" (Sommer and Wolf 1997; Nishikawa et al. 2001). Molecular chaperones bind to improperly folded proteins and target them for retrotranslocation from the ER lumen back to the cytoplasm, where degradation occurs in proteasomes (Travers et al. 2000). The process of retrotranslocation is likely to prove to be universal; its equivalent is known to occur even in bacteria (Cao and Saier 2003).

The *u*nfolded *p*rotein *r*esponse (UPR) pathway was initially discovered in *S. cerevisiae*, in which genetic screens revealed that only three proteins were necessary for signal transduction from the ER to the nucleus. These proteins are Ire1p, a transmembrane serine/threonine kinase, which is believed to act as a sensor in the ER lumen; Hac1p, which directly activates transcription of UPR target genes; and Rlg1p (tRNA ligase), which plays an important role in bridging activation of Ire1p and the production of Hac1p (Mori et al. 2000; Ma and Hendershot 2001; Patil and Walter 2001). The pathway is summarised in **Figure 2.10**.

The most upstream component of the pathway, Ire1p, has three functional domains. The most amino-terminal of these domains is in the ER lumen, where it senses abnormally high levels of unfolded ER proteins. The accumulation of unfolded proteins in the ER lumen causes Ire1p to oligomerise and *trans*-autophosphorylate via its cytosolic kinase domain. The activated kinase then stimulates the activity of the most carboxy-terminal domain of Ire1p. The only known substrate of Ire1p endonuclease is the *HAC1* mRNA, which encodes the basic leucine zipper transcription factor that in turn activates the transcription of the UPR target genes. *HAC1*^{u} ('uninduced') mRNA is constitutively expressed as a 1.4 kb precursor mRNA, but Hac1p is not detectable under normal conditions (Bowring and Llewellyn 2001;

Patil and Walter 2001). This is due to the presence of a 252 bp, non-classical intron near the 3' end of the open reading frame, which prevents *HAC1*^{u} mRNA translation. Upon activation of the UPR, the intron is removed by two site-specific cleavages that are mediated by Ire1p interaction. The 5' and 3' portions of the mRNA are rejoined by tRNA ligase to form a new *HAC1*^{i} ('induced') mRNA molecule, which is efficiently translated to produce the transcription activator Hac1p (Mori et al. 2000; Patil and Walter 2001). Hac1p is translocated to the nucleus, where it activates target gene transcription by binding an UPR-specific upstream activating sequence, the *u*nfolded *p*rotein *r*esponse *e*lement (UPRE).

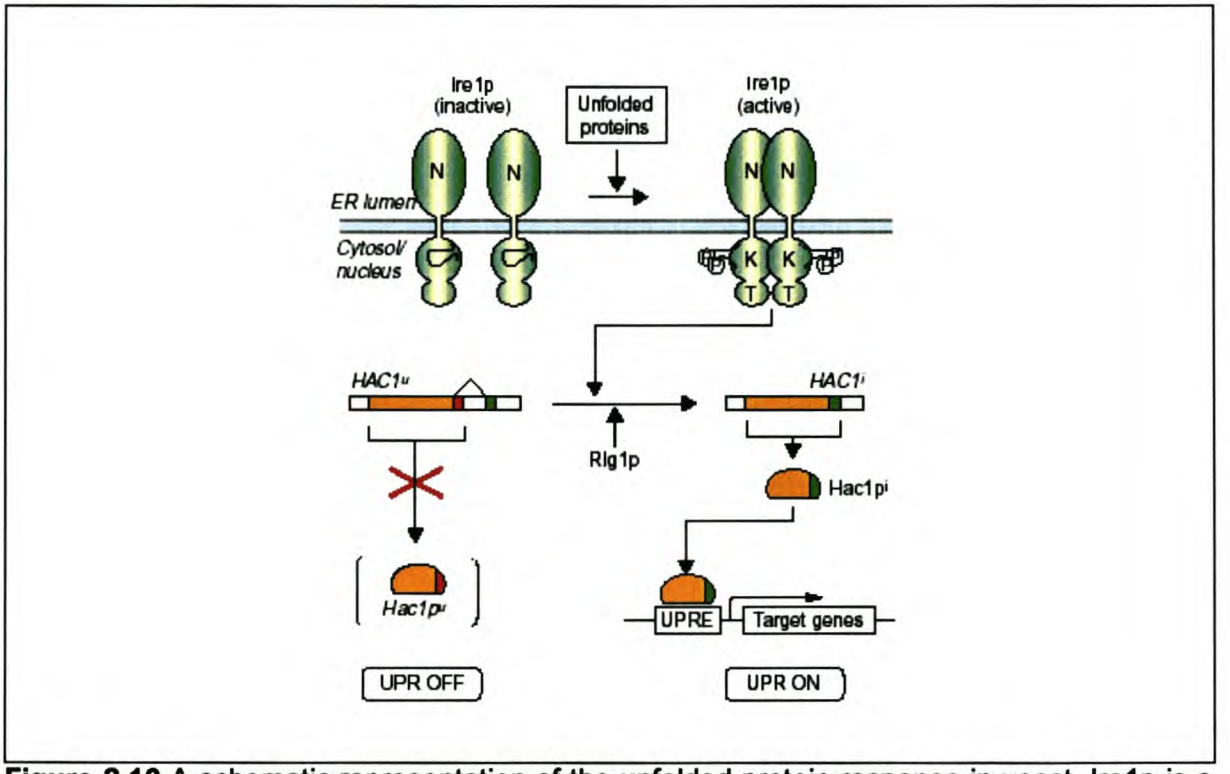


Figure 2.10 A schematic representation of the unfolded protein response in yeast. Ire1p is a transmembrane serine-threonine kinase, oriented with the amino terminal (N) in the ER lumen and the carboxyl terminal in the cytosol. When unfolded proteins accumulate in the ER, Ire1p oligomerises, *trans*-autophosphorylates via the cytosolic kinase domain (K) and activates the endonuclease in the tail domain (T). The endonuclease Ire1p cuts the *HAC1* mRNA at two sites, removing a nonclassical intron, followed by Rlg1p (tRNA ligase) rejoining the two exons. *HAC1*^{u} (uninduced) is not translated due to the presence of the intron, and Hac1p^{u} is not produced (brackets). After Ire1p-mediated splicing, *HAC1*^{i} mRNA is efficiently translated into Hac1p^{i}, a transcriptional activator that upregulates the expression of the UPR target genes after binding to the unfolded protein response element (UPRE) in the promoters of genes encoding ER-resident chaperones and other proteins (Patil and Walter 2001).

The UPRE is found in the promoters of several ER-resident chaperone genes, including in those of *KAR2* and protein disulphide isomerase (*PDI*), and is sufficient enough to upregulate their transcription (Ma and Hendershot 2001).

By upregulating the expression of chaperones and binding proteins, the yeast ER is protected from the accumulation and aggregation of unfolded proteins. Yeast cells are able to activate the entire UPR pathway with three proteins, with the influence of this pathway being more extensive than previously believed. DNA microarrays have

shown that more than 5% of the yeast genome, i.e. more than 350 genes out of 6300, is regulated by the UPR pathway (Patil and Walter 2001). More than half of these regulated genes function in the secretory pathway and include factors involved in protein translocation, glycosylation, ER-associated protein degradation and protein targeting to the cell surface (Ma and Hendershot 2001; Patil and Walter 2001).

2.6 THE YEAST SECRETORY PATHWAY: A BIOTECHNOLOGICAL VIEW

The yeast *S. cerevisiae* has been described as mankind's most domesticated organism. This is largely due to the considerable role it has played in food and beverage production in the past two thousand years. In recent years, however, the importance of *S. cerevisiae* has extended beyond its fermentation uses; it has become a powerful model system for biological research (Gellissen and Hollenberg 1997; Cereghino and Cregg 1999). The *S. cerevisiae* genome has been sequenced and numerous genes have been characterised. The unicellular eukaryote shares many genes with higher eukaryotes. A surprising aspect has been the similarity in structure and function between the *S. cerevisiae* secretory pathway and the mammalian one. The former exhibits the same ability for protein folding, glycosylation, processing and secretion as the human secretory pathway.

In some ways, *S. cerevisiae* is regarded as an optimal host for the production of foreign proteins (see **Table 2.1**; Bitter et al. 1984; Beccera et al. 2001). It does not

Table 2.1 Selected proteins of therapeutic importance that have been produced by recombinant means in *S. cerevisiae.*

Protein	Application
Hepatitis B surface antigen	Vaccination against the disease
Interferon-α	Treatment of cancer and viral diseases
Insulin	Treatment of diabetes mellitus
Human growth hormone	Treatment of short stature
Antibodies/antibody fragments	Used for the *in vivo* detection and treatment of tumours
Human blood factor VIII	Treatment of haemophilia and other blood disorders

produce toxic compounds and the development of a variety of stable vectors with efficient promoters and signal peptides has made this yeast an attractive choice for the production of heterologous proteins (Eckart and Bussineau 1996; Sudberry 1996). The production of a considerable amount of the insulin used worldwide underlines the importance of *S. cerevisiae* in biotechnology (Kjeldsen et al. 1999).

Despite these attributes, the *S. cerevisiae* eukaryotic expression system is also subject to a number of drawbacks (Jenkins and Curling 1994). The expression levels of heterologous proteins are often low, usually representing less than 5% of the total cellular protein. Many of the heterologous proteins produced and secreted by

S. cerevisiae are not released into the extracellular medium, but end up in the periplasmic space. This is especially true in the case of heterologous proteins larger than 30 kDa.

In view of the secretion potential of *S. cerevisiae* and the favourable similarities with the secretion machinery of higher eukaryotes, the low yields of heterologous proteins obtained from the *S. cerevisiae* expression system pose an intriguing question (Conesa et al. 2001; Punt et al. 2002).

In the following section the focus will be on how the acquired knowledge of the secretion pathway has been used in attempts to improve protein secretion from the *S. cerevisiae* system.

2.6.1 TARGETING THE KEY EVENT: TRANSLOCATION

Efforts to improve heterologous protein secretion have focused on translocation and, more specifically, on the components of translocation. It has been argued that, in a heterologous expression system, proteins may encounter translocation difficulties that limit their ability to be secreted, especially in the case of proteins being posttranslationally translocated, and that this problem can possibly be overcome by increasing the levels of chaperones.

Bearing this in mind, Shusta et al. (1998) studied the effect of *KAR2* overexpression on the secretion levels of some proteins. They observed an increase in secretion of five single-chain antibodies from *S. cerevisiae*, but, interestingly, the secretion levels of another heterologous protein, the plant thaumatin, remained the same (Shusta et al. 1998; Conesa et al. 2001). Another contradictory result was obtained for reduced Kar2p levels: a reduction in Kar2p diminished the secretion of three heterologous proteins in *S. cerevisiae*, but increased Kar2p levels did not result in the opposite phenotype (Robinson et al. 1996; Conesa et al. 2001).

Similar results were observed for the overproduction of *PDI*, an enzyme that catalyses the formation of disulphide bonds in secretory proteins. Disulphide-bond formation is an important step in protein assembly, as many proteins, for example antibodies, are oligomers of two or more polypeptide chains held together by disulphide bonds.

PDI is expressed at very low levels in *S. cerevisiae*; it represents less than 0.05% of the total cellular protein (Tuite and Freedman 1994). Overexpression of *PDI* in *S. cerevisiae* increased the secretion levels of numerous heterologous proteins, such as human lysosome, the leech protein antistatin, human platelet-derived growth factor and *Schizosaccharomyces pombe* acid phosphatase (Tuite and Freedman 1994; Conesa et al. 2001). However, as can be expected, PDI overexpression did not enhance the secretion of all the heterologous proteins tested; the secretion of human granulocyte-colony-stimulating factor was unaffected.

Results from the chaperone overexpression studies are far from conclusive. We are observing the different consequences of the cellular functions in which chaperones are involved, as well as the different nature of each secreted protein, and

it is difficult to formulate general rules. As has already been mentioned, Kar2p participates in posttranslational translocation and in retrotranslocation; PDI functions both as a catalyst of protein folding and as a chaperone of properly folded proteins (Conesa et al. 2001).

2.6.2 SYNTHETIC SECRETION LEADERS

Another strategy to improve secretion is the use of synthetic secretion leaders as employed by Kjeldsen et al. (1997) in the secretion of insulin (Sagt et al. 2000). To date, this approach has been the most successful modification to increase heterologous protein secretion from *S. cerevisiae* (Clements et al. 1991; Kjeldsen et al. 1998). The synthetic leaders are based on the *S. cerevisiae* α-factor prepro-peptide and were engineered by combination of semi-random mutations and stepwise optimisation. The amino acid sequence of the synthetic leaders differs considerably from that of the α-factor leader and feature two N-linked glycosylation sites compared to the three of the α-factor leader. Proteolytic processing is improved by the introduction of a removable spacer peptide (EEAEAEAEPK) between the Kex2p site and the N-terminal of the heterologous protein (Kjeldsen et al. 1996). The ability of the synthetic leaders to facilitate secretion was analysed and the experimental evidence indicates a significant increase in fermentation yield for most of the synthetic leaders, with an increase of up to 480% relative to the α-factor leader for the protein fusions of some synthetic leaders (Kjeldsen 2000).

The success of the synthetic leader peptides appears to be based on the quality of the interaction with existing chaperones in the ER, rather than the quantity. Several of the synthetic leaders have heptapeptides with Kar2p-binding scores of +13 (Blond-Elguindi et al. 1993), indicating a very high probability of Kar2p-binding. An improved interaction of the synthetic leaders with Kar2p ensures a quicker entrance into the ER and prolongs the retention of the heterologous protein in the ER lumen, thereby providing additional time for correct folding and subsequent secretion (Kjeldsen et al. 1997; Conesa et al. 2001). It has also been proposed that the synthetic leaders might function as intramolecular chaperones by participating in protein folding (Shinde and Inouye 1993). The synthetic secretion leaders have an amino acid sequence with a highly predictable probability for forming an α-helix. An α-helix in the leader may stabilise the structure by folding back on the heterologous protein, as is the case with several proteins including subtilisin and carboxypeptidase (Zhu et al. 1989; Winther and Sorensen 1991; Eder and Fersht 1995).

2.7 CONCLUSIONS AND PERSPECTIVES

Several longstanding questions about protein translocation have been answered in recent years. The existence of both a co- and posttranslational translocation pathway, with their essential components, has been proven. Reconstitution and crosslinking studies have also identified novel protein complexes, e.g. NAC and

TRAM, which play a role in the translocation process. Major issues currently on the agenda are related to the differences obtained in the secretion of different proteins, namely why some heterologous proteins are secreted and others not; whether there are limits to the transit of proteins through the translocation pathways and, if so, what controls them; and whether there is a proofreading mechanism for SPs.

The answers to these questions may well lead the discovery of additional factors involved in the overall process. Despite its general conservation, the translocation process could vary in different cell types. Novel assays and biochemical methods combined with the existing reconstituted systems are required to address these issues. Finally, the study of nonclassical translocation pathways may lead to the discovery of an entirely new mechanism by which proteins, and more specifically heterologous proteins, are translocated across the ER membrane.

2.8 LITERATURE CITED

Allison DS, Young ET (1988) Single-amino-acid substitutions within the signal sequence of yeast prepro-α-factor affect membrane translocation. Mol Cell Biol 8:1915-1922

Allison DS, Young ET (1989) Mutations in the signal sequence of prepro-α-factor inhibit both translocation into the endoplasmic reticulum and processing by signal peptidase in yeast cells. Mol Cell Biol 9:4977-4985

Andrews DW, Johnson AE (1996) The translocon: more than a hole in the ER membrane? TIBS 21:365-369

Arnold CE, Parekh RN, Yang W, Wittrup KD (1998) Leader peptide efficiency correlates with signal recognition particle dependence in *Saccharomyces cerevisiae*. Biotechnol Bioeng 59:286-293

Beccera M, Diaz Padro S, Gonzalez Siso, MI, Cerdan ME (2001) New secretory strategies for *Kluyveromyces lactis* β-galactosidase. Protein Eng 14:379-386

Bevan A, Brenner C, Fuller RS (1998) Quantitative assessment of enzyme specificity *in vivo*: P_2 recognition by Kex2 protease defined in a genetic system. Proc Natl Acad Sci USA 95:10384-10389

Bird P, Gething MJ, Sambrook J (1987) Translocation in yeast and mammalian cells: not all signals are functionally equivalent. J Biol Chem 105:2905-2914

Bitter GA, Chen KK, Banks AR, Lai P (1984) Secretion of foreign proteins from *Saccharomyces cerevisiae* directed by α-factor fusions. Proc Natl Acad Sci USA 81:5330-5334

Blobel G, Dobberstein B (1975) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains in membrane-bound ribosomes of murine myeloma. J Cell Biol 67:835-851

Blond-Elguindi S, Cwirla SE, Dower WJ, Lipshutz RJ, Sprang SR, Sambrook JF, Gething MJ (1993) Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. Cell 75:717-728

Böhni PC, Deshaies RJ, Schekman RW (1988) *SEC11* is required for signal peptide processing and yeast cell growth. J Cell Biol 106:1035-1042

Bowring CE, Llewellyn DH (2001) Differences in HAC1 mRNA processing and translation between yeast and mammalian cells indicate divergence of the eukaryotic ER stress response. Biochem Biophys Res Com 287: 789-800

Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, Urdea MS, Valenzuela P, Barr PJ (1984) α-Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 81:4642-4646

Brodsky JL (1998) Translocation of proteins across the endoplasmic reticulum. Int Rev Cytol 178:277-328

Brodsky JL, Goeckeler J, Schekman R (1995) BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. Proc Natl Acad Sci USA 92:9643-9646

Cao TB, Saier MH (2003) The general protein secretory pathway: phylogenetic analyses leading to evolutionary conclusions. Biochim Biophys Acta 1609:115-125

Caplan S, Green R, Rocco J, Kurjan J (1991) Glycosylation and structure of the yeast *MFα1* α-factor precursor is important for efficient transport through the secretory pathway. J Bacteriol 173:627-635

Cereghino GPL, Cregg JM (1999) Applications of yeast in biotechnology: protein production and genetic analysis. Curr Opin Biotechnol 10:422-427

Clements JM, Catlin GH, Price MJ, Edwards RM (1991) Secretion of human epidermal growth factor from *Saccharomyces cerevisiae* using synthetic leader sequences. Gene 106:267-272

Cleves AE, Bankaitis VA (1992) Secretory pathway function in *Saccharomyces cerevisiae*. Adv Microbial Physiol 33:73-144

Conesa A, Punt PJ, Van Luijk N, Van den Hondel CAMJJ (2001) The secretion pathway in filamentous fungi: a biotechnological view. Fungal Genet Biol 33:155-171

Connoly T, Rapiejko PJ, Gilmore R (1991) Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor. Science 2:1171-1173

De Nobel JG, Barnett JA (1991) Passage of molecules through yeast cell walls: a brief essay-review. Yeast 7:313-323

Deshaies RJ, Schekman R (1987) A yeast mutant defective at an early stage in import of secretory protein precursors into the endoplasmic reticulum. J Cell Biol 105:633-645

Deshaies RJ, Kepes F, Böhni PC (1989) Genetic dissection of the early stages of protein secretion in yeast. Trends Genet 5:87-93

Deshaies RJ, Sanders SL, Feldheim DA, Schekman R (1991) Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. Nature 349:806-808

Eckart MR, Bussineau C (1996) Quality and authenticity of heterologous proteins synthesized in yeast. Curr Opin Biotechnol 7:525-530

Eder J, Fersht AR (1995) Pro-sequence-assisted protein folding. Mol Microbiol 16:609-614

Feldheim D, Rothblatt J, Schekman R (1992) Topology and functional domains of Sec63p, an ER membrane protein required for secretory protein translocation. Mol Cell Biol 12:3288-3296

Gellissen G, Hollenberg CP (1997) Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. Gene 190:87-97

Gierasch LM (1989) Signal sequences. Biochemistry 28:923-930

Gilmore R, Walter P, Blobel G (1982) Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. J Cell Biol 95:470-477

Görlich D, Hartman E. Prehn S, Rapoport TA (1992) A protein of the endoplasmic reticulum involved in polypeptide translocation. Nature 357:47-52

Görlich D, Rapoport TA (1993) Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. Cell 75:615-630

Haguenauer-Tsapis R (1992) Protein-specific features of the general secretion pathway in yeast: the secretion of acid phosphatase. Mol Microbiol 6:573-579

High S, Dobberstein B (1991) The signal sequence interacts with the methionine-rich domain of the 54-kD protein of signal recognition particle. J Cell Biol 113:229-233

Hikita C, Mizushima S (1992) Effects of total hydrophobicity and length of the hydrophobic domain of a signal peptide on the *in vitro* translocation efficiency. J Cell Biol 267:4882-4888

Holkeri H, Paunola E, Jämsä E, Makarow M (1998) Dissection of the translocation and chaperoning functions of yeast BiP/Kar2p *in vivo*. J Cell Sci 111:749-757

Jenkins N, Curling EMA (1994) Glycosylation of recombinant proteins: problems and prospects. Enzyme Microb Technol 16:354-364

Jiang L, Rogers JC (1999) Functional analysis of a Golgi-localized Kex2p-like protease in tobacco suspension cultures. Plant J 18:23-32

Johnson AE (1993) Protein translocation across the ER membrane: a fluorescent light at the end of the tunnel. TIBS 18:456-458

Jungnickel B, Rapoport TA (1995) A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. Cell 82:261-270

Kaiser CA, Preuss D, Grisafi P, Botstein D (1987) Many random sequences functionally replaces the secretion signal sequence of yeast invertase. Science 235:312-317

Kjeldsen T (2000) Yeast secretory expression of insulin precursors. Appl Microbiol Biotechnol 54:277-286

Kjeldsen T, Brandt J, Andersen AS, Egel-Mitani M, Hach M, Petterson AF, Vad K (1996) A removable spacer peptide in an α-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in *Saccharomyces cerevisiae*. Gene 170:107-112

Kjeldsen T, Petterson AF, Hach M, Diers I, Havelund S, Hansen PH, Andersen AS (1997) Synthetic leaders with potential BiP binding mediate high-yield secretion of correctly folded insulin precursors from *Saccharomyces cerevisiae*. Protein Expression Purification 9:331-336

Kjeldsen T, Andersen AS, Hach M, Diers I, Nikolajsen J, Markussen J (1998) α-Factor pro-peptide N-linked oligosaccharides facilitate secretion of the insulin precursor in *Saccharomyces cerevisiae*. Biotechnol Appl Biochem 27:109-115

Kjeldsen T, Pettersson AF, Hach M (1999) The role of leaders in intracellular transport and secretion of the insulin precursor in the yeast *Saccharomyces cerevisiae*. J Biotechnol 75:195-208

Kurjan J, Herskowitz I (1982) Structure of a yeast pheromone gene (MFα): a putative α-factor precursor contains four tandem copies of mature α-factor. Cell 30:933-943

Larriba G (1993) Translocation of proteins across the membrane of the endoplasmic reticulum: a place for *Saccharomyces cerevisiae*. Yeast 9:441-463

Lütcke H (1995) Signal recognition particle (SRP), a ubiquitous initiator of protein translocation. Eur J Biochem 228:531-550

Lyman SK, Schekman R (1996) Polypeptide translocation machinery of the yeast endoplasmic reticulum. Experientia 52:1042-1049

Ma Y, Hendershot LM (2001) The unfolding tale of the unfolded protein response. Cell 107:827-830

Mason N, Ciufo LF, Brown JD (2000) Elongation arrest is a physiologically important function of a signal recognition particle. EMBO J 19:4164-4174

Matlack KES, Misselwitz B, Plath K, Rapoport TA (1999) BiP acts as a molecular ratchet during posttranslational transport of prepro-α factor across the ER membrane. Cell 97:553-564

Monod M, Haguenauer-Tsapis R, Rauseo-Koenig I, Hinnen A (1989) Functional analysis of the signal-sequence processing site of yeast acid phosphatase. Int J Biochem 182:213-221

Mori K, Ogawa N, Kawahara T, Yanagi H, Yura T (2000) mRNA splicing-mediated C-terminal replacement of transcription factor Hac1p is required for efficient activation of the unfolded protein response. Proc Natl Acad Sci USA 97:4660-4665

Mothes W, Jungnickel B, Brunner J, Rapoport TA (1998) Signal sequence recognition in co-translational translocation by protein components of the endoplasmic reticulum membrane. J Cell Biol 142:355-364

Ngsee JK, Hansen W, Walter P, Smith M (1989) Cassette mutagenic analysis of the yeast invertase signal peptide: effects on protein translocation. Mol Cell Biol 9:3400-3410

Nishikawa SI, Fewell SW, Kato Y, Brodsky JL, Endo T (2001) Molecular chaperones in the yeast endoplasmic reticulum maintain the solubility of proteins for retrotranslocation and degradation. J Cell Biol 153:1061-1070

Novick P, Field C, Schekman R (1980) Identification of 23 complementation groups required for posttranslational events in the secretory pathway. Cell 21:205-215

Palade G (1975) Intracellular aspects of the process of protein synthesis. Science 189:347-358

Patil C, Walter P (2001) Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. Curr Opin Cell Biol 13:349-356

Peberdy JF (1994) Protein secretion in filamentous fungi – trying to understand a highly productive black box. Trends Biotechnol 12:50-57

Plath K, Mothes W, Wilkinson BM, Stirling CJ, Rapoport TA (1998) Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. Cell 94:795-807

Punt PJ, Van Biezen N, Conesa A, Albers A, Mangus J, Van den Hondel, C (2002) Filamentous fungi as cell factories for heterologous protein production. Trends Biotechnol 20:200-206

Rapoport TA (1991) Protein transport across the endoplasmic reticulum membrane: facts, models, mysteries. FASEB J 5:2792-2798

Rapoport TA (1992) Transport of proteins across the endoplasmic reticulum membrane. Science 258:931-936

Rapoport TA, Jungnickel B, Kutay U (1996) Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. Annu Rev Biochem 65:271-303

Robinson AS, Bockhaus JA, Voegler AC, Wittrup KD (1996) Reduction of BiP levels decreases heterologous protein secretion in *Saccharomyces cerevisiae.* J Biol Chem 271:10017-10022

Rothe C, Lehle L (1998) Sorting of invertase signal peptide mutants in yeast dependent and independent on the signal-recognition particle. Eur J Biochem 252:16-24

Rothman JE (1989) Polypeptide chain binding proteins: catalysts of protein folding and related processes in cells. Cell 59:591-601

Sagt CMJ, Kleizen B, Verwaal R, De Jong MDM, Müller WH, Smits A, Visser A, Boonstra J, Verkleij AJ, Verrips CT (2000) Introduction of an N-glycosylation site increases secretion of heterologous proteins in yeasts. Appl Environ Microbiol 66:4940-4944

Sanz P, Meyer DI (1988) Signal recognition particle (SRP) stabilizes the translocation-competent conformation of pre-secretory proteins. EMBO J 7:3553-3557

Schatz G, Dobberstein B (1996) Common principles of protein translocation across membranes. Science 271:1519-1526

Schekman R (1985) Protein localization and membrane traffic in yeast. Annu Rev Cell Biol 1:115-143

Shinde U, Inouye M (1993) Intramolecular chaperones and protein folding. TIBS 18:442-446

Shusta EV, Raines RT, Pluckthun A, Wittrup KD (1998) Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments. Nat Biotechnol 16:773-777

Sidhu RS, Bollon AP (1987) Analysis of α-factor secretion signals by fusing with acid phosphatase of yeast. Gene 54:175-184

Simon S, Blobel G (1991) A protein-conducting channel in the endoplasmic reticulum. Cell 65:371-380

Sommer T, Wolf DH (1997) Endoplasmic reticulum degradation: reverse protein flow of no return. FASEB J 11:1127-1233

Sudberry PE (1996) The expression of recombinant proteins in yeasts. Curr Opin Biotechnol 7:517-524

Suzuki Y, Ikeda N, Kataoka E, Ohsuye K (2000) Effect of amino acid substitution at the P_3 and P_4 subsites of fusion proteins on Kex2 protease activity. Biotechnol Appl Biochem 32:53-60

Taussig R, Carlson M (1983) Nucleotide sequence of the yeast *SUC2* gene for invertase. Nucleic Acids Res 11:1943-1954

Travers KJ, Patil CK, Wodicka L, Lockhart DJ, Weissman JS, Walter P (2000) Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell 101:249-258

Tuite M, Freedman RB (1994) Improving secretion of recombinant proteins from yeast and mammalian cells: rational or empirical design? Trends Biotechnol 12:432-434

VanValkenburgh C, Chen X, Mullins C, Fang H, Green N (1999) The catalytic mechanism of endoplasmic reticulum signal peptidase appears to be distinct from most eubacterial signal peptidases. J Biol Chem 274:11519-11525

Voigt S, Jungnickel B, Hartmann E, Rapoport TA (1996) Signal-sequence-dependent function of the TRAM protein during early phases of protein translocation across the ER membrane. J Cell Biol 134:25-35

Von Heijne G (1981) On the hydrophobic nature of signal sequences. Eur J Biochem 116:419-422

Von Heijne G (1985) Signal sequences: the limits of variation. J Mol Biol 184:99-105

Waters MG, Evans EA, Blobel G (1988) Prepro-α-factor has a cleavable signal sequence. J Biol Chem 263:6209-6214

Wiedman B, Sakai H, Davis TA, Wiedman M (1994) A protein complex required for signal-sequence-specific sorting and translocation. Nature 370:434-440

Wilkinson BM, Tyson JR, Reid PJ, Stirling CJ (2000) Distinct domains within yeast Sec61p involved in post-translational translocation and protein dislocation. J Biol Chem 275:521-529

Willer M, Jermy AJ, Young BP, Stirling CJ (2003) Identification of novel protein-protein interactions at the cytosolic surface of the Sec63p complex in the yeast ER membrane. Yeast 20:133-148

Winther JR, Sorensen P (1991) Propeptide of carboxypeptidase Y provides a chaperone-like function as well as inhibition of the enzymatic activity. Proc Natl Acad Sci USA 88:9330-9334

Young BP, Craven RA, Reid PJ, Willer M, Stirling CJ (2001) Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum *in vivo*. EMBO J 20:262-271

Zheng N, Gierasch LM (1996) Signal sequences: the same yet different. Cell 86:849-852

Zhu X, Yoshiji O, Jordan F, Inouye M (1989) Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. Nature 339:483-484

Zsebo KM, Lu H, Fieschko JC, Goldstein L, Davis J, Duk K, Suggs SV, Lai P, Bitter GA (1986) Protein secretion from *Saccharomyces cerevisiae* directed by the prepro-α-factor leader region. J Biol Chem 261:5858-5865

CHAPTER 3

RESEARCH RESULTS

Optimisation of the secretion of two *Lipomyces kononenkoae* α-amylases in *Saccharomyces cerevisiae*

This manuscript will be submitted for publication in

Applied Microbiology and Biotechnology

3. RESEARCH RESULTS

OPTIMISATION OF THE SECRETION OF TWO *LIPOMYCES KONONENKOAE* α-AMYLASES IN *SACCHAROMYCES CEREVISIAE*

J. H. Kriel[1], N. Ramachandran[1], P. van Rensburg[1], R. R. Cordero Otero[1] and I. S. Pretorius[2]

[1]*Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, Victoria Street, Stellenbosch, ZA 7600, South Africa*
[2]*Australian Wine Research Institute, Waite Road, Urrbrae, Glen Osmond, Adelaide, SA 5064, Australia*

ABSTRACT

There is great interest in the development of amylolytic *Saccharomyces cerevisiae* strains for industrial processes in which the production of amylases, hydrolysis of starch-rich substrates and fermentation of resulting sugars to commercially important products (e.g. bio-fuel ethanol, low carbohydrate beer, grain whisky, etc.) occur in a single step. Previously we have reported the cloning and expression of the *Lipomyces kononenkoae LKA1* and *LKA2* α-amylase-encoding genes in *S. cerevisiae*. Secretion of these *L. kononenkoae* α-amylases in *S. cerevisiae* was directed by the native *LKA1* and *LKA2* encoded leader sequences. The purpose of the present study was to devise a means through which the levels of secreted Lka1p and Lka2p in *S. cerevisiae* could be increased. Three different sets of expression cassettes were prepared that contained the *LKA1* and *LKA2* genes separately under the secretory direction of either their native secretion leader peptides, the *S. cerevisiae* mating pheromone α-factor (MFα1) secretion leader or the MFα1 secretion leader containing a synthetic C-terminal spacer peptide (EEGEPK). These constructs were transformed into CENPK 42, a laboratory strain of *S. cerevisiae*. Southern blot analysis confirmed the stable integration of these constructs into the yeast genome and plate assays revealed amylolytic activity, demonstrating a functional expression system. The *S. cerevisiae* strains secreting the *LKA1* and *LKA2* α-amylases under their native secretion leaders resulted in the highest saccharolytic activity in the extracellular medium. In contrast, *S. cerevisiae* strains utilising the synthetic MFα1 secretion leader produced the highest fermentation yield, with 62.91 mg/l of extracellular protein for the transformant secreting both α-amylases under direction of the synthetic secretion peptides, compared to the 19.05 mg/l for the transformant secreting both amylases under their native leader peptides. We hypothesise that the unexpected differences between the enzymes secreted and the activity of the secreted enzymes are due to additional interactions of the leader with its passenger protein.

3.1 INTRODUCTION

The yeast *Saccharomyces cerevisiae* has proved to be a useful host for the overproduction of proteins, both homologous and heterologous. A wide variety of tools and techniques perfected over the past few years have allowed the expression and secretion of a number of pharmacologically important proteins, including insulin and interferon-α (Cereghino and Cregg 1999; Conesa et al. 2001).

Secretion, as opposed to intracellular expression, has become the preferred route of production for most of the proteins. Correct folding and glycosylation of proteins occur more readily in the secretory pathway and the secretion of proteins removes them from the yeast's proteases. Some proteins are also unstable or toxic when expressed intracellularly and these problems can be circumvented by secretion.

As in higher eukaryotes, protein secretion in yeast is directed by an N-terminal signal peptide that initiates translocation into the endoplasmic reticulum (Brake et al. 1984; Zsebo et al. 1986). The *S. cerevisiae* system for the secretion of recombinant proteins utilises the *S. cerevisiae* α-factor leader and the *KEX2*-encoded endoproteinase for protein maturation. The yeast mating pheromone α-factor is synthesised as a 165 amino acid precursor protein that comprises a signal peptide of 19 amino acids, aptly named the pre region, and a 64 amino acid pro region containing three sites for N-linked glycosylation (Caplan et al. 1991). The pro region is followed by four repeats of the mature 13 amino acid α-factor pheromone sequence, with each of the four peptides being preceded by a short spacer sequence. On a molar basis, α-factor is produced as efficiently as the highly expressed yeast glycolytic enzymes (Bitter et al. 1984). It is therefore not surprising that the components responsible for the secretion of α-factor are utilised in a similar manner to direct secretion of heterologous proteins into the extracellular medium.

The role of leader peptides in secretion is still poorly understood. It is known to play a central role in the targeting and secretion of nearly all proteins in both prokaryotes and eukaryotes, yet it displays a lack of primary sequence similarity (Gierasch 1989; Zheng and Gierasch 1996). The ability of leader peptides to facilitate secretion in different organisms points to a general conservation of the secretory process; however, this generalisation is based solely on evidence from a limited set of secretion leaders (Zheng and Gierasch 1996).

The renewed interest in secretion leaders is mainly due to their practical importance in a booming biotechnological industry. There is a steady increase in demand for more effective heterologous protein expression systems (Sudberry 1996). Kjeldsen et al. (1997) developed a novel *S. cerevisiae* expression-secretion system based on the use of synthetic leader peptides for the secretion of proinsulin. The synthetic leader sequences were engineered by a combination of rational design and stepwise optimisation. They feature potential BiP sites and were shown to increase the retention time of proinsulin in the endoplasmic reticulum (ER), thus providing additional time for the correct folding and subsequent secretion of the

heterologous protein (Simons et al. 1995; Kjeldsen et al. 1998; Kjeldsen 2000). It has also been proposed that the synthetic leader might function as an intramolecular chaperone by participating in protein folding. The leader peptides have amino acid sequences with a high probability for forming an α-helix. An α-helix in the leader may stabilise the structure by folding back on the heterologous protein, as is the case in several proteins (Shinde and Inouye 1993).

The present study forms an integral part of a long-term research programme in which we endeavour to help lay the foundation for the development of *S. cerevisiae* strains optimised for a process in which the production of amylolytic enzymes, hydrolysis of starch-rich substrates and the fermentation of resulting sugars to a desired product (e.g. bio-fuel ethanol, low carbohydrate beer, grain whisky, etc.) occurs in a single step. Previously we reported the cloning and the expression of two α-amylase-encoding genes, *LKA1* and *LKA2*, from the most efficient raw-starch-degrading yeast, *Lipomyces kononenkoae* (Spencer-Martins and Van Uden 1979) in *S. cerevisiae* (Eksteen et al. 2003). The main purpose of the present study was to identify means to increase the secretion of two *L. kononenkoae LKA1* and *LKA2* encoded α-amylases from recombinant *S. cerevisiae* strains. Here we report on the use and comparison of various secretion leader peptides and their capacity to direct heterologous secretion of the α-amylase enzymes in *S. cerevisiae*.

3.2 MATERIALS AND METHODS

3.2.1 MICROBIAL STRAINS AND PLASMIDS

The sources and relevant genotypes of bacterial and yeast strains, together with the plasmids used in this study, are listed in Tables 1 and 2.

3.2.2 GROWTH MEDIA AND CULTURE CONDITIONS

Escherichia coli was grown in Luria Bertani (LB) broth (Sambrook et al. 1989). Bacterial transformants were propagated in LB medium containing ampicillin (100 μg/ml). Yeasts were cultured in Yeast Peptone Dextrose (YPD) medium containing 1% yeast extract, 2% bacto-peptone and 2% glucose. The *S. cerevisiae* transformants were cultured and selected on Synthetic Complete (SC) medium containing 2% glucose and 0.67% yeast nitrogen base without amino acids. Growth factors were added according to the selectable marker used. Starch medium containing potato soluble starch (Sigma-Aldrich, Steinheim, Germany) as only carbon source and SC medium containing Phadebas starch (20 Phadebas tablets per litre; Pharmacia Diagnostics, Uppsala, Sweden) was used to screen for amylolytic activity. All solid media contained 2% agar. Selective plates were incubated at 30°C for 2 to 4 days and the appearance of halos around the colonies was taken to be indicative of starch utilisation (Steyn and Pretorius 1995). Bacteria and yeasts were routinely cultured at 37°C and 30°C respectively.

Table 1 Microbial strains used in this study

Strain	Genotype	Source/Reference
***Escherichia coli* strains:**		
DH5α	*F⁻ recA1 endA1 gyrA96 thi hsdR17 supE44 relA1* Δ(*argF-Lac-ZYA*) (φ80d*lacZ*(M15)λ	Invitrogen
***S. cerevisiae* strains:**		
CENPK 42	*Matα, leu2-3, 112 his3-Δ1 ura3-52 trp1-289*	This laboratory
CENPK MFα1S	*Matα ura3*::pMFα1S	This study
CENPK LKA1	*Matα ura3*::pLKA1	This study
CENPK LKA2	*Matα ura3*::pLKA2	This study
CENPK LKA1/2	*Matα his3 ura3*::pLKA1/pLKA2^{his3}	This study
CENPK MFα1-LKA1	*Matα ura3*::pMFα1-LKA1	This study
CENPK MFα1-LKA2	*Matα ura3*::pMFα1-LKA2	This study
CENPK MFα1-LKA1/2	*Matα his3 ura3*::pMFα1-LKA1/ pMFα1-LKA2^{his3}	This study
CENPK IPLKA1	*Matα* ura3::pIPLKA1	This study
CENPK IPLKA2	*Matα ura3*::pIPLKA2	This study
CENPK IPLKA1/2	*Matα ura3*::pIPLKA1/2	This study
CENPK MFα1S-LKA1	*Matα ura3*::pMFα1S-LKA1	This study
CENPK MFα1S-LKA2	*Matα ura3*::pMFα1S-LKA2	This study
CENPK MFα1S-LKA1/2	*Matα his3 ura3*::pMFα1S-LKA1/ pMFα1S-LKA2^{his3}	This study

3.2.3 DNA MANIPULATIONS, SEQUENCING AND AMPLIFICATION

Techniques employed in the sub-cloning and manipulation of DNA fragments, plasmid DNA isolation, the transformation of *E. coli*, agarose gel electrophoresis and the purification of DNA fragments were carried out as described by Sambrook et al. (1989). Restriction endonucleases and T4 DNA ligase were purchased from

Roche Biochemical Products (Randburg, South Africa) and used as recommended by the supplier.

Bi-directional DNA sequencing was carried out using the ABI PRISM® Big Dye™ Terminator cycle sequencing Ready Reaction kit with an ABI PRISM™ 377 DNA sequencer (PE/Applied Biosystems) (Sanger et al. 1977). Sequencing was performed directly on plasmid DNA using the synthetic oligonucleotides as primers.

Table 2 Plasmids used in this study

Plasmid	Properties[1]	Source/Reference
YEp352	Ap^R Tc^R *$PGK1_{PT}$ URA3 2μm*	This laboratory
$pSTA^{ura3/his3}$	Ap^R Tc^R *$PGK1_{PT}$ HIS3 URA3*	This study
pGEM	Ap^R *f1 ori LacZ $T7_P$ $SP6_P$*	Yanish-Peron et al. 1985
pMFα1S	Ap^R Tc^R *$PGK1_{PT}$ $MF\alpha1_{SS}$ URA3*	This study
pLKA1	Ap^R Tc^R *$PGK1_{PT}$ LKA1 URA3*	This study
pLKA2	Ap^R Tc^R *$PGK1_{PT}$ LKA2 URA3*	This study
$pLKA2^{his3}$	Ap^R Tc^R *$PGK1_{PT}$ LKA2 HIS3*	This study
pMFα1-LKA1	Ap^R Tc^R *$PGK1_{PT}$ $MF\alpha1_S$ LKA1 URA3*	This study
pMFα1-LKA2	Ap^R Tc^R *$PGK1_{PT}$ $MF\alpha1_S$ LKA2 URA3*	This study
$pMF\alpha1\text{-}LKA2^{his3}$	Ap^R Tc^R *$PGK1_{PT}$ $MF\alpha1_S$ LKA2 HIS3*	This study
pIPLKA1	Ap^R Tc^R *$PGK1_{PT}$ $LKA1_S$ LKA1 URA3*	Eksteen et al. 2003
pIPLKA2	Ap^R Tc^R *$PGK1_{PT}$ $LKA2_S$ LKA2 URA3*	Eksteen et al. 2003
pIPLKA1/2	Ap^R Tc^R *$PGK1_{PT}$ $LKA1_S$ LKA1 $PGK1_{PT}$ $LKA2_S$ LKA2 URA3*	Eksteen et al. 2003
pMFα1S-LKA1	Ap^R Tc^R *$PGK1_{PT}$ $MF\alpha1_{SS}$ LKA1 URA3*	This study
pMFα1S-LKA2	Ap^R Tc^R *$PGK1_{PT}$ $MF\alpha1_{SS}$ LKA2 URA3*	This study
$pMF\alpha1S\text{-}LKA2^{his3}$	Ap^R Tc^R *$PGK1_{PT}$ $MF\alpha1_{SS}$ LKA2 HIS3*	This study

[1]*S*, secretion signal; *SS*, secretion signal with spacer peptide

DNA was amplified by the polymerase chain reaction (PCR) technique. The PCR reaction was conducted in a 50 μl volume containing 25 mM of $MgCl_2$, 200 μM of a nucleotide mixture (dNTPs), 200 nM of each primer, 2.6 U of Expand High Fidelity polymerase (Roche Biochemical Products) and 2-10 ng of the appropriate template

DNA. After an initial denaturation step at 94°C for 2 min, the PCR reaction was conducted for 30 cycles with denaturation at 94°C for 30 s, annealing at 55-58°C for 45 s, and extension at 72°C for 1-2 min. There was a final incubation at 72°C for 5 min to fill in ends. The PCR products were purified by ethanol precipitation. Restriction enzyme digestion of the PCR products was performed at the appropriate temperature(s) for 12 h. Primers were purchased from Roche Biochemical Products. The following synthetic oligodeoxyribonucleotides were used as primers:

F-LKA1$_{EcoRI\ and\ HindIII}$ (5'-CGGAATTCCGGAAGCTTGATTGCACTACAGTTACG-3')
R-LKA1$_{XhoI}$ (5'-GGCTCGAGCTACATGGAGCAGATTC-3')
F-MFalpha$_{EcoRI}$ (5'-CGGAATTCATGAGATTTCCTTCAATTTT-3')
R-MFalphaS$_{HindIII}$ (5'-GGAAGCTTTGGTTCGCCTTCCTCTCTTTTATCCAAAGATACC-3')
F-LKA2$_{HindIII}$ (5'-GAAGCTTAAGACCGCGGCAGAATG-3')
R-LKA2$_{XhoI}$ (5'-GCTCGAGTTAAGAACAAAATTTCCCAG-3')

3.2.4 YEAST TRANSFORMATION

S. cerevisiae strain CENPK 42 was transformed using the lithium acetate protocol, as described by Gietz and Schiestl (1991). The different constructs were first linearised by digesting them in the *URA3* and/or *HIS3* gene(s) of the vector, pSTA, with either *Apa*I, *Nhe*I, *Nco*I, *Kpn*I or *Stu*I, depending on the gene involved.

3.2.5 SOUTHERN BLOT ANALYSIS

Southern hybridisation was performed using the DIG Labelling Kit from Roche Biochemical Products, using the method described in the Dig Application Manual.

3.2.6 ISOLATION OF EXTRACELLULAR PROTEIN

Liquid cultures of transformed strains were performed in Erlenmeyer flasks filled with 100 ml of YPD broth. As inocula, a suitable volume of stationary phase pre-culture in YPD medium was added to obtain an initial absorbance of 0.2 at 600 nm. Cells were cultivated at 30°C for 72 h. After centrifugation at 5000 rpm for 5 min, the supernatant containing the extracellular protein was removed and used in the enzyme assays.

3.2.7 ENZYME ASSAYS

The saccharolytic activity of the two amylases was assayed by the dinitrosalicylic acid (DNS) method (Miller et al. 1960). The standard reaction mixture, containing 700 µl of substrate, 100 µl of enzyme solution and 200 µl of buffer, was incubated at 30°C for 60 min. The reaction was stopped with 1.5 ml of DNS reagent and boiled for 15 min. The contents of the assay tubes were diluted with 5 ml of distilled water and absorbance readings (A_{540}) were taken in a Beckman DU-5 spectrophotometer. Potato soluble starch (0.5%) was heated on a magnetic stirrer plate prior to its use as substrate in the enzyme assays. The assay was performed at room temperature and pH 6.9.

3.3 RESULTS AND DISCUSSION

3.3.1 CONSTRUCTION OF EXPRESSION CASSETTES

The plasmids containing the constructed α-amylase secretion cassettes are shown in Figure 1.

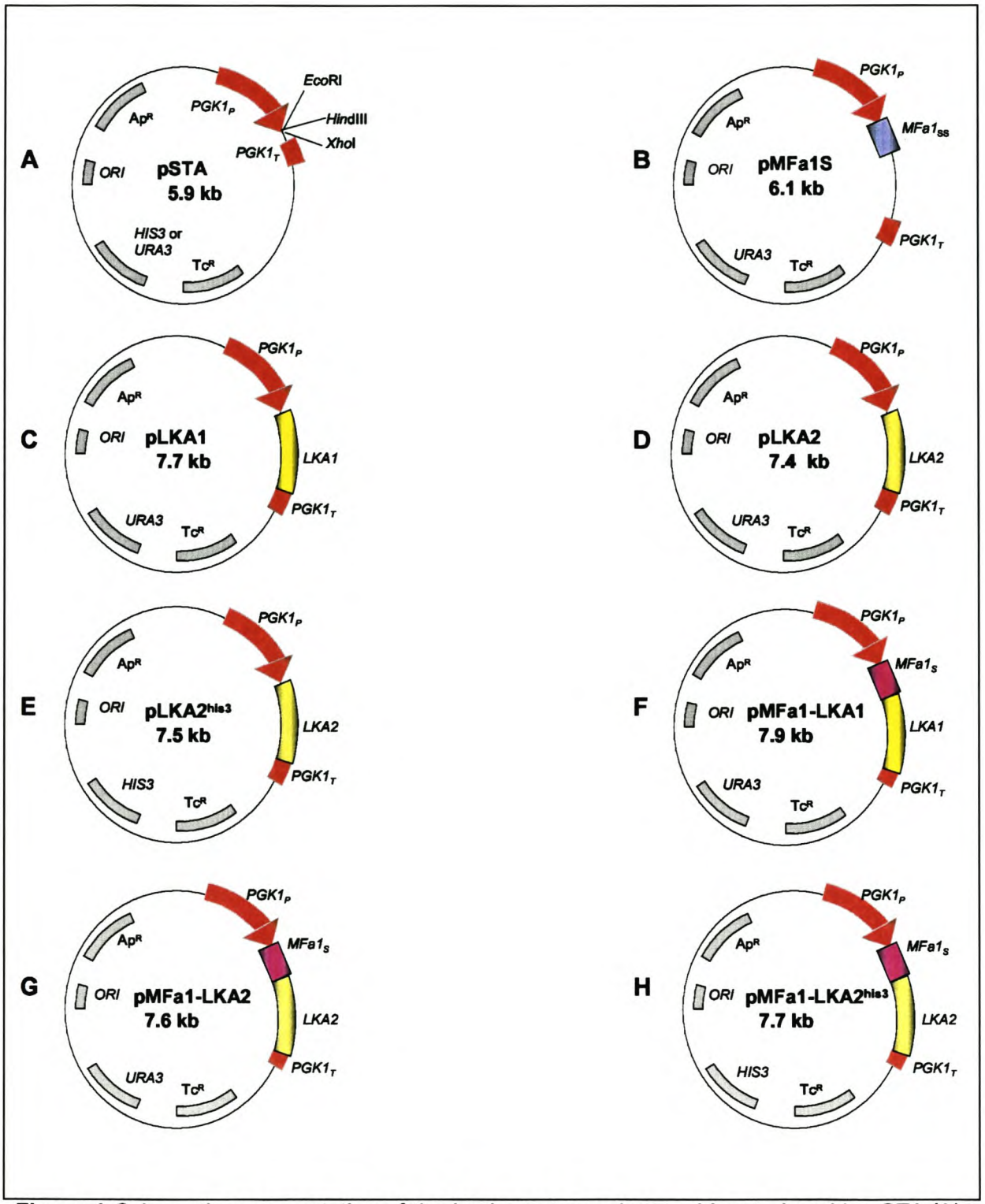


Figure 1 Schematic representation of the basic vector and recombinant plasmids pSTA (A), pMFα1S (B), pLKA1 (C), pLKA2 (D), pLKA2^{his3} (E), pMFα1-LKA1 (F), pMFα1-LKA2 (G) and pMFα1-LKA2^{his3} (H)

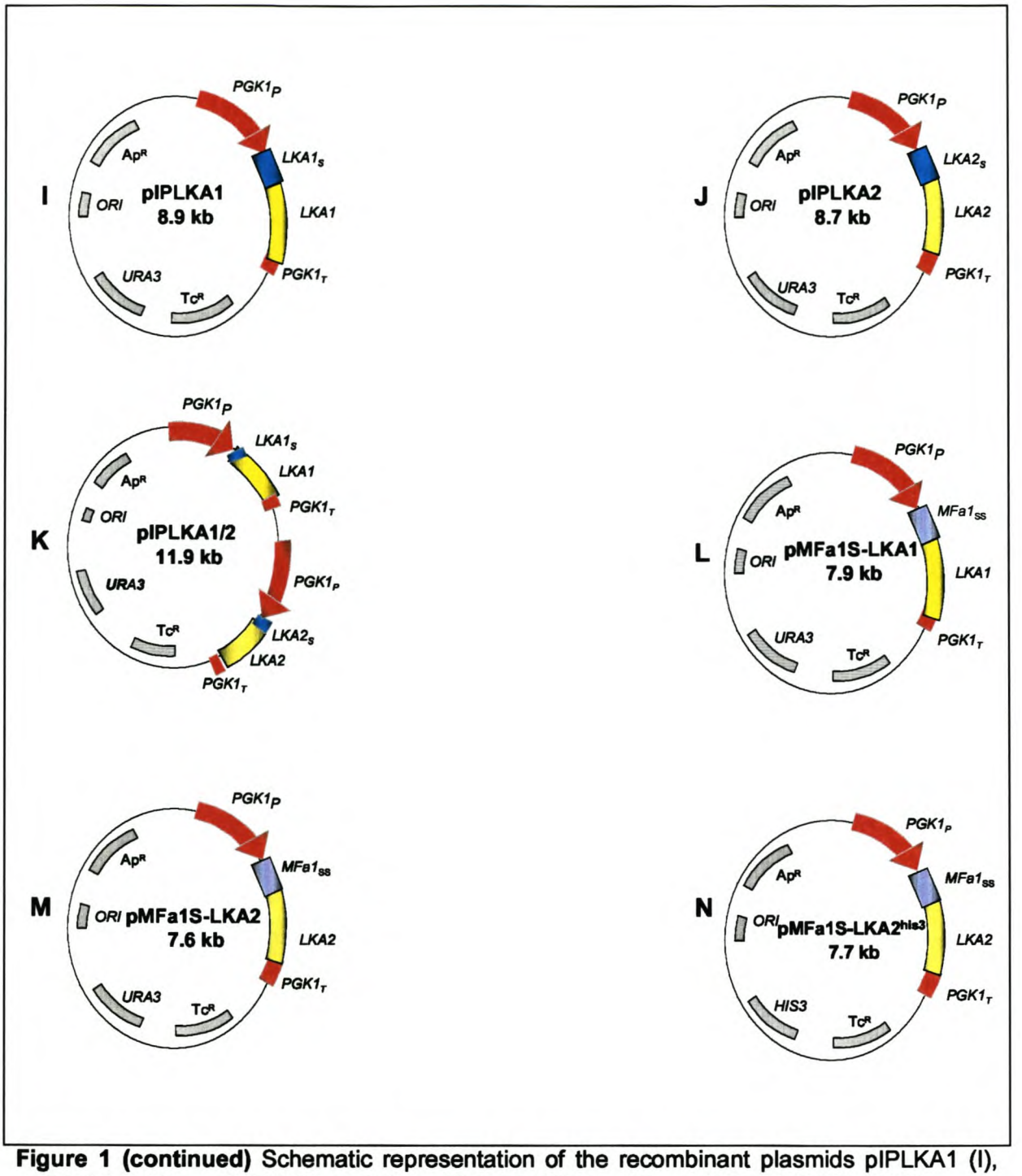


Figure 1 (continued) Schematic representation of the recombinant plasmids pIPLKA1 (I), pIPLKA2 (J), pIPLKA1/2 (K), pMFα1S-LKA1 (L), pMFα1S-LKA2 (M) and pMFα1S-LKA2^{his3} (N)

The pSTA^{ura3} and pSTA^{his3} plasmids (Table 2) allow the expression of the cloned gene under the transcriptional control of the strong constitutive *PGK1* promoter and terminator sequences in *S. cerevisiae*. In addition, these plasmids also have multiple unique restriction sites, namely *Eco*RI, *Hin*dIII and *Xho*I, which are contained between the promoter and the terminator sequences and facilitate the cloning of a secretion block, i.e. a secretion signal and a structural gene.

The first set of constructs, pLKA1, pLKA2 and pLKA2^{his3}, consists of the *L. kononenkoae* α-amylase genes, *LKA1* and *LKA2*, which were expressed without any secretion leader, native or MFα1.

Plasmids pMFα1-LKA1, pMFα1-LKA2 and pMFα1-LKA2^{his3} allow the expression of the cloned genes in-frame to the *S. cerevisiae* α-factor secretion leader.

The third set of constructs, pIPLKA1, pIPLKA2 and pIPLKA1/2, which were previously constructed by Eksteen et al. (2003), allows the expression of the α-amylase genes under their individual native leaders.

Plasmids pMFα1S-LKA1, pMFα1S-LKA2 and pMFα1S-LKA2^{his3} are identical to the second set of constructs, except for the introduction of a spacer peptide (EEGEPK) after the dibasic Kex2p processing site of each secretion leader. The inclusion of the spacer peptides ensures enhanced Kex2p proteolytic processing of the α-factor leader/amylase fusion protein (Kjeldsen et al. 1996).

Plasmid pMFα1S, constructed only with the α-factor secretion leader and spacer peptide, served as control during the enzyme assays.

3.3.2 EXPRESSION OF THE DIFFERENT CONSTRUCTS IN *S. CEREVISIAE*

The 13 constructs were transformed separately into the haploid strain *S. cerevisiae* CENPK 42, resulting in CENPK MFα1Spacer, CENPK LKA1, CENPK LKA2, CENPK LKA1/2, CENPK MFα1-LKA1, CENPK MFα1-LKA2, CENPK MFα1-LKA1/2, CENPK IPLKA1, CENPK IPLKA2, CENPK IPLKA1/2, CENPK MFα1Spacer-LKA1, CENPK MFα1Spacer-LKA2 and CENPK MFα1Spacer-LKA1/2 (Table 1). Southern hybridisations were performed to confirm stable integration into the yeast genome. Once integration had been confirmed, plate assays were used to determine whether the transformants were capable of secreting functional amylolytic enzymes (data not shown).

3.3.3 SACCHAROLYTIC ACTIVITY OF THE DIFFERENT TRANSFORMANTS

The saccharolytic activity of the two amylases, *LKA1* and *LKA2*, and the total extracellular protein present in the culture supernatants were determined for each of the transformants (see Materials and Methods). Table 3 summarises these results. As was expected, no amylase activity was detected in the transformants expressed without a structural gene (CENPK MFα1S as control strain) or secretion leader (CENPK LKA1, CENPK LKA2 and CENPK LKA1/2). The transformant CENPK IPLKA1 resulted in the highest saccharolytic activity, namely 137.36 U/l. However, considerably lower values were obtained for the strains expressing the α-amylase genes under direction of the MFα1 and MFα1Spacer secretion leaders. CENPK MFα1-LKA1 and CENPK MFα1S-LKA1 resulted in values of 90.96 and 90.31 U/l respectively. Transformants expressing the *LKA2* α-amylase, separately and jointly with the *LKA1* α-amylase, follow a similar pattern: CENPK IPLKA2, with 93.12 U/l saccharolytic activity in the culture medium, compared to the 73.41 and 61.59 U/l activity for CENPK MFα1-LKA2 and CENPK MFα1S-LKA2 respectively.

Table 3 Fermentation yield and saccharolytic activity of two different α-amylases expressed in *S. cerevisiae* CENPK 42 by different secretion leaders. One unit (U) of enzyme activity was defined as the amount of enzyme required to liberate 1 mmol of reducing sugar/min. The values are means obtained from four independent experiments; deviations are within 10%

Strain	Pro-peptide	Total protein[1]	Yield[1]	Saccharolytic activity[2]
CENPK MFα1S[a]	*MFα1S*[a]	147.44	0.00	N. D.[3]
CENPK LKA1	-	150.03	2.59	N. D.
CENPK LKA2	-	148.25	0.81	N. D.
CENPK LKA1/2	-	149.05	1.61	N. D.
CENPK MFα1-LKA1	*MFα1*	170.81	23.37	90.96
CENPK MFα1-LKA2	*MFα1*	164.65	17.21	73.41
CENPK MFα1-LKA1/2	*MFα1*	166.28	18.83	83.43
CENPK IPLKA1	*LKA1*	176.28	28.84	137.36
CENPK IPLKA2	*LKA2*	172.09	24.65	93.12
CENPK IPLKA1/2	*LKA1 & 2*	166.49	19.05	100.40
CENPK MFα1S-LKA1	*MFα1S*	177.09	29.65	90.31
CENPK MFα1S-LKA2	*MFα1S*	183.84	36.40	61.59
CENPK MFα1S-LKA1/2	*MFα1S*	210.35	62.91	81.44

[a] S denotes spacer peptide

[1] mg/l

[2] U/l

[3] Not detected

The values obtained for the total extracellular protein produced by these transformants are in sharp contrast to the above-mentioned observations. Transformants expressing the α-amylase genes under secretory direction of the synthetic MFα1Spacer leader resulted in a noticeably higher fermentation yield, especially for those transformants expressing both amylase genes under the synthetic secretion leader.

There are several possibilities for the unexpected differences between the enzymes secreted and the activity of the secreted enzymes, one being that of secretion efficiency versus processing efficiency.

Kjeldsen et al. (1996) indicated the importance of an optimum Kex2p cleavage site. They improved secretion efficiency of the insulin precursor by introducing a

spacer peptide containing the Kex2p target sequences. We employed a similar strategy and the fermentation yield obtained with these constructs is consistent with what was expected.

The *LKA1* and *LKA2* encoded α-amylases, however, are enzymes with complex tertiary structures. We hypothesise that optimal processing of these enzymes requires the (exclusive) participation of their native secretion leader as an intramolecular chaperone, as is the case with several other proteins such as the carboxypeptidase Y of *S. cerevisiae* and subtilisin E of *Bacillus subtilis* (Zhu et al. 1989; Winther and Sorensen 1991). Enzyme activity is based, among other factors, on three-dimensional structure. It is therefore tempting to think that the native amylase leaders might function as chaperones and assist in the correct processing and folding of their passenger proteins, thus increasing processing efficiency. The saccharolytic activity obtained from these constructs suggests that this is solely due to the secretion leader. In support of this hypothesis is the apparent lack of sequence homology between the amylase and MFα1 secretion leaders.

In conclusion, this study has resulted in a definite progress towards laying the foundation for the possible development of efficient starch-degrading *S. cerevisiae* strains that could eventually be optimised for consolidated bioprocessing (CBP) in which the production of amylolytic enzymes, the hydrolysis of starch-rich substrates and the fermentation of resulting sugars to a desired product (e.g. bio-fuel ethanol, low carbohydrate beer, grain whisky, etc.) occurs in a single step.

3.4 ACKNOWLEDGEMENTS

The authors thank the National Research Foundation (NRF) and the South African wine industry (WINETECH) for financial support.

3.5 REFERENCES

Bitter GA, Chen KK, Banks AR, Lai P (1984) Secretion of foreign proteins from *Saccharomyces cerevisiae* directed by α-factor fusions. Proc Natl Acad Sci USA 81:5330-5334

Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, Urdea MS, Valenzuela P, Barr PJ (1984) α-Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae.* Proc Natl Acad Sci USA 81:4642-4646

Caplan S, Green R, Rocco J, Kurjan J (1991) Glycosylation and structure of the yeast *MFα1* α-factor precursor is important for efficient transport through the secretory pathway. J Bacteriol 173:627-635

Cereghino GPL, Cregg JM (1999) Applications of yeast in biotechnology: protein production and genetic analysis. Curr Opin Biotechnol 10:422-427

Conesa A, Punt PJ, Van Luijk N, Van den Hondel CAMJJ (2001) The secretion pathway in filamentous fungi: a biotechnological view. Fungal Genet Biol 33:155-171

Eksteen JM, Van Rensburg P, Cordero Otero RR, Pretorius IS (2003) Starch fermentation by recombinant *Saccharomyces cerevisiae* strains expressing the α-amylase and glucoamylase genes from *Lipomyces kononenkoae* and *Saccharomycopsis fibuligera*. Biotechnol Bioeng (*in press*)

Gierasch LM (1989) Signal sequences. Biochemistry 28:923-930

Gietz RD, Schiestl RH (1991) Applications of high efficiency lithium acetate transformation of intact cells using single-stranded nucleic acids as carrier. Yeast 7:253-263

Kjeldsen T (2000) Yeast secretory expression of insulin precursors. Appl Microbiol Biotechnol 54:277286

Kjeldsen T, Brandt J, Andersen AS, Egel-Mitani M, Hach M, Petterson AF, Vad K (1996) A removable spacer peptide in an α-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in *Saccharomyces cerevisiae.* Gene 170:107-112

Kjeldsen T, Petterson AF, Hach M, Diers I, Havelund S, Hansen PH, Andersen AS (1997) Synthetic leaders with potential BiP binding mediate high-yield secretion of correctly folded insulin precursors from *Saccharomyces cerevisiae*. Protein Expression Purification 9:331-336

Kjeldsen T, Andersen AS, Hach M, Diers I, Nikolajsen J, Markussen J (1998) α-Factor pro-peptide N-linked oligosaccharides facilitate secretion of the insulin precursor in *Saccharomyces cerevisiae*. Biotechnol Appli Biochem 27:109-115

Miller JL, Glennon WE, Burton AL (1960) Measurement of carboxymethylcellulase activity. Anal Biochem 2:127-132

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463-5467

Shinde U, Inouye M (1993) Intramolecular chaperones and protein folding. TIBS 18:442-446

Simons JF, Ferro-Novick S, Rose MD, Helenius A (1995) BiP/Kar2p serves as a molecular chaperone during carboxypeptidase Y folding in yeast. J Cell Biol 130:41-49

Spencer-Martins I, Van Uden N. (1979) Extracellular amylolytic system of the yeast *Lipomyces kononenkoae*. Eur J Appl Microbiol 6:241-250

Steyn AJC, Pretorius IS (1995) Characterization of a novel α-amylase from *Lipomyces kononenkoae* and expression of its gene (*LKA1*) in *Saccharomyces cerevisiae*. Curr Genet 28:569-573

Sudberry PE (1996) The expression of recombinant proteins in yeasts. Curr Opin Biotechnol 7:517-524

Winther JR, Sorensen P (1991) Propeptide of carboxypeptidase Y provides a chaperone-like function as well as inhibition of the enzymatic activity. Proc Natl Acad Sci USA 88:9330-9334

Yanish-Peron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119

Zheng N, Gierasch LM (1996) Signal sequences: the same yet different. Cell 86:849-852

Zhu X, Yoshiji O, Jordan F, Inouye M (1989) Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. Nature 339:483-484

Zsebo KM, Lu H, Fieschko JC, Goldstein L, Davis J, Duk K, Suggs SV, Lai P, Bitter GA (1986) Protein secretion from *Saccharomyces cerevisiae* directed by the prepro-α-factor leader region. J Biol Chem 261 5858-5865

ADDENDUM

DEVELOPMENT OF SYNTHETIC SIGNAL SEQUENCES FOR HETEROLOGOUS PROTEIN SECRETION FROM *SACCHAROMYCES CEREVISIAE*

INTRODUCTION

The classical *Saccharomyces cerevisiae* system for the secretion of heterologous proteins utilises the *S. cerevisiae* mating pheromone α-factor (MFα1) secretion leader and the Kex2p endoprotease for protein maturation (Zsebo et al. 1986). The α-factor leader is composed of a prepro peptide, consisting of a 19 amino acid pre peptide and a 64 amino acid pro region, followed by a Kex2p cleavage site and a short spacer peptide.

Previous studies by Kjeldsen et al. (1996) reported a considerable improvement in proteolytic processing and fermentation yield of an α-factor leader/proinsulin fusion protein by including a spacer peptide after the Kex2p site. In addition to employing a spacer sequence, Kjeldsen et al. (1997) also developed a novel *S. cerevisiae* secretion system based on the use of synthetic α-factor leader peptides. The synthetic leaders, in combination with a spacer peptide, were equal to or better than the native α-factor leader in facilitating secretion of proinsulin.

In this research study we describe the construction and evaluation of synthetic secretion leaders and their ability to direct secretion of the *Lipomyces kononenkoae LKA1* encoded α-amylase enzyme from the *S. cerevisiae* expression-secretion system. The aim of the present study was to improve the fermentation yield obtained from the existing MFα1 leader through the addition of a C-terminal spacer peptide, followed by the introduction of semirandom point mutations within the leader-spacer sequence.

MATERIALS AND METHODS

Microbial strains and plasmids

The sources and relevant genotypes/descriptions of bacterial and yeast strains, together with the plasmids used in this study, are listed in Table A.1.

Growth media and culture conditions

Escherichia coli transformants were cultured in Luria Bertani (LB) broth (Sambrook et al. 1989) containing ampicillin (100 μg/ml). Yeasts were cultured in Yeast Peptone Dextrose (YPD) medium containing 1% yeast extract, 2% bacto-peptone and 2% glucose. The *S. cerevisiae* transformants were grown and selected on Synthetic Complete (SC) medium containing 2% glucose and 0.67% yeast nitrogen base without amino acids. No growth factors were added and uracil was used as the selectable marker. SC^{-Ura} medium containing either Phadebas starch

(20 Phadebas tablets per litre; Pharmacia Diagnostics, Uppsala, Sweden) or 2% corn starch (Sigma-Aldrich, Steinheim, Germany) as only carbon source was used to screen for amylolytic activity. All solid media contained 2% agar. Selective plates

Table A.1 Microbial strains and plasmids used in this study

Strain/plasmid[1]	Relevant genotype/description	Source/reference
***Escherichia coli* strains:**		
DH5α	*F⁻ recA1 endA1 gyrA96 thi hsdR17 supE44 relA1* Δ(*argF-Lac-ZYA*) (φ80d*lacZ*(M15)λ	Invitrogen
***S. cerevisiae* strains:**		
Σ1278b	*Matα URA3*	Liu et al. 1993
Σ1278b-YEp352	Σ1278b transformed with YEp352	This study
Σ1278b-pSL1	Σ1278b transformed with pSL1	This study
Σ1278b-pSL2	Σ1278b transformed with pSL2	This study
Σ1278b-pSL3	Σ1278b transformed with pSL3	This study
Σ1278b-pSL3^M	Σ1278b transformed with library of pSL3^M	This study
Plasmids:		
pGEM	Ap^R *f1 ori LacZ T7$_P$ SP6$_P$*	Yanish-Peron et al. 1985
YEp352	Ap^R Tc^R *PGK1$_{PT}$ URA3 2μm*	This laboratory
pSL1	YEp352 containing the *LKA1* gene without any secretion leader	This study
pSL2	YEp352 containing the *LKA1* gene under secretory direction of MFα1 secretion leader	This study
pSL3	YEp352 containing the *LKA1* gene under secretory direction of MFα1Spacer secretion leader	This study
pSL3^M	YEp352 containing the *LKA1* gene under secretory direction of mutated MFα1Spacer secretion leader	This study

[1] M denotes mutated MFα1Spacer secretion leader

were incubated at 30°C for 2 to 3 days and the appearance of halos around the colonies was taken to be indicative of starch utilisation (Steyn and Pretorius 1995). Bacteria and yeasts were routinely cultured at 37°C and 30°C respectively.

DNA manipulations, sequencing and amplification

Standard techniques employed in the cloning and manipulation of DNA fragments were carried out as described by Sambrook et al. (1989). Restriction endonucleases, T4 DNA ligase and Expand High Fidelity polymerase were purchased from Roche Biochemical Products (Randburg, South Africa) and used as recommended by the supplier.

The isolation of plasmid DNA from the yeast transformants was done according to Adam and Polaina (1991).

Standard polymerase chain reaction (PCR) amplifications and sequencing reactions were performed as described in Chapter 3.

Cloning and generation of the synthetic leaders

The construction of the MFα1Spacer secretion leader is described in Chapter 3.

The various modifications introduced to the above-mentioned 282 bp *Eco*RI-*Hin*dIII DNA fragment were generated by PCR-based mutational reactions. The use of an imbalanced nucleotide mixture (dNTPs) combined with the addition of 0.3 mM $MnCl_2$ to an otherwise standard PCR reaction, gave rise to semirandom point mutations within the MFα1Spacer sequence. Sequencing of the modified leaders confirmed an average mutational frequency of 6-8%, i.e. 17 to 23 bp out of 282 were mutated. The synthetic secretion leaders were isolated, amplified by standard PCR and cloned using standard methods

Yeast transformation

The laboratory strain *S. cerevisiae* Σ1278b was separately transformed with each of the plasmids YEp352, pSL1, pSL2, pSL3 as well as the library pool of $pSL3^{M}$ containing the synthetic leaders.

Isolation of extracellular protein

SC^{-Ura} broth (50 ml) was inoculated with a suitable volume of stationary phase pre-culture to obtain an initial absorbance of 0.1 at 600 nm. Cells were grown at 30°C for 48 h. After centrifugation at 5000 rpm for 5 min, the supernatant containing the extracellular protein was carefully removed and used in the enzyme assays.

Enzyme assays

Fermentation yield was spectrophotometrically (OD_{595}) determined, using 200 µl of culture supernatant, by means of the Bradford Total Protein assay.

Extracellular α-amylase activity was quantitatively measured using the Phadebas Amylase Test (Pharmacia Diagnostics, Uppsala, Sweden) assay (Steyn and

Pretorius 1995). One unit of enzyme activity was defined as the amount of enzyme that released 1 mmol of reducing sugar/min.

RESULTS AND DISCUSSION

Isolation and assay of putative 'positives'

Approximately 12000 transformants, containing the library pool of mutated MFα1Spacer secretion leaders, were screened on minimal media containing Phadebas starch as only carbon source. To reduce background and to assist us in our efforts in isolating the secretion leader(s) with improved secretion abilities, we replica-plated the putative 'positives' onto minimal media containing corn starch as sole carbon source. Transformants were selected only on phenotypical observations, such as the size of a halo-producing colony and the size of the halo itself.

120 putative 'positive' transformants were selected and the α-amylase activity and extracellular protein present in the culture supernatant were determined for each of the transformants (as described in Materials and Methods). Table A.2 summarises the fermentation yield obtained for the different transformants. As was expected, no

Table A.2 Fermentation yield of an α-amylase enzyme expressed in *S. cerevisiae* by different secretion leaders. Values are means from three independent experiments; deviations are within 12%

Strain[1]	48h Culture OD_{600}	Bradford assay OD_{595}	Yield (μg/ml)
Σ1278b-YEp352	16.4	0.002	0
Σ1278b-pSL1	16.9	0.004	0
Σ1278b-pSL2	16.7	0.122	59.22
Σ1278b-pSL3	15.8	0.168	81.55
Σ1278b-pSL3^M (Mut 1)	17.9	0.069	33.5
Σ1278b-pSL3^M (Mut 2)	16.3	0.047	22.82
Σ1278b-pSL3^M (Mut 3)	18.3	0.092	44.66
Σ1278b-pSL3^M (Mut 4)	16.5	0.065	31.55
Σ1278b-pSL3^M (Mut 5)	16.1	0.092	44.66
Σ1278b-pSL3^M (Mut 6)	16.7	0.103	50.0
Σ1278b-pSL3^M (Mut 7)	16.6	0.093	45.15
Σ1278b-pSL3^M (Mut 8)	16.8	0.122	59.22

[1] Σ1278b-pSL3^M (Mut 1-8) represents eight randomly chosen transformants

extracellular protein was detected in the strain expressed without a secretion leader (Σ1278b-pSL1). Transformants Σ1278b-pSL2 and Σ1278b-pSL3 (as the control strain) resulted in fermentation yields of 59.22 and 81.55 μg/ml respectively. In contrast to these values, were those obtained for the transformants expressing the Lka1p enzyme under the secretory direction of the synthetic leaders. Transformants Σ1278b-pSL3^{M} (Mut 1-8) resulted in much lower fermentation yields after 48 h of growth compared to the control strain. This eventually proved to be the norm for all the isolated transformants utilising the synthetic leaders.

Modifications to the Bradford Total Protein assay had no effect on the values obtained for the extracellular protein produced by these transformants. As seen in Table A.3, the addition of a protease inhibitor cocktail to the growth medium, as was expected since *S. cerevisiae* do not secrete any active proteases, made no difference to our observations.

Table A.3 Fermentation yield of an α-amylase enzyme expressed in *S. cerevisiae* by different secretion leaders under conditions of inclusion/omission of protease inhibitors. Values are means from three independent experiments; deviations are within 12%

Strain	48h Culture OD_{600}	Yield with protease inhibitors (μg/ml)	Yield without protease inhibitors (μg/ml)
Σ1278b-YEp352	16.3	0	0
Σ1278b-pSL1	16.1	0	0
Σ1278b-pSL2	15.9	54.08	52.43
Σ1278b-pSL3	16.8	82.05	85.44
Σ1278b-pSL3^{M} (Mut 1)	15.8	36.49	34.37
Σ1278b-pSL3^{M} (Mut 2)	16.3	26.52	22.84
Σ1278b-pSL3^{M} (Mut 3)	16.6	45.89	44.69
Σ1278b-pSL3^{M} (Mut 4)	15.8	33.75	33.45
Σ1278b-pSL3^{M} (Mut 5)	16.7	44.61	44.60
Σ1278b-pSL3^{M} (Mut 6)	16.5	51.60	50.04
Σ1278b-pSL3^{M} (Mut 7)	16.4	41.83	44.76
Σ1278b-pSL3^{M} (Mut 8)	15.9	60.03	64.61

The extracellular amylolytic activity produced by these transformants follows a similar pattern. As summarised in Table A.4, the control strain resulted in the highest α-amylase activity (444 U/l) while considerably lower values were obtained for the strains containing the synthetic leaders. Transformants Σ1278b-pSL3^{M} (Mut 5) and

Σ1278b-pSL3^M (Mut 8) resulted in 155 and 177 U/l of α-amylase activity respectively. Σ1278b-pSL1 showed no amylolytic activity. Similarly to the Bradford assay, modifications to the Phadebas Amylase Test assay, e.g. changes in incubation times and temperatures, enzyme buffer used, etc., had no effect on the observed values.

Table A.4 The extracellular amylolytic activity of an α-amylase enzyme expressed in *S. cerevisiae* by different secretion leaders. The values are means from three independent experiments; deviations are within 12%

Strain	48h Culture OD_{600}	Phadebas Amylase Test OD_{620}	α-Amylase activity (U/l)
Σ1278b-YEp352	17.73	0	0
Σ1278b-pSL1	17.9	0	0
Σ1278b-pSL2	17.8	0.007	<27
Σ1278b-pSL3	17.4	0.735	444
Σ1278b-pSL3^M (Mut 1)	17.18	0.482	306
Σ1278b-pSL3^M (Mut 2)	16.44	0.437	280
Σ1278b-pSL3^M (Mut 3)	16.74	0.739	449
Σ1278b-pSL3^M (Mut 4)	15.14	0.628	390
Σ1278b-pSL3^M (Mut 5)	17.06	0.288	155
Σ1278b-pSL3^M (Mut 6)	17.38	0.627	385
Σ1278b-pSL3^M (Mut 7)	17.6	0.584	362
Σ1278b-pSL3^M (Mut 8)	18.58	0.267	177

Isolation and sequencing of plasmid DNA

Free plasmid DNA was isolated from Σ1278b-pSL3^M (Mut 1-8) and sent for sequencing (see Material and Methods). The sequence analysis showed that the synthetic secretion leaders had an average mutational frequency of 30-40% compared to the non-mutated MFα1Spacer leader. Further sequencing of plasmid DNA isolated from the bulk of the 120 transformants confirmed this result. Reintroducing the isolated plasmids into strain Σ1278b resulted in similar phenotypical conditions and assay values previously observed for these transformants.

The factor(s) responsible for the much higher than anticipated mutational frequency is still open for debate. We hypothesise that it is a question of '$MnCl_2$ contamination' and that the (6-8%) mutated MFα1Spacer fragment used as template DNA in the standard PCR amplification reactions, could have contained residual

$MnCl_2$. This would explain the higher rate of mutation and the resulting deleterious effect on the secretion efficiency of the synthetic secretion leaders.

This approach, the use of PCR-based mutational reactions to improve secretion, was abandoned in favour of the work described in Chapter 3.

REFERENCES

Adam AC, Polaina J (1991) Construction of a *Saccharomyces cerevisiae* strain able to ferment cellobiose. Curr Genet 20:5-8

Kjeldsen T, Brandt J, Andersen AS, Egel-Mitani M, Hach M, Petterson AF, Vad K (1996) A removable spacer peptide in an α-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in *Saccharomyces cerevisiae*. Gene 170:107-112

Kjeldsen T, Petterson AF, Hach M, Diers I, Havelund S, Hansen PH, Andersen AS (1997) Synthetic leaders with potential BiP binding mediate high-yield secretion of correctly folded insulin precursors from *Saccharomyces cerevisiae*. Protein Expression Purification 9:331-336

Liu H, Styles CA, Fink GR (1993) Elements of the yeast pheromone response pathway required for filamentous growth of diploids. Science 262:1741-1744

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Steyn AJC, Pretorius IS (1995) Characterization of a novel α-amylase from *Lipomyces kononenkoae* and expression of its gene (*LKA1*) in *Saccharomyces cerevisiae*. Curr Genet 28:569-573

Yanish-Peron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119

Zsebo KM, Lu H, Fieschko JC, Goldstein L, Davis J, Duk K, Suggs SV, Lai P, Bitter GA (1986) Protein secretion from *Saccharomyces cerevisiae* directed by the prepro-α-factor leader region. J Biol Chem 261:5858-5865

CHAPTER 4

GENERAL DISCUSSION AND CONCLUSIONS

4. GENERAL DISCUSSION AND CONCLUSIONS

4.1 CONCLUDING REMARKS AND OTHER PERSPECTIVES

Research efforts have increased our specific knowledge of the secretion process in yeast, highlighting its similarities with the secretion pathway of higher eukaryotes. This knowledge has been used in our efforts to improve and extend the biotechnological uses of the yeast secretory machinery (Sagt et al. 2000; Conesa et al. 2001). Bottlenecks responsible for low product yields have been identified and novel ideas for and approaches to overcoming them have been employed (Clements et al. 1991; Cereghino and Cregg 1999).

This research study is written almost 30 years after the groundbreaking proposal for a signal peptide was made by Blobel and Dobberstein (1975). Today we know that most of the predictions made in 1975 were in fact correct; all proteins are synthesised with an N-terminal signal sequence of varying composition and length that directs their transport to and across the ER membrane (Brake et al. 1984; Zsebo et al. 1986; Cleves and Bankaitis 1992).

Interest in signal peptides is high. While considerable progress has been made in the last few years in elucidating the mechanisms involved in protein translocation and transport, much confusion still exists about the specific role(s) of signal peptides in protein folding and secretion (Zheng and Gierasch 1996; Kjeldsen et al. 1996, 1997). We know that these molecular "zip codes" contain information necessary for the targeting and translocation of all proteins; that a seemingly unimportant interaction with the signal recognition particle prevents the passenger protein from folding into a translocation-incompetent conformation; and that overall length and hydrophobicity of the signal peptide core are crucial to its function (Von Heijne 1981, 1985; Bird et al. 1987; Gierasch 1989; Caplan et al. 1991). However, evidence is accumulating that indicates that the secretion leader has additional, passenger protein-specific functions.

It has already been demonstrated that the pro region renders the passenger protein catalytically inactive, especially in the case of enzymes (Winther and Sorensen 1991; Winther et al. 1994; Eder and Fersht 1995). It is now believed that the pro regions might also facilitate the proper, three-dimensional folding of the proteins they are attached to and thus function as intramolecular chaperones. This was shown to occur for a variety of proteins, with the carboxypeptidase Y of *S. cerevisiae*, the subtilisin E of *Bacillus subtilis* and bovine pancreatic trypsin inhibitor (BPTI) being the best studied to date (Zhu et al. 1989; Kobayashi and Inouye 1992; Weissman and Kim 1992; Takahashi et al. 2001). Intramolecular chaperones differ from the molecular chaperones present in the ER in that they are covalently attached to the N-terminus of proteins and only mediate the folding of their passenger protein. They function in a 1:1 ratio with their substrate and facilitate

folding by lowering the activation energy required for these processes (Shinde and Inouye 1993; Le Loir et al. 2001).

The present work suggests that the *LKA1* and *LKA2* α-amylase gene products of *L. kononenkoae* also require the participation of their secretion leaders in the folding process. We examined several leader peptides with the aim of improving the secretion of the amylase enzymes from the heterologous host *S. cerevisiae*. The yeast *L. kononenkoae* is legendary for its ability to degrade raw starch and the improved secretion of its 'most effective' amylase genes from *S. cerevisiae* paves the way for a future one-step starch utilisation process (Spencer-Martins and Van Uden 1979; Eksteen et al. 2003).

Three different sets of constructs containing the *LKA1* and *LKA2* genes under the secretory direction of either their own native secretion leader, the MFα1 leader, or a synthetic MFα1 secretion leader with a C-terminal spacer peptide (EEGEPK) were successfully expressed in *S. cerevisiae*. The inclusion of a spacer peptide in the latter set of constructs ensures improved Kex2p proteolytic processing of the leader/protein fusion, a prerequisite for efficient secretion. The strategy is based on the Lys-Arg spacer sequences preceding each of the four α-factor pheromone peptides of ppαF and was used with notable success by Kjeldsen in the secretion of insulin (Kjeldsen 2000). Plate assays revealed amylolytic activity for the three sets of constructs, demonstrating a functional expression system. When assayed for saccharolytic activity in the extracellular medium, strains bearing the *LKA1* and *LKA2* constructs under their native secretion signals resulted in the highest activity. These observations are in marked contrast to the values obtained for the total extracellular protein produced by these transformants. Transformants expressing the amylase genes under secretory direction of the synthetic leader peptide resulted in a significantly higher fermentation yield. Surprisingly, this increase in yield was not reflected in the saccharolytic activity; strains utilising the synthetic leader resulted in considerable lower extracellular activity compared to the native *LKA1* and *LKA2* leaders.

There are several possible reasons for the unexpected differences between the enzymes secreted and the activity of the secreted enzymes. One possibility, and certainly the most appealing, is that of secretion efficiency versus processing efficiency.

Previous secretion studies have indicated the importance of an optimum Kex2p cleavage site; Kjeldsen (2000) improved MFα1-directed secretion of the insulin precursor by introducing a spacer peptide containing the Kex2p target sequences. We employed a similar strategy with our synthetic leaders and the fermentation yield obtained with these constructs is consistent with what was expected.

The *LKA1* and *LKA2* α-amylases, however, are enzymes with complex tertiary structures and unique three-dimensional folding patterns. We hypothesise that the (optimal) processing of these enzymes requires the additional participation of their native secretion leaders. Enzyme activity is based, among other factors, on

three-dimensional conformation. It is therefore tempting to think that the native amylase leaders might function as intramolecular chaperones and assist in the correct processing and folding of their passenger proteins. The saccharolytic activity obtained from these constructs suggests that this is solely due to the secretion leader. In support of this hypothesis is the apparent lack of sequence homology between the LKA1/LKA2 and MFα1 secretion leaders, which could explain the latter's inability to function as a "surrogate" chaperone for the amylases.

In summary, by optimizing proteolytic processing of the MFα1 secretion leader, we improved the secretion efficiency of both amylase enzymes from the current *S. cerevisiae* expression system. The activity of these enzymes requires efficient and complete folding and we hypothesise that the native leaders may function as intramolecular chaperones in these processes. However, additional work needs to be done to elucidate the structural relationship between the leader and the protein.

4.2 LITERATURE CITED

Bird P, Gething MJ, Sambrook J (1987) Translocation in yeast and mammalian cells: not all signals are functionally equivalent. J Biol Chem 105:2905-2914

Blobel G, Dobberstein B (1975) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains in membrane-bound ribosomes of murine myeloma. J Cell Biol 67:835-851

Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, Urdea MS, Valenzuela P, Barr PJ (1984) α-Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae.* Proc Natl Acad Sci USA 81:4642-4646

Caplan S, Green R, Rocco J, Kurjan J (1991) Glycosylation and structure of the yeast *MFα1* α-factor precursor is important for efficient transport through the secretory pathway. J Bacteriol 173:627-635

Cereghino GPL, Cregg JM (1999) Applications of yeast in biotechnology: protein production and genetic analysis. Curr Opin Biotechnol 10:422-427

Clements JM, Catlin GH, Price MJ, Edwards RM (1991) Secretion of human epidermal growth factor from *Saccharomyces cerevisiae* using synthetic leader sequences. Gene 106:267-272

Cleves AE, Bankaitis VA (1992) Secretory pathway function in *Saccharomyces cerevisiae*. Adv Microbial Physiol 33:73-144

Conesa A, Punt PJ, Van Luijk N, Van den Hondel CAMJJ (2001) The secretion pathway in filamentous fungi: a biotechnological view. Fungal Genet Biol 33:155-171

Eksteen JM, Van Rensburg P, Cordero Otero RR, Pretorius IS (2003) Starch fermentation by recombinant *Saccharomyces cerevisiae* strains expressing the α-amylase and glucoamylase genes from *Lipomyces kononenkoae* and *Saccharomycopsis fibuligera*. Biotechnol Bioeng (*in press*)

Eder J, Fersht AR (1995) Pro-sequence-assisted protein folding. Mol Microbiol 16:609-614

Gierasch LM (1989) Signal sequences. Biochemistry 28:923-930

Kjeldsen T (2000) Yeast secretory expression of insulin precursors. Appl Microbiol Biotechnol 54:277-286

Kjeldsen T, Brandt J, Andersen AS, Egel-Mitani M, Hach M, Petterson AF, Vad K (1996) A removable spacer peptide in an α-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in *Saccharomyces cerevisiae*. Gene 170:107-112

Kjeldsen T, Petterson AF, Hach M, Diers I, Havelund S, Hansen PH, Andersen AS (1997) Synthetic leaders with potential BiP binding mediate high-yield secretion of correctly folded insulin precursors from *Saccharomyces cerevisiae*. Protein Expression Purification 9:331-336

Kobayashi T, Inouye M (1992) Functional analysis of the intramolecular chaperone: mutational hot spots in the subtilisin pro-peptide and a second-site suppressor mutation within the subtilisin molecule. J Mol Biol 226:931-933

Le Loir Y, Nouaille S, Commissaire J, Brétigny L, Gruss A, Langella P (2001) Signal peptide and propeptide optimization for heterologous protein secretion in *Lactococcus lactis*. Appl Environ Microbiol 67:4119-4127

Sagt CMJ, Kleizen B, Verwaal R, De Jong MDM, Müller WH, Smits A, Visser A, Boonstra J, Verkleij AJ, Verrips CT (2000) Introduction of an N-glycosylation site increases secretion of heterologous proteins in yeasts. Appl Environ Microbiol 66:4940-4944

Shinde U, Inouye M (1993) Intramolecular chaperones and protein folding. TIBS 18:442-446

Spencer-Martins I, Van Uden N. (1979) Extracellular amylolytic system of the yeast *Lipomyces kononenkoae*. Eur J Appl Microbiol 6:241-250

Takahashi S, Ueda M, Tanaka A. (2001) Function of the prosequence for in vivo folding and secretion of active *Rhizopus oryzae* lipase in *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol 55:454-462

Von Heijne G (1981) On the hydrophobic nature of signal sequences. Eur J Biochem 116:419-422

Von Heijne G (1985) Signal sequences: the limits of variation. J Mol Biol 184:99-105

Weissman JS, Kim PS (1992) The pro region of BPTI facilitates folding. Cell 71:841-851

Winther JR, Sorensen P (1991) Propeptide of carboxypeptidase Y provides a chaperone-like function as well as inhibition of the enzymatic activity. Proc Natl Acad Sci USA 88:9330-9334

Winther JR, Sorensen P, Kielland-Brandt MC (1994) Refolding of a carboxypeptidase Y folding intermediate *in vitro* by low-affinity binding of the proregion. J Biol Chem 269:22007-22013

Zheng N, Gierasch LM (1996) Signal sequences: the same yet different. Cell 86:849-852

Zhu X, Yoshiji O, Jordan F, Inouye M (1989) Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. Nature 339:483-484

Zsebo KM, Lu H, Fieschko JC, Goldstein L, Davis J, Duk K, Suggs SV, Lai P, Bitter GA (1986) Protein secretion from *Saccharomyces cerevisiae* directed by the prepro-α-factor leader region. J Biol Chem 261:5858-5865