Proteolytic processing in the secretory pathway of *Aspergillus niger*

Ruud Jalving

Promotor:	Prof. dr. ir. A.J.J. van Ooyen Hoogleraar in de Toegepaste Microbiële Genetica Wageningen Universiteit
Co-promotor:	Dr. P.J. Schaap Universitair docent, leerstoelgroep Microbiologie Wageningen Universiteit
Promotiecommissie:	Prof. dr. D. B. Archer, University of Nottingham Prof. dr. C.A.M.J.J van den Hondel, TNO Voeding Prof. dr. ir. F.M. Rombouts, Wageningen Universiteit Prof. dr. W.M. de Vos, Wageningen Universiteit

Proteolytic processing in the secretory pathway of *Aspergillus niger*

Ruud Jalving

Proefschrift
Ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit
Prof. dr. ir. L. Speelman
in het openbaar te verdedigen op
vrijdag 4 februari 2005
des namiddags om half twee in de Aula

Proteolytic processing in the secretory pathway of *Aspergillus niger*

Ruud Jalving

Thesis Wageningen University. – With ref. – With summary in Dutch

ISBN 90-8504-129-5

Subject headings: *Aspergillus* / protease / maturation / intracellular / proteolysis / processing

Contents

Chapter 1

Introduction

Fungi play an important role in our daily life. They are both harmful and beneficial for our health, agriculture and environment. Fungal pathogens are lethal for immunocompromised patients, but other non-pathogenic fungi are used for the production of pharmaceutical products. Fungi are the most important group of plant pathogens, however, mycorrhizal fungi that grow interdependently with plant roots are critical for nutrient uptake by plants. In addition, filamentous fungi are saprophytes and play an important role in nutrient recycling.

The number of species within the kingdom of fungi is currently estimated to be around 1.5 million (Borkovich et al., 2004). The majority of these species is filamentous. Their substrates are often insoluble polymers, and need to be degraded into soluble monomers or small soluble oligomers before these can be used as nutrition. For this purpose filamentous fungi secrete a wide spectrum of enzymes, such as glycosyl hydrolases, proteases, lipases, and phosphatases.

While filamentous fungi naturally secrete a large number of enzymes, the yeast *Saccharomyces cerevisiae* does not. Because of its simplicity, the secretory pathway of *S. cerevisiae* has been studied in great detail and functions as model for the secretory pathway of eukaryotes in general. To outline the fundamentals of the eukaryotic secretory pathway and to illustrate the environment for the main topic of this chapter, proteolysis, a description of the yeast secretory pathway follows below.

The fungal secretory pathway

Proteins are synthesized by ribosomes in the cytosol. If their final destination is the ER, the Golgi, the vacuole, the plasma membrane or extracellular environment they will reach their target after a journey through the secretory pathway (Fig. 1). The endoplasmic reticulum is the entry point into the secretory pathway. The synthesized precursor proteins contain a hydrophobic prepeptide which mediates their translocation into the ER. This prepeptide is removed from most proteins by signal peptidase upon entry. Some proteins keep their prepeptide as membrane anchor.

As the precursor proteins progress on their journey through the secretory pathway they are gradually altered until they obtain their mature form (Shinde and Inouye, 2000). These post-translational alterations include site specific proteolysis and amino acid modifications (such as phosphorylation, glycosylation, deamination, and acetylation). Via the ER the proteins travel

towards the Golgi, which is in effect a branching point of the secretory pathway and the vacuolar pathway (Nakano, 2004). There are at least two protein sorting pathways leading from the Golgi to the vacuole, and there are indications that more than one pathway exists for protein sorting to the plasma membrane. Protein transport from one organelle to another is mediated by vesicles. Membrane proteins without sorting signals are transported via the endosomes to the vacuole. Soluble proteins are transported by default towards the plasma membrane, where they are released into the extracellular environment.

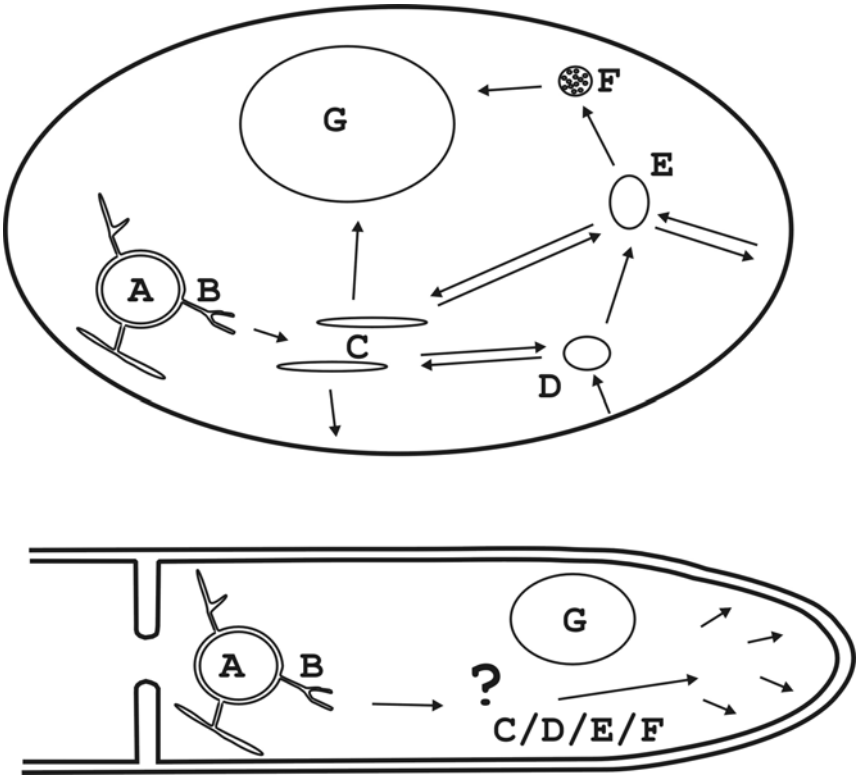


Fig. 1. The secretory pathway of yeast (upper panel) and filamentous fungi (lower panel). The indicated organelles are: A) nucleus; B) endoplasmic reticulum; C) Golgi; D) post Golgi endosome (PGE) / early endosome; E) prevacular endosome (PVE) / late endosome; F) multivesicular body; G) vacuole. The question mark indicates organelles that have not been visualised in filamentous fungi. The arrows represent the protein transport pathways between the organelles.

The secretory pathway of filamentous fungi is assumed to be similar to the secretory pathway of yeast. A comparative analysis of the genome of *Neurospora crassa* with the genome of *S. cerevisiae* revealed that *N. crassa* has homologues of all the Rab-type GTPases present in the yeast genome (Borkovich et al., 2004). Rab-type GTPases are required for vesicle targeting and fusion with the acceptor organelle. In yeast, these GTPases have a high specificity for a single transport pathway, making them very effective molecular markers for the individual organelles. Despite the presence of these putative Rab genes, indicating the presence of yeast-like secretory pathway structures, most of these structures, such as the Golgi, have never been visualised in filamentous fungi. However, the comparative analysis revealed additional Rab genes being present in the *N. crassa* genome which are also present in mammalian genomes, but not in the yeast genome, suggesting that the secretory pathway of filamentous fungi is more complex than the yeast secretory pathway. Considering the high secretion potential of filamentous fungi compared to yeast, this is not surprising.

Protein production by filamentous fungi

Recognition of this secretion potential has resulted in large scale industrial use of filamentous fungi. A variety of proteins has been isolated from filamentous fungi. An important group of enzymes isolated from fungi are the pectic enzymes (Alkorta et al., 1998). These enzymes have a long history of use in the yield increasing and clarifying of fruit juices. A few other fungal enzyme families that are isolated on an industrial scale are the glucoamylases, the cellulases, and the xylanases. Production improvement programs have made protein yields above 1 g l^{-1} medium quit common (Gouka et al., 1997; Verdoes et al., 1995). The yield of several proteins is even above 10 g l^{-1}.

A more recent application of filamentous fungi is their use as cell factories for heterologous protein production (Archer, 2000). The advantage of the use of these organisms in comparison to prokaryotes is their capability to produce eukaryotic proteins needing post-translational modifications like glycosylation and disulfide bridge formation (Maras et al., 1999). The production of heterologous proteins in yeasts, such as *S. cerevisiae* and to a lesser extent *Pichia pastoris*, resulted in several cases in unwanted hyperglycosylation of the heterologous protein (Cereghino and Cregg, 2000). This hyperglycosylation is not found if these proteins are produced in filamentous fungi.

Nevertheless, heterologous non-fungal proteins are still produced at far lesser quantities by filamentous fungi than fungal proteins (Gouka et al., 1997). Human lactoferrin production levels vary from 5 mg l^{-1} to more than 2 g l^{-1}. Attempts to produce human interleukin-6 resulted in yields between 1 mg l^{-1} and 40 mg l^{-1}. Obviously non-fungal proteins have not been optimised during evolution to be secreted efficiently by the fungal secretory pathway. As a result the production efficiency of these proteins can be compromised at any stage of the secretion process. Problems were found to occur at the levels of: mRNA stability, codon-usage, translocation, folding, sorting, and protease susceptibility (Maras et al., 1999).

Especially proteases are a major problem for heterologous protein production, they are present at every stage of the secretion process and their action on the foreign protein has a catastrophic influence on the yield (van den Hombergh, 1997). Disruption of the major extracellular proteases of *Aspergillus niger* resulted in an enormous reduction of extracellular product degrading activity. Degradation of the proteolytic susceptible *A. niger PelB* protein was reduced 1000-fold in an *A. niger* production strain, in which three major proteases were disrupted. Also intracellular proteases are important for protein secretion. They play an important role in protein modification and regulation. However until recently not much was known about proteases present in the secretory pathway of filamentous fungi.

Protease nomenclature

Before a detailed description is given of intracellular proteases in the fungal secretory pathway a short introduction into protease nomenclature is needed. The two most important characteristics used for the classification of proteases are the substrate specificity and composition of the catalytic site.

Starting with substrate specific nomenclature, exoproteases hydrolyse proteins at the termini, while endoproteases cleave internal amino-acid bonds. Aminopeptidases and carboxypeptidases are exoproteases which liberate an amino acid from the amino-terminus and carboxy-terminus of the protein respectively. The protease name may also be based on the specificity of the protease. For example, oligopeptidases prefer small peptides, dipeptidases prefer dipeptides, and methionine aminopeptidases remove the amino acid at the amino-terminus, if the penultimate amino acid is a methionine residue.

A different way of classification of the proteases is based upon the composition of the active site. Using this classification method, proteases can be classified in four groups, the serine, metallo, aspartic and cysteine

proteases. A brief explanation of these four groups is given here; a more elaborate description is given by van den Hombergh (1996). Each group has a different proteolytic mechanism and therefore different enzymatic characteristics and inhibitors. The cysteine proteases are a rare group of proteases, and none have been found in filamentous fungi so far.

The serine proteases contain a catalytic triad, which is formed by a serine, a histidine and an aspartic residue. PMSF is a common inhibitor of serine proteases. The serine protease families are grouped into 6 clans. The two major clans are the (chymo)trypsin-like and the subtilisin-like clan.

The catalytic site of aspartic proteases is formed by two conserved aspartic residues. The exact mechanism of this clan of proteases has been unknown until recently (Northrop, 2001). The active site appears to be mono-protonated. The single hydrogen bond keeps both aspartic residues in a single plane together with a water molecule. Pepstatin A inhibits most aspartic proteases. The aspartic proteases are roughly divided in pepsin-like and non-pepsin-like aspartic proteases.

The metallo proteases need a metal ion for their activity. Most metallo endoproteases contain zinc as metal ion (Jiang and Bond, 1992) and contain three or four amino acids that function as zinc ligand. Based on the first two zinc ligands (both histidine residues) the zinc metallo proteases can be divided roughly in two categories. One category of metallo proteases is characterised by the presence of the HXXEH motif. This category includes bacterial protease III and the human insulin degrading enzyme (Jiang and Bond, 1992). The other category contains metallo proteases with a HEXXH motif. This second category can be subdivided on the basis of the third zinc ligand. One of these subgroups, containing the thermolysin family, uses glutamic acid as third ligand. The other group, containing the astacin family, uses histidine as third ligand and contains tyrosine as an additional fourth ligand. EDTA and 1,10-phenanthroline are effective inhibitors of the metallo proteases. The largest superfamily of metallo proteases is the zincins family of proteases (Hooper, 1994).

To indicate how the amino acid residues of the substrate are positioned in relation to the active sites of the protease the nomenclature as shown in Fig. 2 is used. The specificity of most proteases depends on the pre-scissile amino acid residue(s) of the substrate. This is probably the reason why the effect of the post-scissile amino acid residue(s) on the protease efficiency is often not examined.

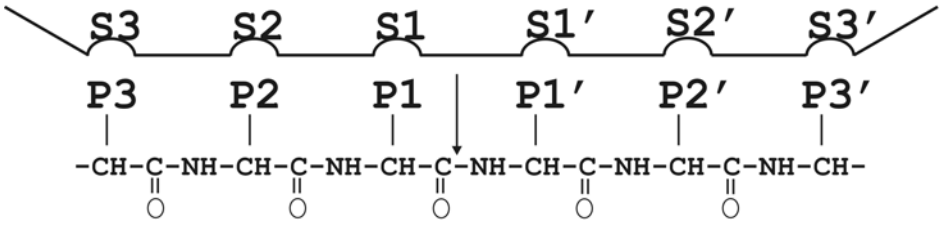


Fig. 2. Proteolytic cleavage site definition. The protein sequence below represents the substrate; the upper structure represents the protease. The amino acid bond cleaved by the protease is called the scissile bond and is indicated in the figure by the arrow. The pre-scissile amino acid residues are numbered from the scissile bond towards the N-terminus in the form of P1, P2, P3, and so on. The first amino acid following the scissile bond is designated P1'. The remaining amino acid residues are numbered in a similar fashion P2', P3', and so on. The protease subsites that interact with these amino acid residues are numbered in a similar fashion S1- S3, and S1' - S3'.

Proteolysis in the fungal secretory pathway

Several proteases present in the secretory pathway have a proteolytic processing function. In contrast to their extracellular degrading partners, their function is needed to ensure that the proteins of interest are correctly produced. The proteolysis of the precursor protein can result in cleavage of different subunits, cleavage of identical copies, removal of a peptide fragment, or removal of amino acid residues at either one of the termini. The proteolytic processing by yeast of killer toxin, pheromones, and carboxypeptidase Y are all well documented and together serve as good examples for these four types of processing.

K28 preprotoxin, a yeast killer toxin precursor consists of the two polypeptide chains of mature killer toxin, connected by an internal peptide (Schmitt and Breinig, 2002). This combined polypeptide is preceded by the so called prepropeptide (Fig. 3). After covalent linkage of the two domains via a disulphide bridge the connecting peptide is removed via proteolysis by the KEX2 endoprotease. The cleavage sites at the borders of the internal peptide contain both a specific amino acid recognition motif -Ser-Arg- and -Lys-Arg-, respectively. The remaining carboxy-terminal basic amino acid residues in the α- and β-chain are removed by the KEX1 carboxypeptidase. The mature killer toxin contains the HDEL endoplasmic reticulum retention

signal. This retention signal is necessary for the retrograde transport of the toxin towards the endoplasmic reticulum after receptor-mediated endocytosis in susceptible yeast strains. This retention signal is masked in the protoxin, but after the removal of the carboxy-terminal basic residues it is activated.

The yeast pheromone α-factor is produced as a polypeptide consisting of either 2 or 4 copies of the end product. Proteolytic processing of these precursor proteins involves the removal of the N-terminal propeptide and separation of the different copies of the end product. Dibasic amino acid motifs function as recognition sites for this endoproteolytic cleavage. The liberated α-factor copies contain N-terminal dipeptidyl residues that need to be removed. The penultimate alanine present in these residues ensures recognition and subsequent removal by the yeast dipeptidyl aminopeptidase DPAP A. This exopeptidase, which belongs to the proline specific proteases recognises both penultimate alanine and proline amino acid residues.

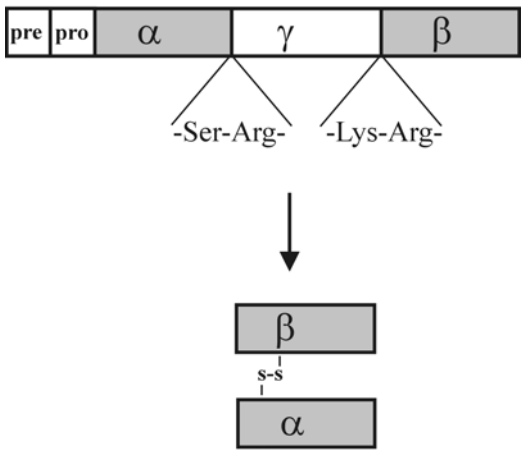


Fig. 3. Proteolytic processing of the yeast k28 killer toxin precursor. The processing of the killer toxin involves the removal of the N-terminal prepropeptide, and removal of the internal γ peptide following the formation of a sulphide bridge between the α and β domains.

In contrast to **α**-factor, of which an N-terminal peptide is removed, the yeast mating pheromone **a**-factor has a carboxy-terminal peptide that needs to be removed (Tam et al., 2001). This carboxy-terminal peptide is removed by CAAX proteases after the cysteine near the carboxy-terminus has been prenylated.

The proteolytic processing of carboxypeptidase Y (CPY) has been studied extensively. After entry into the ER the propeptide of proCPY remains attached to the carboxypeptidase until it has arrived at its final destination, the vacuole. The propeptide of CPY is essential for correct folding of the protease.

An N-terminal propeptide is present at the N-terminus of many proteins. Of the examples above, the precursors for killer toxin, α-factor and carboxypeptidase Y all contain an N-terminal propeptide. Many proteases remove their propeptide in an autocatalytic manner. Most of the propeptides that are not removed in an autocatalytic event contain a dibasic motif as the recognition site. Dibasic recognition sites are processed by the kexin family of endoproteases. Some propeptides have a monobasic amino acid cleavage site. These monobasic cleavage sites are recognised in yeast by the yapsin family of endoproteases. The function of the N-terminal propeptide is very divers. It can play a role in folding, sorting, hierarchical organisation, and regulation of activity (Shinde and Inouye, 2000). For subtilisin, the influence of the propeptide on folding has been analysed. Unfolded mature subtilisin can fold into a molten-globule intermediate in the absence of the propeptide (Fig. 4). *In vitro*, this molten globule subtilisin can be converted to the native subtilisin by addition of the propeptide as a separate peptide. This indicates that the folding transition state between the molten globule state and the native state is the rate-limiting factor in the folding process. Furthermore it appears that the propeptide decreases the amount of energy necessary for this transition state.

In summary, three examples of protein processing in yeast were discussed, being killer toxin, α/**a** pheromones, and carboxypeptidase Y. The proteases involved in this processing are the CAAX proteases, the kexin protease, the yapsin proteases, proline specific exoproteases and exoproteases specific for basic residues. In the next paragraph these proteases will be discussed in more detail.

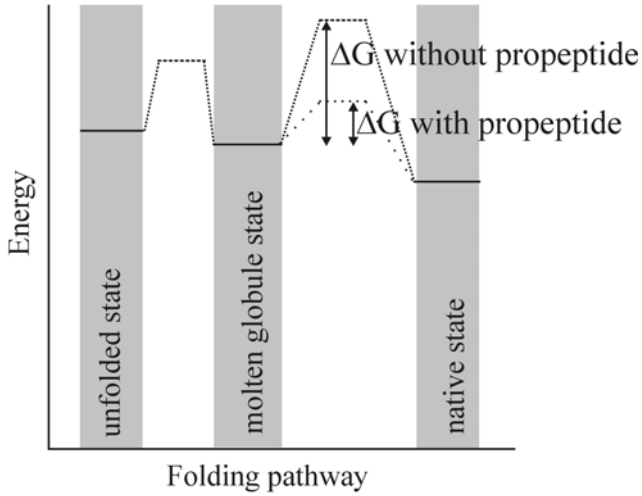


Fig. 4. The partial free energy diagram for folding of subtilisin. The energy needed for the transition state between the molten globule state and native state is reduced in the presence of the propeptide.

CAAX specific proteases

The c-terminal CAAX motif is a motif consisting of a cysteine (C), two aliphatic residues (A), and one less conserved amino acid (X). Eukaryotic proteins containing this motif will be subjected to CAAX processing (Tam et al., 2001), resulting in prenylation of the protein. In contrast to most other forms of protein maturation, this type of post- translational modification takes place at the cytosolic side of the endoplasmic reticulum. Examples of protein families containing this motif are the Ras family of small GTPases, nuclear lamina proteins, and kinetechore proteins (Maurer-Stroh et al., 2003). CAAX processing is a combination of three ordered post-translational processing events (Fig. 5). These three modifications are: protein prenylation, proteolytic removal of the AAX tripeptide, and carboxyl methylesterification.

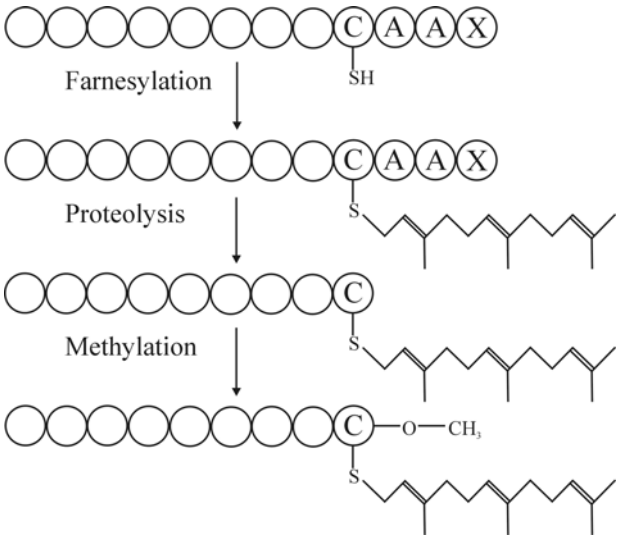


Fig. 5. The sequence of CAAX processing. A protein undergoing CAAX processing is subjected to protein prenylation, proteolytic removal of the AAX tripeptide and carboxyl methylestrification. The CAAX residues are cysteine (C), aliphatic residues (A) and a less conserved residue (X).

The first modification process, protein prenylation, is the attachment of an isoprenoid lipid (either farnesyl or geranylgeranyl) to the cysteine (Fig. 5). Whether the farnesyl or the geranylgeranyl moiety is attached depends on the exact sequence of the CAAX motif (Trueblood et al., 2000). The attachment occurs via a thioether linkage. After the lipid attachment the AAX tripeptide is removed by a CAAX specific protease. As a result the cysteine of the former CAAX motif is now the C-terminal residue. In the final step, this prenylated cysteine residue (N-acetyl-S-farnesyl-L-cysteine or N-acetyl-S-geranylgeranyl-L-cysteine) is methylated (Zhang and Casey, 1996). The prenylation is necessary for membrane association of the protein.

Fungal CAAX specific proteases

CAAX specific proteolysis has been well studied in yeast. Two proteases, identified in *S. cerevisiae,* were shown to be capable of this kind of processing, Afc1 (also called STE24) and Rce1. Both proteases are integral membrane proteins. Afc1 is a type I CAAX protease (Pei and Grishin, 2001). Type I CAAX proteases are zinc dependent proteases with a

HEXXH sequence motif. Rce1 is a type II CAAX protease. The exact catalytic mechanism of type II CAAX proteases is unknown, but they are thought to be metallo-dependent enzymes.

A well characterised fungal substrate for CAAX processing is **a**-factor (Schmidt et al., 1998). This yeast pheromone is exported via a non-classical secretory pathway. The c-terminal CVIA sequence of **a**-factor is a valid CAAX motif triggering the CAAX processing. In the first CAAX processing step a farnesyl isoprenoid is transferred to the cysteine of the **a**-factor CAAX motif, by the yeast cytosolic Ram1/Ram2 farnesyltransferase complex. Both Rce1 and Afc1 are capable of removal of the c-terminal VIA tripeptide, however Afc1 also removes 7 amino acids at the N-terminus of **a**-factor. The CAAX processing is completed by Ste14, the yeast methyltransferase. As both CAAX proteases, Ste14 is located at the ER membrane.

Dibasic residue specific proteases (kexins)

An important group of processing proteases is the kexin family of proteases. The first identified member of this family was the yeast Kex2 endoprotease. In a search for enzymes involved in sexual reproduction of this organism, Kex2 was identified as maturation protease of the yeast **α**-factor, one of the two oligopeptide hormones triggering sexual conjugation of the haploid **α** and **a** cell types (Wagner et al., 1987).

The kexin family of proteases recognizes dibasic amino acid motifs and removes N-terminal propeptides. Members of this family have been identified among yeast, fungi and mammals. An overview of all known members, their substrate specificity and subcellular location is given in Table 1. In mammals the kexin endoproteases are responsible for correct processing of a wide range of proteins such as serum proteins, hormones, growth factors, cell-surface receptors, extracellular matrix proteins, bacterial toxins and viral coat proteins (Molloy et al., 1999).

Table 1. The kexin family of maturases.

	Protease	Cleavage site specificity	Cellular location	Reference
mammalian kexins	Furin (PACE)	-Arg-X-Lys/Arg-Arg↓ I) P_1 arginine is essential, lysine is cleaved less efficient II) At least 2 residues of P_2, P_4, and P_6 have to be basic III) No P_1' amino acid with hydrophobic side chain	TGN/endosomal	Rockwell et al., 2002
	PC1 / PC3	-Lys/Arg-Arg↓	secretory granules	Seidah and Chrétien, 1997
	PC2	-Lys/Arg-Arg↓	secretory granules	
	PACE4	-Arg-X-Lys/Arg-Arg↓	TGN/endosomal	Rockwell et al., 2002
	PC4	-Lys-X-X-Arg↓	Testicular germ cells and ovarian macrophages	Basak et al., 2004
	PC5 / PC6	-Arg-X-X-Arg↓X-Lys-	A: secretory granules B: TGN/endosomal	Seidah and Chrétien, 1997
	PC7 (LPC)	-Lys/Arg-Arg↓	TGN/endosomal	Munzer et al., 1997
yeast kexins	Kex2	-Lys/Arg-Arg↓ I) Extremely stringent for P_1 arginine II) Proline is tolerated at P_2, but cleaved less efficient III) Basic or aliphatic P_4 residue needed	TGN/PVC	Rockwell et al., 2002
	Krp1	-Lys/Arg-Arg↓	N.D.	Davey et al., 1994
	Xpr6	-Lys/Arg-Arg↓	N.D.	Enderlin and Ogrydziak, 1994
k.f.f.	KexB	X-Arg↓ where X can be Arg, Lys, Ala, Pro and Ser	N.D.	Chapter 4

N.D = Not determined; TGN = Trans Golgi Network; PVC = prevacuolar compartment; k.f.f. = kexins of filamentous fungi

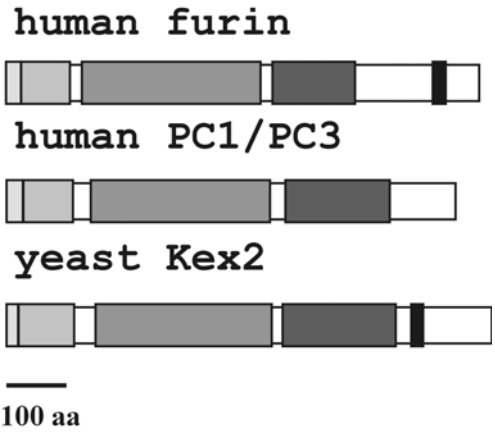


Fig. 6. Three model kexin endoproteases. The *Saccharomyces cerevisiae* Kex2 protease functions as model for fungal kexin-like endoproteases. The human furin protease is a model for Golgi located kexin-like mammalian endoproteases. The human prohormone convertase PC1 (also called PC3) is a model for secretory granules located kexin-like mammalian endoproteases. The grey boxes indicate from left to right the prepeptide, the propeptide, the subtilisin domain, the P-domain, and the transmembrane domain (black).

The kexin proteins contain secretion signals, a subtilisin domain as active domain, an additional P-domain, and depending on their location a transmembrane domain (Fig. 6). As in other proteins, the signal peptide of a kexin-like protease, ensures correct entering of the protease into the endoplasmic reticulum. The kexin prodomain plays an important role in correct folding and activation of the protease. The prodomain is removed in an autocatalytic event, before the protease is secreted. This process has been studied in great detail for furin (Anderson et al., 1997). Furin separates its 83-amino acid N-terminal propeptide in an autocatalytic fashion, soon after it has entered the endoplasmic reticulum (Fig. 7). This first cleavage is a prerequisite for transport of furin from the endoplasmic reticulum into the Golgi. After cleavage the propeptide remains associated with furin, as long as the complex remains in the environment of the endoplasmic reticulum, that has a neutral pH. The associated propeptide functions as an autoinhibitor of furin. In the trans-Golgi network the mildly acid pH of 6.0 invokes a second cleavage event within the propeptide. The second cleavage occurs at -Arg^{70}-Gly-Val-Thr-Lys-Arg^{75}↓. Bhattacharjya et al. (2001) demonstrated that the C-terminal part of the propeptide, containing the second cleavage site, has a molten-globule like conformation at neutral pH,

and is unfolded at acidic pH. They therefore proposed that the second cleavage site is buried within the folded propeptide core at the neutral pH of the endoplasmic reticulum. Furthermore they reason that the acidic pH of the TGN results in a destabilizing of the ionic bondings, existing within the globular state, resulting in a unfolding of the propeptide. As a result, the second cleavage site of the now unfolded propeptide becomes exposed and is processed by furin. After this second cleavage event the fragments dissociate from furin, after which the enzyme is capable of *in trans* cleavage of appropriate substrates.

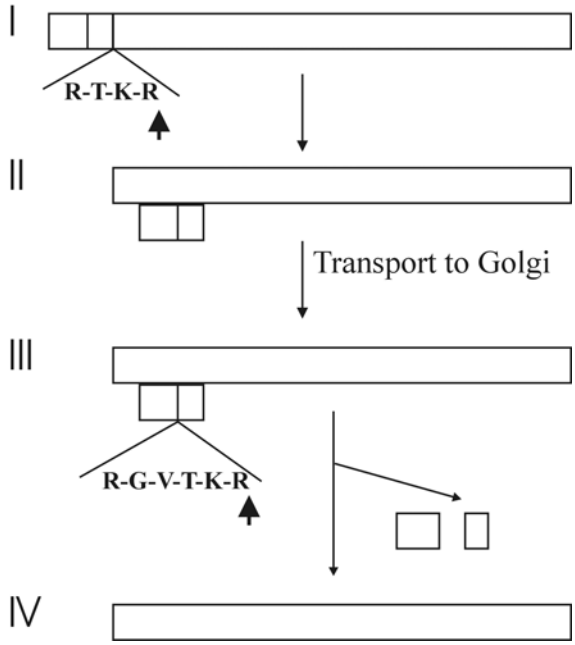


Fig. 7. The autocatalytic processing of the furin propeptide. The furin precursor contains an N-terminal propeptide, with a dibasic cleavage site (I). Following the autocatalytic cleavage of the indicated processing site, the propeptide remains attached to the furin polypeptide (II). This complex travels to the Golgi complex and contains an additional internal cleavage site in the propeptide (III). In the Golgi, the additional cleavage site is processed, resulting in release of the propeptide fragments and mature furin (IV). The bold arrows indicate the cleavage sites.

The active domain of the kexin maturases is a subtilisin – like active domain. The kexin proteases are therefore classified as one of the subclasses of the subtilase superfamily of serine proteases (Siezen and Leunissen,

1997). However in contrast to most members of this superfamily, which are degradative proteases, the kexin proteases have a narrow substrate specificity. Analysis of the crystal structure of soluble Kex2 (Holyoak et al., 2003) and soluble furin (Henrich et al., 2003), both bound to an inhibitor, provided insight into the structural differences between the kexins and the degradative subtilases. Without Ca^{2+} the kexin proteases are not active, this in contrast to other subtilases that do not need a metal ion to be functional. Two of the Ca^{2+} sites in the subtilisin-like domain present in Kex2 are conserved, while the third site, present in both Kex2 and furin, is not found in the degradative subtilases. Furthermore one of the sides of this Ca^{2+} binding pocket is formed by an insertion in the amino acid sequence, which is characteristic for the kexin proteases. The site containing the extra Ca^{2+} ion appears to be essential for the recognition of arginine at P_1. Furthermore, the catalytic domain of the kexins has in comparison with bacterial subtilisins a larger number of negatively charged residue (Glu and Asp).

The P-domain, which is also called the homoB domain, plays an important role in the stabilization of the protease. The P-domain is essential for recognition of pH, calcium binding and substrate specificity. The protein domain is an eight- (furin) or nine (Kex2) -stranded β-barrel structure, which is connected to the subtilisin like domain via an extended tether (Henrich et al., 2003; Holyoak et al., 2003).

Both furin and the fungal kexin proteases contain a transmembrane domain (Rockwell et al., 2002). This domain is absent in the majority of the prohormone convertases. The prohormone convertases lacking a transmembrane domain are expressed by neuroendocrine cells and are localised to the regulated secretory pathway. Furin and the fungal kexin proteases are localised to late compartments of the constitutive secretory pathway.

The kexin proteases which contain a transmembrane domain also contain Golgi retention signals (Voos and Stevens, 1998). Yeast membrane proteins are transported by default to the vacuolar compartment, unless they contain such retention signals. Currently the requirements for a functional Golgi retention signal are unknown. However, it has been established that aromatic residues present in the cytosolic C-terminal part of the Kex2 protein are critical for the retrieval of the endoprotease from the prevacuolar compartment.

Fungal kexin proteases

The kexin endoprotease Kex2 was first identified in *S. cerevisiae*. Homologs of this protease have since then been found in yeast as *Candida albicans* and filamentous fungi such a *Aspergillus niger*. Analysis of this endoprotease in fungi has revealed that the protease is involved in the processing of a wide variety of proteins, is important for the hyphal formation, and is necessary for the pathogenicity of pathogenic fungi.

The completion of the genome of *C. albicans* enabled the identification of possible Kex2 substrates encoded by the genome of this yeast (Newport et al., 2003). In total 147 possible substrates were proposed containing a putative leader peptide followed by a Lys-Arg dibasic amino acid motif. In addition to this *in silico* derived processing, *in vivo* processing by the *S. cerevisiae* Kex2 endoprotease was demonstrated for exoglucanase (Basco et al., 1996), α-factor (Davey et al., 1998), killer toxin (Schmitt and Breinig, 2002), and *Rhizopus oryzae* lipase produced in *S. cerevisae* (Takahashi et al., 1999).

Deletion of the fungal kexin endoprotease appeared to be lethal for *Schizosaccharomyces pombe* (Davey et al., 1994). In several other fungi in which deletion of the protease was attempted it resulted in major morphological changes. Both *Yarrowia lipolytica* and *C. albicans* kexin mutants completely fail to form hyphae, and lack cell separation after budding. Disruption of the *A. niger* (chapter 2) and *A. oryzae* (Mizutani et al., 2004) kexin endoproteases on the other hand results in hyper-branching. The wild type morphology could be restored in *A. oryzae* under growth conditions of high osmolarity (Mizutani et al., 2004). Expression of several cell wall biogenesis related genes appeared to be increased in the *A. oryzae kexB* disrupted strain in comparison with the wild type strain. The *A. oryzae mpkA* gene also appeared to be upregulated in the disrupted strain and encodes a mitogen-activated protein kinase. MpkA plays a role in the cell integrity pathway and is involved in the regulation of cell wall biogenesis related genes. How the *kexB* disruption affects the regulation of these genes is not known.

Several pathogenic fungi in which the kexin endoprotease is disrupted lose their pathogenicity (Newport et al., 2003). As was demonstrated for *C. albicans*, this is caused by the inability to form hyphae and to correctly process secreted aspartic proteases. Furthermore the adhesin of the pathogenic fungus *Coccidioides immites* was also shown to contain a kexin cleavable prosequence (Hung et al., 2002), which demonstrates that

multiple proteins involved in pathogenesis of fungi are targets for kexin processing.

Basic residue specific proteases (yapsins)

Overexpression of α-factor in yeast mutants with a disrupted *kex2* gene resulted in a partially processed product (Egel-Mitani et al., 1990). An aspartic protease, designated yapsin 1, was found to be responsible for this activity. Yapsin-like proteases have specificity for basic amino acid residues. The yapsin protease family currently consist of *S. cerevisiae* yapsin 1 (also known as YAP3), yapsin 2 (also known as MKC7) and yapsin 3, of *C. albicans* Sap9p and of yapsin A (formerly called pro-opiomelanocortin converting enzyme) from bovine pituitary secretory granules (Cawley et al., 2003).

The posttranslational processing of Yapsin 1 has been well studied. The protease is synthesized as a protein with an N-terminal prepropeptide. After removal of the prepeptide, the propeptide is removed and the protease is processed into an α and β subunit. The two subunits are associated with each other via a disulfide bond. The proteases of the yapsin family are anchored to the transmembrane domain with a glycosylphosphatidylinositol (GPI) anchor. This GPI anchor is added to the protein backbone in the ER.

Fungal yapsin proteases

In the genome of S. cerevisiae five potential yapsin proteases are encoded (Olsen et al., 1999). Of these five, three (yapsin1, yapsin2 and yapsin3) have been characterised and shown to be involved in proteolytic posttranslational processing. In *S. pombe*, one yapsin, Yps1 has been identified and characterised (Ladds and Davey, 2000). Overexpression of the *S. pombe* yapsin resulted in partial complementation of temperature sensitive mutants in which the *S. pombe* kexin encoding gene *krp1* has been disrupted.

Only the combined disruption of *kex2*, yapsin1 and yapsin2 in *S. cerevisiae* resulted in a severe reduction of cell growth, suggesting an overlap in substrate specificity between these three proteases (Komano et al., 1999). Moreover overexpression of either yapsin1 or yapsin2 in *kex2* disrupted mutants could partially restore the wildtype phenotype.

Proline specific proteases

Many proteases have a broad specificity and can cleave almost any amino acid bond, while other proteases can only cleave very specific amino acid motifs. The proline residue is unique due to its cyclic structure (Cunningham and O'Conner, 1997). Most proteases cannot cleave the proline imino acid bond. Therefore, proline residues provide the polypeptide with a means of protection against proteolysis. As a result, a specific set of proline specific proteases has evolved that is able to cleave these chemical bonds. A wide range of proline specific proteases exist in nature. A summary of these different types of proteases and their substrate specificity is given in Table 2. The combination of resistance of proline towards unspecific proteolysis and the limited proteolytic activity of these proline specific proteases allowed proline residues to develop a second biological function, regulation of activity.

The eukaryotic cell produces a wide spectrum of biological active enzymes. Obviously, most of these enzymes do not need to be active in the secretory pathway. To protect them from proteolysis and to protect the cell from their activity, many of those enzymes are produced as inactive precursors. The N-terminal, C-terminal, or internal peptides that keep the precursor inactive need to be removed by an endoprotease. The resulting enzyme often contains additional amino acids at the new protein terminus that need to be removed. In many cases these additional amino acids are either alanine or proline residues, which are removed by proline specific proteases.

Table 2. Proline specific proteases.

Protease	EC number	Substrate specificity	Identified in fungal species*
prolyl aminopeptidase	3.4.11.5	releases N-terminal proline from a peptide	*Aspergillus niger* (Q96VT2)
X-Pro aminopeptidase	3.4.11.9	releases proline linked N-terminal amino acids, including proline, even from a di- or tripeptide	
X-Pro dipeptidase	3.4.13.9	Hydrolyses X↓Pro dipeptides, except Pro-Pro; also acts on aminoacyl-hydroxyproline analogs	*Aspergillus nidulans* (Q96WX8)
cytosol non-specific dipeptidase	3.4.13.18	Hydrolyses dipeptides, preferentially hydrophobic dipeptides including prolyl amino acids, including Pro↓X.	
dipeptidyl-peptidase II	3.4.14.2	Releases N-terminal dipeptides, X-Z↓, preferentially when Z is Ala or Pro. Substrates are oligopeptides, preferentially tripeptides.	
dipeptidyl-peptidase IV	3.4.14.5	Releases N-terminal dipeptides, U-X↓Z-, from a polypeptide, when X is Pro (preferred), Ala or hydroxyproline, provided Z is neither Pro or hydroxyproline	*Aspergillus fumigatus* (O14425) *Aspergillus niger* (Q96VT7) *Saccharomyces cerevisiae* (P18962) *Saccharomyces cerevisiae* (P33894)
X-Pro dipeptidyl-peptidase	3.4.14.11	Hydrolyses X-Pro↓ bonds to release unblocked, N-terminal dipeptides from substrates including Ala-Pro↓*p*-nitroanilide and (sequentially) Tyr-Pro↓Phe-Pro↓Gly-Pro↓Ile	
lysosomal Pro-X carboxypeptidase	3.4.16.2	Hydrolyses -Pro↓X to release a C-terminal amino acid	
membrane Pro-X carboxypeptidase	3.4.17.16	Releases C-terminal residues other than proline, by preferentially cleavage of a prolyl bond	

Protease	EC number	Substrate specificity	Identified in fungal species*
prolyl oligopeptidase	3.4.21.26	Hydrolysis of -Pro↓ in oligopeptides, -Ala↓ is cleaved at lower efficiency	
microbial collagenase	3.4.24.3	Digests native collagen in the triple helical region at ↓Gly bonds. With synthetic peptides it has a preference for Gly at P_3 and P_1', Pro and Ala at P_2 and P_2', and hydroxyproline, Ala or Arg at P_3'	
IgA-specific metalloendopeptidase	3.4.24.13	Hydrolyses -Pro↓Thr- in the hinge region of the heavy chain of human IgA	
Procollagen N-endopeptidase	3.4.24.14	Cleaves the N-propeptide of collagen chain α1(I) at Pro↓Gln and of α1(II) and α2(I) at Ala↓Gln	
Neurolysin	3.4.24.16	Preferential cleavage in neurotensin: Pro^{10}↓Tyr	
Astacin	3.4.24.21	hydrolyses substrates with five or more amino acids, preferentially with Ala in P_1', and Pro in P_2'	
Saccharolysin	3.4.24.37	Cleavage of -Pro↓Phe- and -Ala↓Ala- bonds	*Saccharomyces cerevisiae* (P25375)

*Only proteases of which a (deduced) protein sequence is known and the proteolytic activity has been determined are named here

Fungal proline specific proteases

Four different types of the proline specific protease summarised in Table 2 have so far been characterised in fungi. Dipeptidyl peptidases of type IV have been characterised in filamentous fungi and yeasts. Prolidase (X-Pro dipeptidase) and prolyl aminopeptidase have been characterised in filamentous fungi, and saccharolysin (protease D) has been characterised in yeast.

Dipeptidyl aminopeptidase A of yeast was the first identified dipeptidyl peptidase in fungi. Disruption of this protease resulted in sterile mutants, which were not able to produce functional α-factor. Dipeptidyl aminopeptidase A (DPAP A) is located in the Golgi and the exoprotease has been shown to be involved the processing of α-factor, killer toxin, and Kex2. Dipeptidyl aminopeptidase B is a vacuolar homologue of DPAP A, present in *S. cerevisiae*. Its biological function has not been established but it is thought to be important for the recycling of proline.

In filamentous fungi both an intracellular and extracellular DPP IV have been characterised. The intracellular DPP IV is probably a membrane located protease (chapter 4) and might play a role in DPAP A-like maturation in filamentous fungi. The extracellular DPP IV is probably involved in proline catabolism (Beauvais et al., 1997).

The dipeptidyl peptidases of type IV produce X-Pro dipeptides as product of their action. These dipeptides can be hydrolysed by prolidase. Prolidase has been well characterised in the filamentous fungus *Aspergillus nidulans* (chapter 5). It is thought to be located intracellularly.

The prolyl aminopeptidase has been characterised in both yeast and filamentous fungi. It removes N-terminal prolyl residues from oligopeptides, and is located intracellular. It is not known if this protease performs a role in proteolytic processing. In a similar fashion saccharolysin is located intracellular in yeast and is thought to play a role in the late stage of protein degradation.

Other processing proteases

S. cerevisiae Kex1p is a membrane associated carboxypeptidase and removes carboxyterminal basic residues from its substrate (Cooper and Bussey, 1989). Both α-factor and killer toxin are substrates for Kex1p, which removes the arginine and lysine residues remaining after cleavage of both substrates by Kex2. This protease has not been characterised very well.

In mammals an additional processing protease of the subtilase protease family was recently identified (Elagoz et al., 2002). The protease named subtilisin kexin isozyme-1 (SKI-1) preferentially cleaves the following sequence: -(R/K)-X(hydrophibic)-Z↓, where Z is preferably Leu or Thr, but certainly not Val, Pro, Glu, Asp or Cys. Whether any fungal homologues of this protease exist is currently not known.

Outline and aim of this thesis

The proteases described in this introduction play an important role in proper folding, regulation and activation of enzymes produced by filamentous fungi. Filamentous fungi produce and secrete a broad spectrum of enzymes, and often in large quantities. As a result, fungi are capable of degrading many different substrates and are able to grow in many different environments. Fungi are also interesting organisms for many industrial applications because of their enzyme producing capabilities. Proteases involved in proteolytic processing are therefore a major topic of interest, but nevertheless little is known about their existence and role in filamentous fungi. Therefore, within the framework of STW, the characterisation of these proteases in the secretory pathway of the industrial used filamentous fungus *Aspergillus niger* was initiated. The research was conducted at the former section Molecular Genetics of Industrial Microorganisms. Upon joining the Microbiology department the group was recently renamed Fungal Genomics.

The aim of this thesis is the identification and characterisation of proteases, involved in proteolytic processing in the secretory pathway of *A. niger*. In chapters 2 and 3, KexB is described. KexB is the major propeptide processing protease of *A. niger*. As demonstrated in chapter 2, the protease recognises dibasic amino acid motifs and removes the propeptide from the newly synthesized proteins. The focus of chapter 3 is on proteins containing only one basic amino acid as cleavage site for propeptide removal. As shown in chapter 3, KexB is also capable of processing these substrates. Chapters 4 and 5 are focused on the characterisation of two proline specific proteases. In chapter 4, the cloning and characterisation of DapB is described. DapB is a dipeptidyl aminopeptidase and removes dipeptides from the N-terminus of a protein if this protein contains a penultimate alanine or proline residue at the N-terminus. Prolidase, a proline specific dipeptidase is described in chapter 5. Purified prolidase has been used to analyse the substrate specificity of this protease. In the chapters 6 and 7 the results presented in this thesis will be discussed and summarised.

References

Alkorta, I., Garbisu, C., Llama, M. J., and Serra, J. L. (1998). Industrial applications of pectic enzymes: a review. *Process Biochem.* **33: 21-28.**

Anderson, E. D., VanSlyke, J. K., Thulin, C. D., Jean, F., and Thomas, G. (1997). Activation of the furin endoprotease is a multiple-step process: requirements for acidification and internal propeptide cleavage. *EMBO J.* **16: 1508-1518.**

Archer, D. B. (2000). Filamentous fungi as microbial cell factories for food use. *Curr. Opin. Biotechnol.* **11: 478-483.**

Basak, S., Chrétien, M., Mbikay, M., and Basak, A. (2004). *In vitro* elucidation of substrate specificity and bioassay of proprotein convertase 4 using intramolecularly quenched fluorogenic peptides. *Biochem. J.* **380: 505-514.**

Basco, R. D., Cueva, R., Andaluz E., and Larriba G. (1996). In vivo processing of the precursor of the major exoglucanase by KEX2 endoprotease in the *Saccharomyces cerevisiae* secretory pathway. *Biochim. Biophys. Acta*. **1310: 110-118.**

Beauvais, A., Monod, M., Wyniger, J., Debeaupuis, J. P., Grouzmann, E., Brakch, N., Svab, J., Hovanessian, A. G., and Latgé, J. P. (1997). Dipeptidyl-peptidase IV secreted by *Aspergillus fumigatus*, a fungus pathogenic to humans. *Infect. Immun.* **65: 3042-3047.**

Bhattacharjya, S., Xu, P., Xiang, H., Chrétien, M., Seidah, N. G., and Ni, F. (2001). pH-induced conformational transitions of a molten-globule-like state of the inhibitory prodomain of furin: Implications for zymogen activation. *Protein Sci.* **10: 934-942.**

Borkovich, K. A., Alex, L. A., Yarden, O., Freitag, M., Turner, G. E., Read, N. D., Seiler, S., Bell-Pedersen, D., Paietta, J., Plesofsky, N., Plamann, M., Goodrich-Tanrikulu, M., Schulte, U., Mannhaupt, G., Nargang, F. E., Radford, A., Selitrennikoff, C., Galagan, J. E., Dunlap, J. C., Loros, J. J., Catcheside, D., Inoue, H., Aramayo, R., Polymenis, M., Selker, E. U., Sachs, M. S., Marzluf, G. A., Paulsen, I., Davis, R., Ebbole, D. J., Zelter, A., Kalkman E. R., O'Rourke, R., Bowring, F., Yeadon, J., Ishii, C., Suzuki, K., Sakai, W., and Pratt, R. (2004). Lessons from the genome sequence of *Neurospora crassa*: Tracing the path from genomic blueprint to multicellular organism. *Microbiol. Mol. Biol. Rev.* **68: 1-108.**

Cawley, N. X., Chino, M., Maldonado, A., Rodriguez, Y. M., Loh, Y. P., and Ellman, J. A. (2003). Synthesis and characterization of the first potent inhibitor of yapsin 1. *J. Biol. Chem.* **278: 5523-5530.**

Cereghino, J. L., and Cregg, J. M. (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* **24: 45-66.**

Cooper, A., and Bussey, H. (1989). Characterization of the yeast *KEX1* gene product: a carboxypeptidase involved in processing secreted precursor proteins. *Mol. Cell. Biol.* **9: 2706-2714.**

Cunningham, D. F., and O'Connor, B. (1997). Proline specific peptidases. *Biochim. Biophys. Acta*. **1343: 160-186.**

Davey, J., Davis, K., Imai, Y., Yamamoto, M., and Matthews, G. (1994). Isolation and characterization of krp, a dibasic endopeptidase required for cell viability in the fission yeast *Schizosaccharomyces pombe EMBO J*. **13: 5910-5921.**

Davey, J., Davis, K., Hughes, M. Ladds, G., and Powner, D. (1998). The processing of yeast pheromones. *Semin. Cell Dev. Biol.* **9: 19-30.**

Egel-Mitani, M., Flygenring, H. P., and Hansen, M. T. (1990). A novel aspartyl protease allowing KEX2-independent MF alpha propheromone processing in yeast. *Yeast* **6: 127-137.**

Elagoz, A., Benjannet, S., Mammarbassi, A., Wickham, L., and Seidah, N. G. (2002). Biosynthesis and cellular trafficking of the convertase SKI-1/S1P. *J. Biol. Chem.* **277: 11265-11275.**

Enderlin, C. S., and Ogrydziak, D. M. (1994). Cloning, nucleotide sequence and functions of *xpr6*, which codes for a dibasic processing endoprotease from the yeast *Yarrowia lipolytica*. *Yeast* **10: 67-79.**

Gouka, R. J., Punt, P. J., and van den Hondel, C. A. M. J. J. (1997). Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Appl. Microbiol. Biotechnol.* **47: 1-11.**

Henrich, S., Cameron A., Bourenkov, G. P., Kiefersauer, R., Huber, R., Lindberg, I., Bode, W., and Than, M. E. (2003). The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. *Nat. Struct. Biol.* **10: 520-526.**

Holyoak, T., Wilson, M. A., Fenn, T. D., Kettner, C. A., Petsko, G. A., Fuller, R. S., and Ringe D. (2003). 2.4 Å Resolution crystal structure of the prototypical hormone-processing protease Kex2 in complex with an Ala-Lys-Arg Boronic Acid Inhibitor. *Biochemistry* **42: 6709-6718.**

van den Hombergh, J. P. T. W. (1996). An analysis of the proteolytic system in *Aspergillus* in order to improve protein production. *Thesis.* Wageningen University, Wageningen, The Netherlands.

van den Hombergh, J. P. T. W., van de Vondervoort, P. J. I., Fraissinet-Tachet, L., and Visser, J. (1997). *Aspergillus* as a host for heterologous

protein production: the problem of proteases. *Trends biotechnol*. **15: 256-263.**

Hooper, N. M. (1994). Families of zinc metalloproteases. *FEBS Lett*. **354: 1-6.**

Hung, C.-Y., Yu, J.-J., Seshan, K. R., Reichard, U., and Cole, G. T. (2002). A parasitic phase-specific adhesin of *Coccidioides immitis* contributes to the virulence of this respiratory fungal pathogen**.** *Infect. Immun.* **70: 3443-3456**.

Jiang, W., and Bond, J. S. (1992). Families of metalloendoproteases and their relationships. *FEBS Lett.* **312: 110-114**.

Komano, H., Rockwell, N., Wang, G. T., Krafft, G. A., and Fuller, R. S. (1999). Purification and characterization of the yeast glycosylphosphatidylinositol-anchored monobasic-specific aspartyl protease yapsin 2 (Mkc7p). *J. Biol. Chem.* **274: 24431-24437.**

Ladds, G., and Davey, J. (2000). Identification of proteases with shared functions to the proprotein processing protease Krp1 in the fission yeast *Schizosaccharomyces pombe*. *Mol. Microbiol.* **38: 839-853.**

Maras, M., van Die, I. Contreras, R., and van den Hondel, C.A.M.J.J. (1999). Filamentous fungi as production organisms for glycoproteins of bio-medical interest. *Glycoconj. J*. **16: 99-107.**

Maurer-Stroh, S., Washietl, S., and Eisenhaber, F. (2003). Protein prenyltransferases. *Genome Biol.* **4: 212.**

Mizutani, O., Nojima, A., Yamamoto, M., Furukawa, K., Fujioka, T., Yamagata, Y., Abe, K., and Nakajima, T. (2004). Disordered Cell Integrity Signaling Caused by Disruption of the *kexB* Gene in *Aspergillus oryzae*. *Eukaryotic Cell* **3: 1036-1048.**

Molloy, S. S., Anderson, E. D., Jean, F. and Thomas, G. (1999). Bi-cycling the furin pathway: from TGN localization to pathogen activation and embryogenesis. *Trends Cell Biol*. **9: 28-35.**

Munzer, J. S., Basak, A., Zhong, M., Mamarbachi, A., Hamelin, J., Savaria, D., Lazure, C., Benjannet, S., Chrétien, M., and Seidah, N. G. (1997). *In vitro* characterization of the novel proprotein convertase PC7. *J. Biol. Chem.* **272: 19672-19681.**

Nakano, A. (2004). Yeast Golgi apparatus – dynamics and sorting. *Cell. Mol. Life Sci.* **61: 186-191.**

Newport, G., Kuo, A., Flattery, A., Gill, C. , Blake, J. J., Kurtz, M. B., Abruzzo, G. K., and Agabian, N. (2003). Inactivation of kex2p diminishes the virulence of *Candida albicans*. *J. Biol. Chem.* **278: 1713-1720.**

Northrop, D. B. (2001). Follow the protons: A low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Acc. Chem. Res.* **34: 790-797.**

Olsen, V., Cawley, N. X., Brandt, J.. Egel-Mitani, M., and Loh, Y. P. (1999). Identification and characterization of *Saccharomyces cerevisiae* yapsin 3, a new member of the yapsin family of aspartic proteases encoded by the *YPS3* gene. *Biochem. J.* **339: 407-411.**

Pei, J., and Grishin, N. V. (2001). Type II CAAX prenyl endopeptidases belong to a novel superfamily of putative membrane-bound metalloproteases. *Trends Biochem. Sci.* **26: 275-277.**

Rockwell, N. C., Krysan, D. J., Komiyama, T., and Fuller, R. S. (2002). Precursor processing by Kex2/Furin proteases. *Chem. Rev.* **102: 4525-4548.**

Schmidt, W. K., Tam, A., Fujimura-Kamada, K., and Michaelis, S. (1998). Endoplasmic reticulum membrane localization of Rce1p and Ste24p, yeast proteases involved in carboxyl-terminal CAAX protein processing and amino-terminal a-factor cleavage. *Proc. Natl. Acad. Sci.* **95: 11175-11180. Schmitt, M. J., and Breinig, F. (2002).** The viral killer system in yeast: from molecular biology to application. *FEMS Microbiol. Rev.* **26: 257-276.**

Seidah, N. G., and Chrétien, M. (1997). Eukaryotic protein processing: endoproteolysis of precursor proteins. *Curr. Opin. Biotechnol.* **8: 602-607.**

Shinde, U., and Inouye, M. (2000). Intramolecular chaperones: polypeptide extensions that modulate protein folding. *Semin. Cell Dev. Biol.* **11: 35-44.**

Siezen, R. J., and Leunissen, J. A. M. (1997). Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Sci.* **6: 501-523.**

Takahashi, S., Ueda, M., and Tanaka A. (1999). Independent production of two molecular forms of a recombinant *Rhizopus oryzae* lipase by *KEX2*-engineered strains of *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **52: 534-540.**

Tam, A., Schmidt, W. K., and Michaelis, S. (2001). The multispanning membrane protein Ste24p catalyzes CAAX proteolysis and NH_2-terminal processing of the yeast a-factor precursor. *J. Biol. Chem.* **276: 46798-46806.**

Trueblood, C. E., Boyartchuk, V. L., Picologlou, E. A., Rozema, D., Poulter, D., and Rine J. (2000). The CaaX proteases Afc1p and Rce1p, have overlapping but distinct substrate specificities. *Mol. Cell. Biol.* **20: 4381-4392.**

Verdoes, J. C., Punt, P. J., and van den Hondel. C. A. M. J. J. (1995). Molecular genetic strain improvement for the overproduction of fungal proteins by filamentous fungi. *Appl. Microbiol. Biotechnol.* **43: 195-205.**

Voos, W., and Stevens, T. H. (1998). Retrieval of resident late-Golgi membrane proteins from the prevacuolar compartment of *Saccharomyces cerevisiae* is dependent on the function of Grd19p. *J. Cell. Biol.* **140: 577-590.**

Wagner, J. -C., Escher C., and Wolf, D. H. (1987). Some characteristics of hormone (pheromone) processing enzymes in yeast. *FEBS Lett.* **218: 31-34.**

Zhang, F. L., and Casey, P. J. (1996). Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* **65: 241-269.**

Chapter 2

Characterization of the kexin-like maturase of *Aspergillus niger*

Jalving, R., P. J. I. van de Vondervoort, J. Visser, and P. J. Schaap
(2000)
Applied and Environmental Microbiology **66**:363-368

Characterization of the kexin-like maturase of *Aspergillus niger*

Ruud Jalving, Peter J. I. van de Vondervoort, Jaap Visser and Peter J. Schaap

Section Molecular Genetics of Industrial Microorganisms, Wageningen University, Wageningen, The Netherlands

Abstract

Secreted yields of foreign proteins may be enhanced in filamentous fungi through the use of translational fusions in which the target protein is fused to an endogenous secreted carrier protein. The fused proteins are usually separated *in vivo* by cleavage of an engineered Kex2 endoprotease recognition site at the fusion junction. We have cloned the kexin-encoding gene of *Aspergillus niger* (*kexB*). We constructed strains that either overexpressed KexB or lacked a functional *kexB* gene. Kexin-specific activity doubled in membrane-protein fractions of the strain overexpressing KexB. In contrast, no kexin-specific activity was detected in the similar protein fractions of the *kexB* disruptant. Expression in this loss-of-function strain of a glucoamylase human interleukin-6 fusion protein with an engineered Kex2 dibasic cleavage site at the fusion junction resulted in secretion of unprocessed fusion protein. The results show that KexB is the endoproteolytic proprotein processing enzyme responsible for processing of (engineered) dibasic cleavage sites in target proteins that are transported through the secretion pathway of *A. niger*.

Introduction

Many secreted eukaryotic proteins contain a signal peptide and an adjacent propeptide at the amino terminus. The signal peptide specifies a sequence for translocation over the endoplasmic reticulum membrane and is normally removed in the lumen during translocation by a signal peptidase. Propeptides have been implicated in correct folding and in subcellular sorting of proteases. They also often function as (auto)inhibitors. The processing of most of these propeptides occurs at either a monobasic or a dibasic cleavage site (Baker et al., 1993). Propeptides, which are cleaved

after a Lys-Arg or Arg-Arg basic doublet at the P2 and P1 positions (the nomenclature used is according to Schechter and Berger (1967)), are specifically recognized and processed in the trans-Golgi network by the kexin family of proteases, a subfamily of the subtilase family of proteases (Siezen and Leunissen, 1997). The kexin family consists of the yeast Kex2-like proteases (E.C. 3.4.21.61), the mammalian prohormone convertases (PCs) (E.C. 3.4.21.93 and E.C. 3.4.21.94), and the furins (E.C. 3.4.21.75). All members of the kexin subfamily are calcium-dependent, neutral, serine proteases that are activated by the removal of the amino-terminal propeptide at a kexin-specific (auto)processing site. The active proteases all contain two additional domains, a subtilisin-like domain containing the catalytic triad and a conserved P or Homo B domain of approximately 150 residues. The P domain, which is absent in other subtilases, is essential for the catalytic activity (Nakayama**,** 1997) and the stability of the protein (Lipkind et al., 1998). Kex2-like yeast proteases, furins and some of the PCs also have a single transmembrane domain (Molloy et al., 1992; Nakayama, 1997; Seidah et al., 1996). In its cytoplasmic tail, yeast Kex2 contains a Golgi retrieval signal, necessary to remain in the trans-Golgi network (Wilcox et al., 1992).

In *Aspergillus* spp., kexin-like activity has been detected through the cleavage of artificial fusion proteins. Here artificial fusion proteins are used for the production of foreign proteins, exploiting the efficient production, sorting, and processing of endogenous proteins (reviewed in Gouka et al., 1997). In these constructs, a consensus Kex2 cleavage site often separates the foreign protein from the endogenous protein. Studies with such fusion proteins in *Aspergillus niger* showed that amino acid residues directly adjacent to the cleavage site can affect correct processing of an engineered Kex2 site (Spencer et al., 1998).

Examples of physiological substrates for an *A. niger* kexin-like endoprotease are the endopolygalacturonase (Pga) family of proteins (Bussink et al., 1990; Bussink et al., 1991; Bussink et al., 1992; Pařenicová et al., 1998). They all contain a dibasic cleavage site at the carboxy-terminal end of their amino-terminal propeptide, except for PgaII, which has a single arginine residue preceding the cleavage site (Bussink et al., 1991).

Our objectives in this study were (i) to show that *A. niger* expresses a Kex2-like dibasic endoprotease and (ii) to demonstrate that this endoprotease is responsible for the cleavage of a fusion protein with an engineered Kex-2 site.

Materials and methods

Strains, transformation, and DNA and RNA techniques. The *A. niger* strains used in this study are listed in Table 1. NW219, NW249, and NW266 were used for the transformation of *A. niger*, as previously described (Kusters-van Someren et al., 1991). Strain NW266 was constructed by transformation of NW249 with pIM4003. (The construction of the pIM plasmids is described below.) MCK-5 was constructed by the cotransformation of NW249 with pGW635, containing the *A. niger pyrA* gene, and pIM4002. MCGI and MCGIΔ were constructed by the cotransformation of, respectively, NW219 and NW266 with pGW635 and pFGPDGLAHIL6 (Contreras et al., 1991).

Table 1. *A. niger* strains used in this study[a]

Strain	Genotype	Reference
N402	*cspA1*	Bos et al., 1998
NW219	*pyrA6 leuA1 nicA1 cspA1*	Kusters-van Someren, 1991
NW249	*ΔargB, pyrA6, nicA1, leuA1, cspA1*	van de Vondervoort and Muller, unpublished data
MCK-5	NW249 (*pyrA*) $(kexB)_{15\text{-}25}$	This study
NW266	NW249 (*argB::ΔkexB*)	This study
MCGI	NW219 (*pyrA*) $(PgpdA\text{-}glaA\text{-}KEX2\text{-}hIL6)_{12\text{-}15}$	This study
MCGIΔ	NW266 (*pyrA*) $(PgpdA\text{-}glaA\text{-}KEX2\text{-}hIL6)_{10\text{-}12}$	This study

[a]Genes introduced by transformation are indicated in parentheses. Estimated copy numbers of relevant genes were determined by Southern analysis and are indicated in subscript. *KEX2* symbolizes an engineered dibasic processing site.

Escherichia coli LE392 (Promega, Madison, Wis.) was used for phage amplification and λ DNA isolation. *E. coli* DH5α was used for plasmid transformation and propagation. Standard DNA manipulations were carried out essentially as described by Sambrook *et al.* (1989). Plasmid pUC19 or phagemid pBluescript SK^+ was used as a cloning vector for genomic DNA fragments. Cloned hybridizing fragments and cDNA clones were sequenced with a Thermo Sequenase fluorescent-labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham Pharmacia Biotech, Uppsala, Sweden) and an ALF automated sequencer (Amersham Pharmacia Biotech, Uppsala, Sweden).

For Southern analysis, total DNA from *Aspergillus* strains was isolated as previously described (de Graaff et al., 1988). Hybridizations

were done in standard hybridization buffer (SHB) (6 x SSC, 5 x Denhardt's solution (Schechter and Berger, 1967), 0.5% sodium dodecyl sulphate (SDS), and 100 μg of denatured herring sperm DNA ml^{-1}). 1 x SSC contains 0.15 M NaCl and 0.015 M sodium citrate. Washing was performed at 65°C to a final stringency of 0.1 x SSC and 0.1% SDS. For nonstringent conditions, hybridization was executed at 56^0C and washing was done twice in 4 x SSC and 0.1% SDS at the same temperature.

For Northern analysis, strains were grown for 17 h in minimal medium (Pontecorvo et al., 1953) supplemented with 1% glucose as carbon source and 0.5% yeast extract in 50-ml cultures in 250-ml Erlenmeyer flasks in an Innova incubator shaker (New Brunswick Scientific Co., Inc., Edison, N.J.) at 250 rpm at 30^oC. Mycelium was harvested by filtration over a nylon membrane (mesh size, 100 μm) and the mycelium was ground with a Braun II dismembrator (B. Braun Melsungen AG, Melsungen, Germany). Total RNA was isolated from mycelium samples with Trizol reagent (Life Technologies, Rockville, Md.). RNA concentrations were determined spectrophotometrically and equal amounts of RNA were denatured with glyoxal by standard techniques (Sambrook et al., 1989) and separated on a 1.2% (wt/vol) agarose gel. RNA blots were hybridized at 42°C in SHB to which 10% (wt/vol) dextran sulfate and 50% (vol/vol) formamide were added. Washing was performed at 65°C to a final stringency of 0.1 × SSC and 0.1% SDS. As a control, Northern blots were hybridized with *A. niger* ribosomal protein gene *rpS28*.

Cloning of *kexB*. A *kexB* PCR product was generated by PCR on genomic DNA of *A. niger* N400 with two degenerate primers: the forward primer was 5'-CAYGGNACIMGITGYGCNGG-3', encoding HGTRCAGE and the reverse primer was 5'-TAYTGNACRTCNCKCCA-3', encoding WRDVQY (Standard IUB-IUPAC symbols are used to indicate the nucleotide mixes, I indicates inosine). A standard program of 30 thermal cycles was composed of 1 min at 94°C, 1 min at 48°C, and 1 min at 72°C preceded by an incubation of 4 min at 95°C and followed by an incubation of 5 min at 72°C. The amplified fragment was cloned in vector pGEM-T (Promega) and identified by sequencing. Genomic sequences of the *kexB* gene were obtained by screening a λ EMBL4 genomic library of *A. niger* N402 by standard methods (Sambrook et al., 1989), with the *kexB* PCR-product as a probe. Four λ clones were isolated and from one of these positive phages, two *Sal*I fragments of about 2 kb were subcloned in pUC19.

Plasmid construction. The upstream and downstream *kexB Sal*I fragments in pUC19 were each recloned in pBluescript SK(+) digested with *Sal*I and *Xho*I. The resulting plasmids were selected for having the *Sal*I sites

upstream and downstream, respectively, of *kexB* ligated in the *Xho*I sites of the pBluescript vectors. Next, the downstream *kexB* fragment in pBluescript was ligated as a *Sal*I-*Eco*R1 fragment downstream of the upstream *kexB* fragment, yielding pIM4002 (see Fig. 2). For the construction of pIM4003, a 590-bp *Cla*I-*Sal*I fragment was first removed from the open reading frame (ORF) of the upstream *Sal*I fragment in pBluescript. The resulting plasmid was digested with *Hin*dIII and *Pst*I, and the *arg*B gene was inserted 3' of the truncated *kexB* fragment. In the resulting plasmid, the *kexB* downstream fragment was inserted as a *Pst*I-*Bam*HI fragment yielding pIM4003 (see Fig. 3).

Nucleotide and protein sequence analyses. Sequence analysis was performed with the sequence analysis software package PC/Gene (IntelliGenetics, Inc., Geneva, Switzerland). Public databases were searched with the Blast search tools (Altschul et al., 1997). Multiple alignment was done in Clustal W (Thompson et al., 1994). The SignalP program was used to identify the signal sequence for secretion (Nielsen et al., 1997). Putative transmembrane regions were identified with the program SOSUI (Hirokawa et al., 1998).

Isolation of KexB containing membrane-protein fractions. Strains were grown and mycelium was treated as described for the Northern analysis. Ground mycelium (3 g {wet weight}) was extracted with 6 ml of 50 mM sodium HEPES (pH 7.6) and 10 mM EDTA. The extract was clarified by centrifugation (10,000 × *g* for 15 min at 4°C) and the clarified supernatant was centrifuged again at 100,000 × *g* for 90 min at 4°C. The supernatant was discarded and the pelleted membrane-containing fraction was extracted with 5 ml of 50 mM sodium HEPES (pH 7.6), 1 mM EDTA, 50 mM NaCl, and 2% (wt/vol) sodium deoxycholate in 20% glycerol. The extract was clarified again by centrifugation at 100,000 × *g* for 90 min at 4°C. The supernatant was stored at -20°C until use.

Kexin enzyme assay. 7-Amino-4-methylcoumarin (AMC) and tert-butyloxycarbonyl (Boc) methylcoumarinamide (MCA) derivatives were purchased from Sigma Chemical Co. (St. Louis, Mo.). The reaction mixture (100 μl) contained 200 mM sodium HEPES (pH 7.0), 1.5 mM $CaCl_2$, and 100 μM MCA derivative (Table 2). Reaction mixtures were incubated at 37° C for 0 to 4 h and the reactions were terminated by the addition of 1.8 ml of 125 mM $ZnSO_4$ and 0.2 ml of saturated $Ba(OH)_2$. The precipitate was removed by centrifugation for 3 min at 10,000 × *g*, and the amount of AMC liberated was measured with a Hitachi F4500 fluorescence spectrophotometer calibrated with known amounts of AMC (λ_{ex} = 370 nm;

λ_{em} = 445 nm). One unit was defined as 1 pmol of AMC released per min (Fuller et al., 1989). Protein concentrations were determined by the Bicinchoninic acid method as described by the manufacturer (Sigma Chemical Co.).

Western analysis of strains MCGI and MCGIΔ. Shake flask cultures were used to express the *PgpdA-glaA-hIL*-6 fusion gene from strains MCGI and MCGIΔ, each harboring 10 to 15 copies of the construct (Table 1). Culture conditions were as described for Northern analysis except that xylose was used as a carbon source to suppress the endogenous glucoamylase gene. Culture fluid was concentrated by deoxycholate-trichloroacetic acid precipitation. Medium samples were subjected to SDS-polyacrylamide gel electrophoresis and standard protocols were used for detection of recombinant human Interleukin-6 fusion protein by Western analysis (Sambrook et al., 1989). Mouse monoclonal human interleukin-6 antibody (R&D Systems, Inc., Minneapolis, Minn.) was used for the detection. Recombinant human interleukin-6 (Sigma-Aldrich, St. Louis, Mo.) was used as a positive control.

Nucleotide sequence accession number. The sequence data of the *kexB* gene has been submitted to the DDBJ, EMBL, and GenBank databases under accession no. Y18127.

Results

Cloning of an *A. niger* kexin homologue. We designed degenerate primers based on amino acid sequences shared by several kexin proteases and amplified a 603-bp PCR product from *A. niger* genomic DNA. A genomic *Sal*I fragment of approximately 2 kb hybridized with the probe. We recovered two *Sal*I fragments of similar size from an *A. niger* genomic λ–library. The fragments contained adjacent parts of the complete ORF of the gene we designated *kexB* (Fig. 1). The ORF encodes a protein of 844 amino acids interrupted by a putative 51-bp intron located 1,763 bp downstream of the start codon. We identified a partial 1-kb cDNA clone in a cDNA library of *A. niger* that confirmed that the intron was positioned correctly. The cDNA ended in a poly(A) tail 133 bp downstream of the inferred stop codon. Southern analysis performed under conditions of low stringency indicated that the *kexB* gene is a single copy gene.

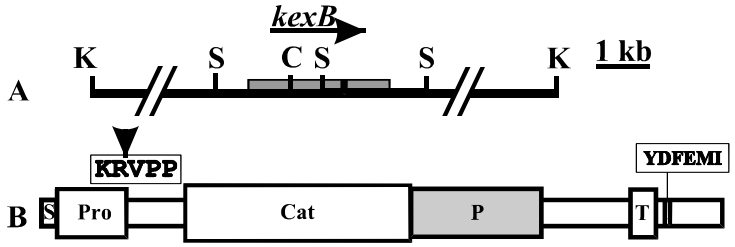


Fig. 1. Sequence characteristics of the *A. niger kex*B gene. (A) Partial restriction map of the *kexB* genomic region. The position of the *Cla*I (C), *Kpn*I (K), and *Sal*I (S) used in cloning strategies are indicated. The distance between the *Kpn*I sites flanking the cloned region is approximately 10 kb. The ORF is indicated with a gray box, the arrow indicates the direction of transcription. The position of the single intron in *kex*B is indicated with a black box. (B) Putative domains of the KexB protease. The KexB amino acid sequence is indicated by an open box. The following domains could be distinguished: signal-sequence (S), prosequence (Pro), subtilisin-like domain (Cat), P domain (P), and transmembrane domain (T). The putative auto cleavage site and the putative Golgi retention signal are also indicated.

Analysis of the *kexB* ORF sequence. The encoded protein is highly similar to the kexin subfamily of proteases and also shows significant similarity with several other subtilisins. The highest similarity was with the Kex2-like protease from the yeast *Yarrowia lipolytica* (Enderlin and Ogrydziak, 1994) yielding 42% overall identity.

As has been described for other kexin proteases, five putative domains were identified in the deduced amino acid sequence (Fig. 1). The first 19 amino acid residues of the ORF were identified by the signalP neural network (Nielsen et al., 1997) as a signal sequence for translocation over the endoplasmic reticulum membrane. The adjacent propeptide probably encompasses amino acids 20 to 129. At the carboxy-terminal end of this segment, there is a Lys-Arg dibasic (auto)cleavage site. Amino acids 190 to 435 contain a subtilisin-like domain. This domain has 53 to 62% identity with the same domains of the yeast Kex2-like proteases and contains the active site residues of the Asp, His and Ser catalytic triad and a conserved Asn residue which stabilizes the oxyanion hole in the transitional state (Bryan et al., 1986). The P-domain (Fig. 1) is not found in other subtilase subfamilies and its presence therefore strongly suggests that we have indeed cloned a member of the kexin subfamily. There also is a putative single membrane-spanning domain. In the putative cytoplasmic tail, 15 amino acids downstream of the transmembrane domain, we found the

peptide sequence YDFEMI. The underlined amino acid residues are identical to the late Golgi retention signal (consensus, YXFXXI) in the cytoplasmic tail of the *Saccharomyces cerevisiae* protease (Wilcox et al., 1992).

In a multiple alignment, the *A. niger* KexB protease was compared to other proteases of the kexin subfamily of proteases and PepC (Frederick et al., 1993) and PepD (Jarai and Buxton, 1994), two *A. niger* subtilases of the proteinase K subfamily (Siezen and Leunissen, 1997). The alignment indicated that the *A. niger* endoprotease is more similar to the yeast-like kexin proteases than to the furins and PCs and is only weakly related to the *A. niger* PepC and PepD subtilases.

Characterization of *kexB* disruptants and overproducers. We transformed strain NW249 (Table 1) with plasmid pIM4002 and identified three multicopy transformants with elevated expression of the *kexB* gene (Fig. 2). We disrupted *kexB* by transformation with plasmid pIM4003. In this plasmid, a part of the subtilisin-like domain, including the Ser residue of the catalytic triad and the conserved Asn, was replaced by the *argB* gene of *A. niger* (Fig. 3). We identified four transformants in which the *kexB* gene was replaced by the non-functional gene and four other transformants in which the disrupted gene was integrated ectopically (Fig. 3). On agar plates, starting from single spores, the four *kexB* disruptants, but not the ectopic transformants, formed compact colonies with no sporulation at the edges. The hyphae of these strains also branched more often and the individual cell segments were shorter, resulting in a dense appearance of the mycelium. In liquid shake flask cultures, there was no clear difference in growth rate and the mycelium appeared much more like wild type mycelium. No unusual phenotype was associated with the *kexB* multicopy transformants.

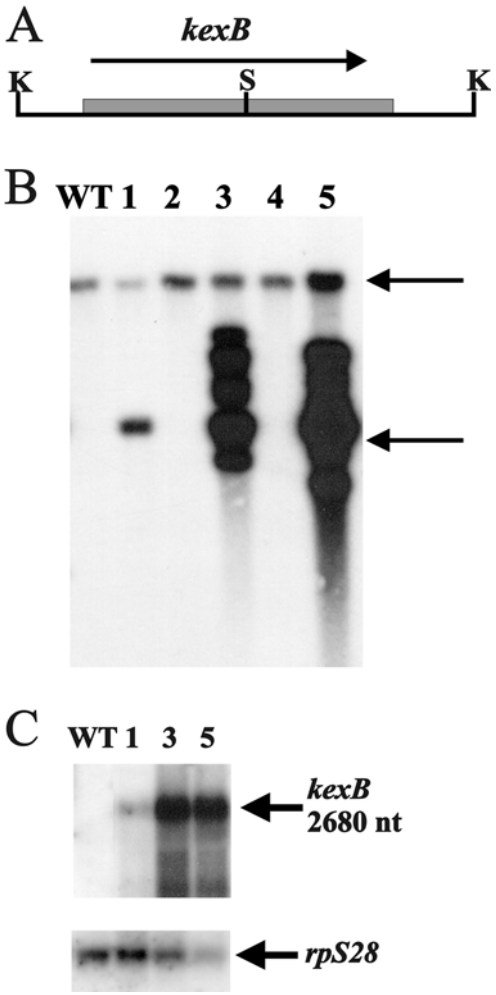


Fig. 2. Molecular characterization of *kexB* multicopy strains. (A) Map of the insert of pIM4002 used for transformation. The *Sal*I sites in the 5' and 3' untranslated regions of *kexB* are replaced by *Kpn*I (K) sites. (B) Southern analysis of *kexB*-overexpressing transformants. *Kpn*I-digested genomic DNA of NW249 and of five transformants was analyzed. The 10 kb band of the endogenous gene and the 4 kb band originating from intact integrated copies of pIM4002 are indicated by arrows. Some scattered integration of pIM4002 is also observed. (C) Northern analysis of *kexB* expression in multicopy transformants. Transformant numbers are indicated above the lanes. In the lower panel, the membrane is rehybridized with ribosomal protein gene *rpS28* to provide a loading control. WT, wild-type.

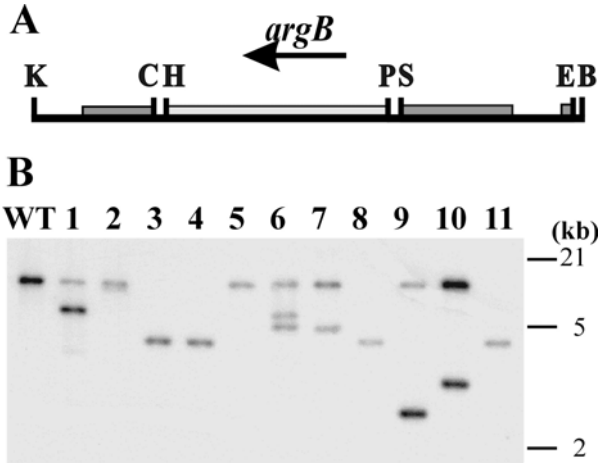


Fig. 3. Molecular characterization of *kexB* disruption strains. (A) Map of the insert of pIM4003 used for transformation. A part of the *kexB* ORF substituted with the *argB* selection marker. The *Bam*H1 (B) *Cla*I (C), *Eco*RI (E), *Kpn*I (K) , *Hin*dIII (H), *Pst*1 (P), and *Sal*I (S) used in the cloning strategy are indicated. (B) Southern analysis of arginine prototrophic transformants. *Pst*I-digested genomic DNA of NW249 and of 11 transformants was hybridized with a *Sal*I-*Eco*RI fragment of pIM4003. The endogenous *kexB* gene hybridizes as a 15-kb fragment (WT lane). A 4.4-kb fragment replaces this fragment if the *arg*B gene replaces a part of the *kexB* coding region (lanes 3, 4, 8, 11). Transformants (lanes 1, 6, 7, 9, and 10) show an ectopic pattern of integration. Transformants (lanes 2 and 5) have integrated only a functional *argB* gene.

Characterization of KexB activity. We measured the relative kexin activity of the detergent-solubilized membrane-protein fraction (DSP) for 4h with Boc-L-K-R-MCA as a fluorogenic substrate in a wild-type (NW249), a *kexB* disruptant (NW266), and a *kexB* multicopy transformant (MCK-5) (Table 1). The DSP fraction of MCK-5 had more than twice as much hydrolyzing activity towards the MCA substrate as the DSP fraction from the wild-type strain. The DSP fraction of the *kexB*-disrupted strain had no significant hydrolyzing activity (Fig. 4). The amount of AMC liberated increased linearly with time if DSP fractions of NW249 or MCK-5 were used, indicating that under the assay conditions the enzyme activity is stable.

Kexin activity is strictly Ca^{2+} dependent (Fig. 5). Without Ca^{2+}, no hydrolyzing activity towards the substrate was observed. Free Ca^{2+} concentration of 1 to 2 mM appears to be optimal for activity of the DSP extract.

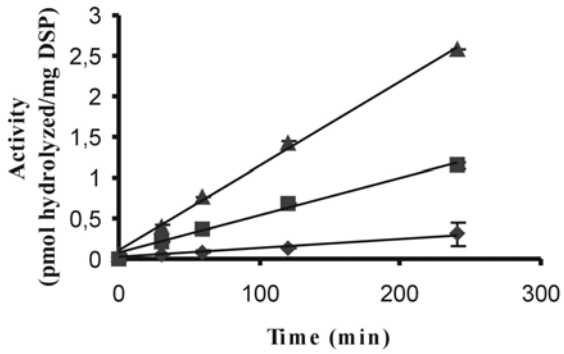


Fig. 4. Time dependence of hydrolysis of Boc-Leu-Lys-Arg-MCA by the DSP fractions of NW249 (■), MCK-5 (▲), and NW266 (♦). The data represent the means of two independent experiments. Standard errors are indicated by bars or are within each symbol.

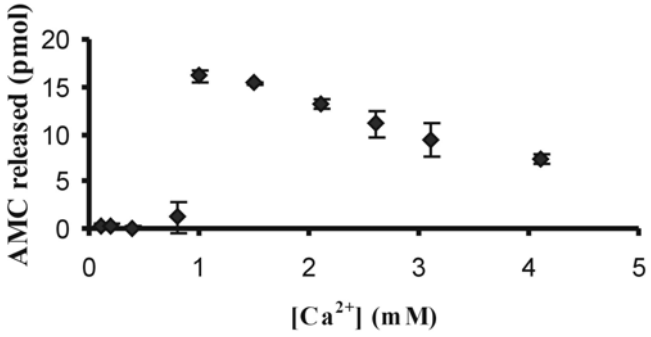


Fig. 5. Ca^{2+} dependency of KexB activity. Incubations were done for 4 h in 200 mM HEPES, 0.2 mM EDTA (pH 7.0) and a varying Ca^{2+} concentration with 10 µl of DSP extract of NW249 (1.33 mg protein/ml) and 100 µM Boc-L-K-R-MCA as a substrate. The data represent the means of two independent experiments. Bars indicate standard errors.

We tested several fluorogenic substrates to determine the substrate specificity of the *A. niger* kexin (Table 2). In all cases, a lysine residue at P2 performs better than an arginine residue. Similarly, a glutamine residue at P3 performs better than a leucine residue. A glycine residue at P3 is not preferred. Furthermore, a typical furin recognition site, the tetrapeptide Arg-Val-Arg-Arg (Nakayama, 1997) is also a good substrate for KexB. The

tripeptide Glu-Lys-Lys and the tetrapeptide Leu-Ser-Thr-Arg are poor substrates for KexB. The membrane-protein fraction of the *kex*B-disrupted strain showed no significant activity towards the MCA substrates tested (Table 2).

Table 2. Relative efficiency of cleavage of peptidyl-MCA substrates[a]

Peptide substrate						Relative activity %		
P5	P4	P3	P2	P1	P1'	NW249	MCK-5	NW266 (*ΔkexB*)
	Boc	Gly	Arg	Arg	MCA	9	24	5
	Boc	Gly	Lys	Arg	MCA	34	81	9
	Boc	Leu	Arg	Arg	MCA	30	74	8
	Boc	Leu	Lys	Arg	MCA	100	220	17
	Boc	Leu	Gly	Arg	MCA	3	7	5
	Boc	Gln	Arg	Arg	MCA	71	180	14
	Boc	Gln	Gly	Arg	MCA	10	29	14
	Boc	Glu	Lys	Lys	MCA	2	5	3
	Boc	Val	Pro	Arg	MCA	29	50	7
Boc	Leu	Ser	Thr	Arg	MCA	11	21	2
Boc	Arg	Val	Arg	Arg	MCA	68	130	7

[a]The activity is expressed as a percentage of the activity obtained with the DSP fraction of the wild type strain (NW249) on Boc-Leu-Lys-Arg-MCA (0.51 pmol AMC liberated/μg DSP). The concentrations of the DSP fractions of the wild-type strain, MCK-5, and NW266 were 1.33, 0.49, and 0.67 mg/ml respectively.

KexB cleavage of an engineered Kex2 site in a reporter construct. Strains MCGI (wild-type) and MCGIΔ (*kexB* disruptant) (Table 1) express a fusion gene encoding a glucoamylase-human interleukin-6 fusion protein separated by a Kex2 site at the fusion junction. Western analysis of the reporter showed that while the wild-type strain secretes glucoamylase and interleukin-6 separately, the *kexB* disruptant is unable to hydrolyze the Kex2 site at the fusion junction and therefore only secretes the intact fusion protein (Fig. 6).

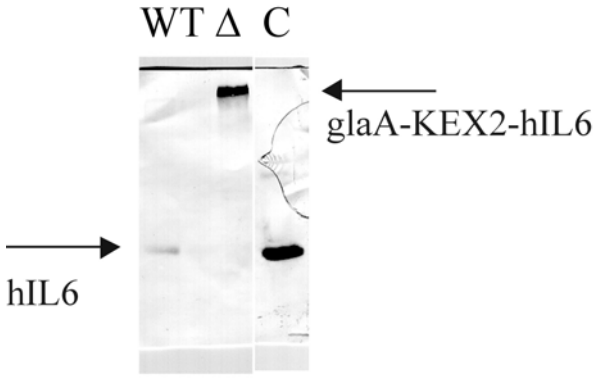


Fig. 6. Western analysis of glucoamylase-human interleukin-6 processing. Medium samples (0.5 ml) of MCGI (wild-type) and MCGIΔ (*kexB* disruptant) (Δ) were analyzed for the presence of unprocessed human interleukin-6 (glaA-KEX2-hIL6). The control lane (C) contains 0.5 μg of recombinant human interleukin-6 (hIL6).

Discussion

We used a PCR approach to clone a kexin homologue in *A. niger*. Kexins have been cloned and characterized from yeasts and mammalian cells, and thus from an evolutionary perspective, the existence of such a function was predictable. The work done with polygalacturonases (Bussink et al., 1990; Bussink et al., 1991; Bussink et al., 1992; Pařenicová et al., 1998) and with fusion proteins that use engineered Kex2 sites at the fusion junction (reviewed in Gouka et al., 1997) suggested that this function exists in *A. niger*. The *kexB* gene we cloned encodes a Kex2-like dibasic endoprotease that has significant similarity to expressed sequence tags (ESTs) of two other filamentous fungi. One EST (genbank accession no. AA787123) originated from *Aspergillus nidulans,* and a translation from that EST has 64% identity with KexB. The other EST (genbank accession no. AIO68899) was from the pyrenomycete *Magnaporthe grisea* and a translation from that EST also showed 64% identity with KexB. These findings suggest that this endoprotease function is ubiquitously expressed in filamentous fungi.

Strain MCK-5 has more than 15 copies of the *kexB* gene, and a high level of overexpression of *kexB* in that strain was confirmed by Northern analysis. The relatively low increase in kexin activity in the DSP fraction of strain MCK-5 could be due to a regulatory mechanism operative at the posttranscriptional level. When the Kex2 protease was overproduced in the

yeast *S. cerevisiae*, it was transported to the vacuole at an increased rate and degraded (Wilcox et al., 1992).

The *kexB* disruptant grows very poorly on agar plates and its hyphae have an unusual morphology. Under these circumstances, KexB function appears to be essential for normal growth. In shake flask cultures, however, the disruption of *kexB* has less severe consequences. Biomass is not reduced in liquid culture and there are no indications that secretion of the reporter is hampered in the *kexB*-disrupted strain. When strains with comparable copy numbers were used, the amount of glucoamylase produced as a separate protein in the control strain was comparable to the amount of fusion protein produced by the *kexB* disruptant (NW266).

The lack of processing of MCA derivatives by the DSP of the *kexB* disruptant proves that hydrolysis of the MCA derivatives in the membrane-protein fractions of the wild-type strain and the *kexB* multicopy strain depends on KexB. *In vitro,* the specificity of KexB is more comparable to that of the yeast Kex2 maturase (Brenner and Fuller, 1992) than it is to that of mammalian furin (Nakayama, 1997). However, *in vitro Aspergillus* KexB appears to process the furin-specific sequence Arg-Val-Arg-Arg better than yeast Kex2. Like Kex2, KexB does not process the tripeptide Glu-Lys-Lys. In yeast, the amino acid at the P4 position is important for the cleavage of a Lys-Lys dibasic cleavage site (Rockwell and Fuller, 1998). If a phenylalanine residue is present at P4, then normal processing of the substrate results. In *Aspergillus niger* PgaI, a similar cleavage site (F-A-K-K) is processed *in vivo* (Bussink et al., 1991).

The results from the *in vitro* experiments are consistent with the inability of the *kexB* disruptant to process the glucoamylase-interleukin-6 fusion protein with an engineered Kex2 site at the fusion junction. The yeast monobasic aspartyl protease Mkc7 can also hydrolyze a dibasic cleavage site (Komano and Fuller, 1995) and monobasic cleavage is operative in *A. niger* (Bussink et al., 1991). However, such compensation for the loss of kexin activity was not observed with this reporter. Clearly, our results identify the *kex*B gene as the kexin-encoding gene of *A. niger*.

Acknowledgments

We thank R. Contreras for providing plasmid pFGPDGLAHIL6 and Y. Müller for technical assistance. This work was supported by Dutch Technology Foundation (STW) grant WBI .4100.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**.** *Nucleic Acids Res***. 25: 3389-3402.**

Baker, D., Shiau, A. K., and Agard, D. A. (1993). The role of pro regions in protein folding. *Curr. Opin. Cell Biol.* **5: 966-970.**

Bos, C. J., Debets, A. J. M., Swart, K., Huybers, A., Kobus, G., and Slakhorst, S. M. (1998). Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. *Curr. Genet.* **14: 437-443.**

Brenner, C., and Fuller, R. S. (1992). Structural and enzymatic characterization of a purified prohormone-processing enzyme: secreted, soluble Kex2 protease. *Proc. Natl. Acad. Sci. USA* **89: 922-926.**

Bryan, P., Pantoliano, M. W., Quill, S. G., Hsiao, H. Y., and Poulos., T. (1986). Site-directed mutagenesis and the role of the oxyanion hole in subtilisin. *Proc. Natl. Acad. Sci. USA* **83: 3743-3745.**

Bussink, H. J. D., Kester, H. C. M., and Visser, J. (1990). Molecular cloning, nucleotide sequence and expression of the gene encoding pre-pro-polygalacturonase II of *Aspergillus niger*. *FEBS Lett.* **273: 127-130.**

Bussink, H. J. D., Brouwer, K. B., de Graaff, L. H., Kester, H. C. M., and Visser, J. (1991). Identification and characterization of a second polygalacturonase gene of *Aspergillus niger*. *Curr. Genet.* **20: 301-307.**

Bussink, H. J. D., Buxton, F. P., Fraaye, B. A., de Graaff, L. H., and Visser, J. (1992). The polygalacturonases of *Aspergillus niger* are encoded by a family of diverged genes. *Eur. J. Biochem.* **208: 83-90.**

Contreras, R., Carrez, D., Kinghorn, J. R., van den Hondel, C. A. M. J. J., and Fiers, W. (1991). Efficient Kex2-like processing of a glucoamylase-interleukin-6 fusion protein by *Aspergillus nidulans* and secretion of mature interleukin-6. *Bio/Technology* **9: 378-381.**

Enderlin, C. S., and Ogrydziak, D. M. (1994). Cloning, nucleotide sequence and functions of XPR6, which codes for a dibasic processing endoprotease from the yeast *Yarrowia lipolytica***.** *Yeast* **10: 67-79.**

Frederick, G. D., Rombouts, P., and Buxton, F. P. (1993). Cloning and characterization of *pepC*, a gene encoding a serine protease from *Aspergillus niger*. *Gene* **125: 57-64.**

Fuller, R. S., Brake, A., and Thorner, J. (1989). Yeast prohormone processing enzyme (Kex2 gene product) is a Ca2+-dependent serine protease. *Proc. Natl. Acad. Sci. USA* **86: 1434-1438.**

Gouka, R. J., Punt, P. J., and van den Hondel, C. A. M. J. J. (1997). Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Appl. Microbiol. Biotechnol.* **47: 1-11.**

de Graaff, L., van den Broek, H., and Visser, J. (1988). Isolation and expression of the *Aspergillus nidulans* pyruvate kinase gene. *Curr. Genet.* **13: 315-321.**

Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14: 378-379.**

Jarai, G., and Buxton, F. P. (1994). Cloning and characterization of the *pepD* gene of *Aspergillus niger* which codes for a subtilisin-like protease. *Gene* **139: 51-57.**

Komano, H., and Fuller, R. S. (1995). Shared functions in vivo of a glycosyl-phosphatidylinositol-linked aspartyl protease, Mkc7, and the proprotein processing protease Kex2 in yeast. *Proc. Natl. Acad. Sci. USA* **92: 10752-10756.**

Kusters-van Someren, M. A., Harmsen, J. A. M., Kester, H. C. M., and Visser, J. (1991). Structure of the *Aspergillus niger pel*A gene and its expression in *Aspergillus niger* and *Aspergillus nidulans*. *Curr. Genet.* **20:293-299.**

Lipkind, G. M., Zhou, A., and Steiner, D. F. (1998). A model for the structure of the P domains in the subtilisin-like prohormone convertases. *Proc. Natl. Acad. Sci. USA* **95: 7310-7315.**

Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R., and Thomas, G. (1992). Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen. *J. Biol. Chem.* **267: 16396-16402.**

Nakayama, K. (1997). Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. *Biochem. J.* **327: 625-635.**

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* **10: 1-6.**

Pařenicová, L., Benen, J. A. E., Kester, H. C. M., and Visser, J. (1998). *pgaE* encodes a fourth member of the endopolygalacturonase gene family from *Aspergillus niger*. *Eur. J. Biochem.* **251: 72-80.**

Pontecorvo, G., Roper, J. A., Hemmons, L. J., MacDonald, K. D., and Bufton, A. W. J. (1953). The genetics of *Aspergillus nidulans*. *Adv. Genet.* **5: 141-238.**

Rockwell, N. C., and Fuller, R. S. (1998). Interplay between S_1 and S_4 subsites in Kex2 protease: Kex2 exhibits dual specificity for the P_4 side chain. *Biochemistry* **37: 3386-3391.**

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: a Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Schechter, I., and Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem. Biophys Res. Commun.* **27: 157-162.**

Seidah, N. G., Hamelin, J., Mamarbachi, M., Dong, W., Tadros, H., Mbikay, M., Chretien, M., and Day, R. (1996). cDNA structure, tissue distribution, and chromosomal localization of rat PC7, a novel mammalian proprotein convertase closest to yeast kexin-like proteinases. *Proc. Natl. Acad. Sci. USA* **93: 3388-3393.**

Siezen, R. J., and Leunissen, J. A. M. (1997). Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Sci.* **6: 501-523.**

Spencer, J. A., Jeenes, D. J., MacKenzie, D. A., Haynie, D. T., and Archer, D. B. (1998). Determinants of the fidelity of processing glucoamylase-lysozyme fusions by *Aspergillus niger*. *Eur. J. Biochem.* **258: 107-112.**

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22: 4673-4680.**

Wilcox C. A., Redding, K., Wright, R., and Fuller, R. S. (1992). Mutation of a tyrosine localization signal in the cytosolic tail of yeast Kex2 protease disrupts Golgi retention and results in default transport to the vacuole. *Mol. Biol. Cell* **3: 1353-1371.**

Chapter 3

KexB of *Aspergillus niger* is responsible for monobasic proteolytic processing of secreted proteins

Jalving, R., H. C. M. Kester, and P. J. Schaap
Manuscript in preparation

KexB of *Aspergillus niger* is responsible for monobasic proteolytic processing of secreted proteins.

Ruud Jalving, Harry C. M. Kester, and Peter J. Schaap

Wageningen University and Research Centre, Laboratory of Microbiology, Section Fungal Genomics, Wageningen, The Netherlands

Abstract

KexB is an *Aspergillus niger* Golgi located endoprotease that removes the propeptide from proteins destined for secretion. The recognition site for this processing event is the dibasic R-R or K-R motif, present in the target proteins. In a *kexB* disrupted strain this activity is absent, resulting in pH-dependent hyper-branching morphology. At pH 4 it has a normal branching morphology, hyper-branching occurs at pH 6. Furthermore we found that the secreted enzyme endopolygalacturonase II, containing a monobasic processing site was not matured in a Δ*kexB* strain at either pH. So far these monobasic processing sites were thought to be specifically processed by other maturases. To analyse a possible activity of KexB on monobasic processing sites, we constructed a soluble secreted form of KexB. We measured the monobasic processing activity of this soluble maturase *in vitro* using artificial substrates. The soluble form of KexB efficiently cleaves A-R, P-R and S-R, while no activity was found on G-R.

Introduction

Protein maturation is a commonly occurring phenomenon in eukaryotic cells. For proteins destined for secretion, maturation includes the removal of information for sorting and modification, often stored in N-terminally located pre- and propeptides (Blobel, 2000). The prepeptide is removed by signal peptidase upon entry of the endoplasmatic reticulum. The propeptide is removed upon or before departure from the secretory pathway by maturases (Baker et al., 1993). This group of endoproteases resides in either the late stage of the Golgi, the secretory vesicles or is extracellularly anchored to the cytoplasmic membrane with a GPI-anchor.

One or two basic amino acid residues on the target protein comprise the common recognition site for a maturase to remove the propeptide. The kexin family of maturases, identified and characterised in both lower and higher eukaryotes, recognises dibasic amino acid motifs (Fuller et al., 1989, Zhou et al., 1999). Kexin-like maturases in lower eukaryotes, for which the yeast Kex2 protease is a common model, prefer the processing sequences K-R and R-R. Higher eukaryotes have several active kexin-like maturases, which can be grouped into prohormone convertases and furins. The prohormone convertases cleave after (K/R) – $(X)_n$-(K/R) sequences, where n = 0, 2, 4 or 6, and X is any amino acid except Cys and rarely Pro (Seidah and Chrétien, 1999). Furins have a preference for the sequences R-X-K-R and R-X-R-R, but at a lower efficiency they also cleave the sequence R-X-X-R (Molloy et al., 1999). Deviation from these preferred sequences results in a dramatic decrease in the processing rate or even complete loss of processing of the target protein (Ledgerwood et al., 1995). Monobasic processing sites (X-R), which are the preferred sequences for aspartic yapsin-like maturases, are only cleaved by some kexin-like maturases at a very low frequency.

The yapsin-like family, of which four members have been characterised so far, processes the majority of monobasic cleavage sites (Olsen et al., 1999). Members of this family are GPI anchored cell surface aspartyl proteases. Three of these four characterised yapsin-like maturases yapsin 1 (yps1), yapsin 2 (yps2), and yapsin 3 (yps3) are from *Saccharomyces cerevisiae*,. Yapsin A is a mammalian yapsin-like maturase. No fungal yapsin-like maturases have yet been identified, however monobasic specific cleavage has been reported to occur in *A. niger* as was demonstrated on artificial glucoamylase-interleukin fusions (Spencer et al., 1998). In the vicinity of a well-cleaved Kex2 site, located in the linker between the two fused proteins, monobasic processing sites were introduced. Before the changes the fusion protein was cleaved only at the Kex2 site, after the changes aberrant processing occurred after these single basic residues.

Fungal kexin-like maturases have been cloned and characterised in *A. niger*, designated KexB (Jalving et al., 2000), and in *Aspergillus nidulans*, designated KpcA (Kwon et al., 2001). Disruption of the maturase resulted in both strains in a hyper-branching morphology. Here we describe the morphology of the disrupted *kexB* strain in more detail and characterise the monobasic proteolytic processing activity of KexB.

Materials and methods

Strains, media and culture conditions. The *A. niger* strains used in this study are listed in Table 1. The strains NW289 and NW266 were used for transformation as previously described (Kusters - van Someren et al., 1991). All transformations were co-transformations with one of the used pIM vectors and vector pGW635 (Goosen et al., 1987), containing *A. niger pyrA*. *Escherichia coli* DH5α was used for plasmid transformation and propagation.

Table 1. *A. niger* strains used in this study

Strain	Genotype	reference
NW249	*ΔargB, pyrA6, nicA1, leuA1, cspA1*	van de Vondervoort and Muller, unpublished data
NW266	*$argB_{nig}$::ΔkexB, ΔargB, pyrA6, nicA1, leuA1, cspA1*	Jalving et al., 2000
635.15	*pkiA-pgaI* multicopy derivative of NW156 (*cspA1, pyrA6, leu-13, prtF28*)	Benen et al., 1999
607.16	*pkiA-pgaII* multicopy derivative of NW154 (*cspA1, pyrA6, prtF28*)	Kester and Benen, unpublished data
NW266:: pkiA-pgaI	*pkiA-pgaI* multicopy derivative of NW266	this study
NW266:: pkiA-pgaII	*pkiA-pgaII* multicopy derivative of NW266	this study
NW289	*$argB_{nid}$::ΔpepA, $argB_{nig}$::ΔkexB, pyrA6, leuA1, prtF, cspA1*	this study
NW289:: pIM4006#4	pIM4006 multicopy derivative of NW289	this study
NW289:: pIM4006#12	pIM4006 multicopy derivative of NW289	this study
NW289:: pIM4008#1	pIM4008 multicopy derivative of NW289	this study

All *A. niger* strains were grown on minimal medium (MM), complete medium (CM) (Pontecorvo et al., 1953) or in McIlvaine buffered CM. The strains were grown in 250 ml shake flasks or 1 l shake flasks in an Inova incubator shaker at 250 rpm at 30°C. As carbon source 100 mM glucose was used, except for the ssKexB and polygalacturonase I (PGI) / polygalacturonase II (PGII) purification for which respectively 100 mM

xylose and 100 mM fructose were used. To stimulate growth 0.1% yeast extract was added to the MM for the purification of PGI and PGII.

For the pH buffered experiments the *A. niger* strains were grown on McIlvaine buffered CM. In McIlvaine buffered CM medium the KH_2PO_4 was replaced by $Na_2HPO_4 \cdot 2H_2O$ (pH 4: 17.7 g/l; pH 5 25.4 g/l; pH 6 35.3 g/l) and citric acid monohydrate (pH 4: 15.8 g/l; pH 5: 13.6 g/l; pH 6: 10.7 g/l). The amount of KCl was increased to 30.3 g/l, 26 g/l, 19.9 g/l respectively for pH 4, 5, 6 to produce a similar ionic strength for each pH.

DNA techniques and plasmids. The *pkiA-pgaI* and *pkiA-pgaII* constructs, pIM3700 and pIM3710, have been described by Benen *et al.* (1999). The vectors pIM4006 and pIM4008 are described in Fig. 1. Site directed mutagenesis was used to generate a stop-codon in the downstream *Sal*I fragment of *kexB*. The primers CGATGATGAAGAACTCTAGAAGACCCCTAGTAC and GTACTAGGGGTCTTCTAGAGTTCTTCATCATCG were used in the mutagenesis procedure. The site directed mutagenesis was performed with the Quick Change™ Site-directed Mutagenesis kit from Stratagene, according to the 1995 instruction manual, provided by the manufacturer. The altered *kexB* fragment was inserted behind the downstream *Sal*I *kexB* fragment resulting in pIM4006, a pBluescript SK(+) based vector. The pIM4006 plasmid was used to construct pIM4008. The *Sal*I-*Nco*I fragment, containing the original *kexB* promoter was replaced by a *Pst*I-*Sal*I *xlnA* promoter fragment, obtained from an altered pIM103, in which a *Sal*I site was incorporated via site directed mutagenesis. The primers used for the incorporation of the *Sal*I site are GTCGACATTGAAGAAAGCTGAATGCTG and CAGCATTCAGCTTTCTTCAATGTCGAC. For a detailed construction scheme see Fig 1b. Standard DNA manipulations were carried out essentially as described by Sambrook *et al.* (1989). For Southern analysis of genomic DNA, total DNA from *A. niger* was isolated as previously described (de Graaff et al., 1988).

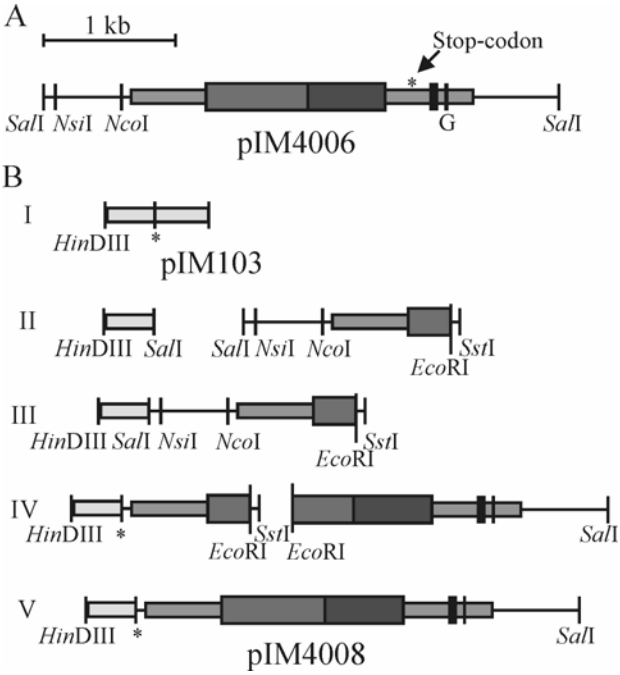


Fig. 1. The constructs used for soluble secreted KexB expression. A. pIM4006. The grey box indicates the *kexB* gene, the two thick grey boxes indicate from the left to the right, the subtilisin domain and P domain. The black box indicates the transmembrane domain. In addition the Golgi retrieval signal (G) and restriction sites have been indicated. The asterisk indicates the residue changed by site directed mutagenesis. B. Construction scheme of pIM4008. I) Site directed mutagenesis of pIM103 to generate a *Sal*I site. II) Cloning of the upstream region of *kexB* behind the *xlnA* promoter (light grey box). III) Removal of the *Sal*I-*Nco*I fragment containing the *kexB* promoter. IV) Cloning of the downstream *kexB* fragment behind the first part to generate V) pIM4008.

Enzyme purifications. PGI and PGII were isolated from 250 ml cultures. After 20 h of cultivation, the mycelium was removed by filtration over nylon gauze. The culture filtrates were adjusted to the appropriate pH, pH 6 for PGI and pH 3.8 for PGII. PGI was subsequently bound to a Streamline-Q matrix (Amersham Pharmacia Biotech). A 10 mM piperazine buffer, pH 6.0 was used to equilibrate and wash the matrix. PGII was bound to a Streamline-SP matrix. The Streamline-SP matrix was equilibrated and washed with 10 mM sodium citrate buffer, pH 3.8. Bound protein was eluted with either the sodium citrate buffer or the piperazine buffer, containing an additional 1 M NaCl. Fractions containing PGI or PGII were identified by SDS-Page. After dialysis with the mentioned equilibration

buffer, the fractions containing PGI were loaded on a Source 30 Q column (for analysis a 1 ml Resource Q column), those containing PGII on a Source 30 S column (for analysis a 1 ml Resource S column). The Source 30 Q column was equilibrated and washed with 20 mM sodium acetate buffer, pH 4.8, the Source 30 S column with 10 mM sodium citrate buffer, pH 3.8. Bound protein was eluted with buffer containing a 0-0.4 M NaCl gradient. The protein containing fractions were identified by measuring the A_{280}. Both the PGI and PGII fractions were identified by SDS-Page.

For the purification of a soluble secreted form of KexB (ssKexB), *A. niger* strain NW289::pIM4006 was grown as two 300 ml cultures for 24 h. The culture filtrate was adjusted to pH 6.0 and diluted 5 times. The diluted culture filtrate was bound to a Streamline-DEAE matrix. The matrix was equilibrated and washed with a 10 mM piperazine buffer, containing 1.5 mM CaCl, pH 6 and eluted with a 10 mM piperazine buffer, containing 1.5 mM CaCl and 1 M NaCl. The resulting fractions were analysed for KexB activity. After dialysis, the 40-60 ml. elute containing most KexB activity was loaded on a Source 30Q column and eluted with a 0 to 0.3 M NaCl gradient. These fractions were also analysed for KexB activity and after dialysis against 10 mM Tris, 1.5 mM CaCl, pH 7.0 the 20-25 ml elute containing most activity was loaded on a Resource-Q column. The buffer used for equilibrating and washing of the column had the same composition as the buffer used for dialysis. Bound protein was eluted from the column with a 0.05-0.4 M NaCl gradient. The amount of KexB activity was determined for each fraction, and the fractions, containing significant KexB activity were analysed by SDS-Page.

Biochemical and morphological characterization of ssKexB. KexB activity was measured using methylcoumarinamide (MCA) derivatives, purchased from Sigma Chemical Co. (St. Louis, Mo.). The assay used has been described previously (Jalving et al., 2000). Protein concentrations were determination by the bichinchoninic acid method as described by the manufacturer (Sigma Chemical Co.). Edman degradation and N-terminal sequencing was done by Eurosequence (Groningen, the Netherlands).

Samples of fungal hyphae were transferred to microscope slides and studied under a microscope equipped with Nomarski's differential interference contrast (DIC) optics. Kodak imagelink negative film, rated at 25 ISO, was used for the Nomarski microscopy.

Results

Characterisation of the *kexB* disrupted strain. We inoculated plates buffered at pH 4, 5 and 6 with the *kexB* disrupted strain NW266 and the parental strain NW249 to determine if external pH influences the morphology of the *kexB* disrupted strain, The ionic strength of the McIlvaine buffer, used to stabilize the pH of the plates, was kept at 0.5 M to distinguish between the influence of ionic charge and pH. At pH 6 we could clearly see reproducible growth retardation and hyper-branching of the Δ*kexB* strain on agar plates (Fig. 2), but at pH 5 and pH 4, it resembles the branching pattern and growth rate of the wild-type strain (only pH 5 is shown). Thus, the effect of the *kexB* disruption on the morphology is pH dependent.

Analysis of endopolygalacturonase proteolytic processing in the Δ*kexB* strain. Previously, we have reported the processing activity of the KexB maturase in Golgi fractions, isolated from a KexB overproducing transformant and a wild-type strain. The processing activity was demonstrated using artificial substrates that mimic the dibasic cleavage site of target protein precursors, such as the processing site of endopolygalacturonase I (PGI) (Jalving et al., 2000). In addition to processing activity of dibasic cleavage sites, the same protein fractions showed a significant activity on Q-A-R-MCA, an artificial substrate that mimics a monobasic cleavage site (data not published). This sequence strongly resembles the monobasic cleavage site -E-A-R- of endopolygalacturonase II (PGII). To analyse whether KexB is directly involved in the maturation of PGI and PGII, we transformed a *kexB* disrupted strain with, respectively, the *pgaI* gene, encoding PGI, or the *pgaII* gene, encoding PGII, both expressed with the strong *pki* promoter. Transformants were selected and screened for the presence of PGI or PGII in the culture fluid by SDS-Page analysis and activity measurements. Both proteins were isolated as active enzymes from the culture fluid of the best producers, using either anion or cation exchange chromatography. The chromatography elution profiles of PGI and PGII isolated from the *kexB* deficient transformants differed from the elution profiles of both proteins isolated from wild-type derived transformants (Fig. 3a), suggesting that no significant maturation of PGI and PGII has occurred in the *kexB* disrupted strain. Mass spectrometric analysis of the isolated proteins from *kexB* deficient transformants also indicated that both proteins were not properly matured in these transformants. To confirm these unexpected results for

PGII, the N-terminal sequence of PGII produced by the *kexB* disruptant was determined (Fig. 3b). The sequence analysis clearly shows that PGII is not properly matured in the *kexB* disrupted strain.

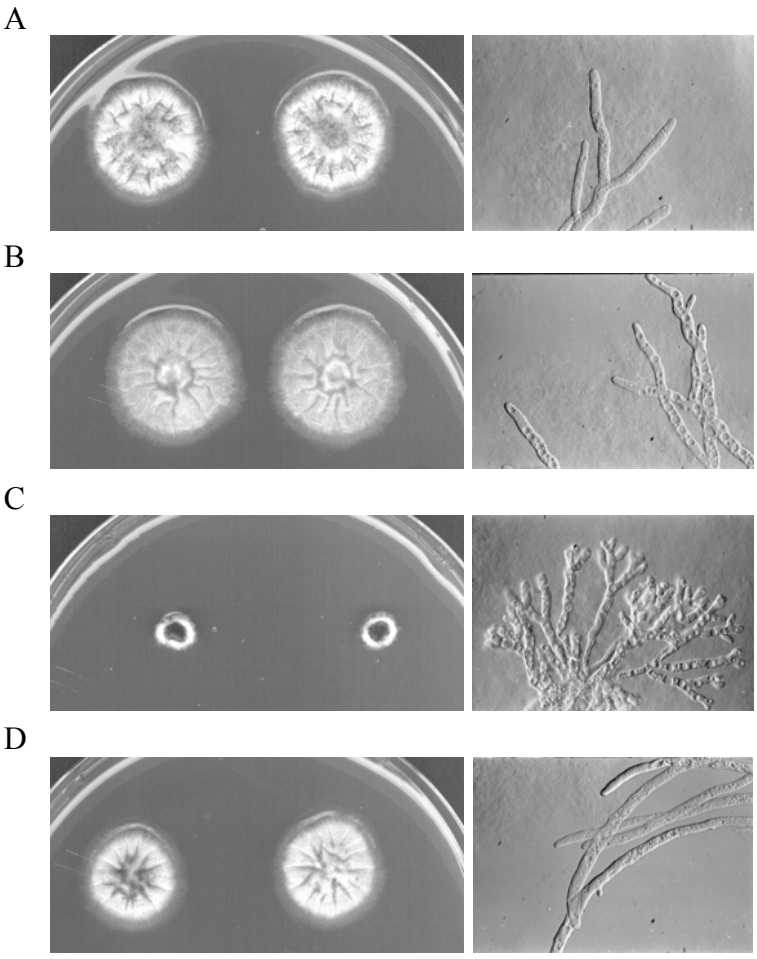
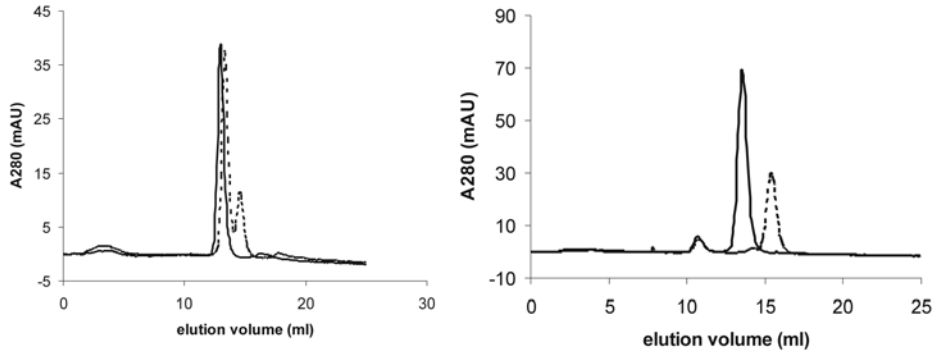


Fig. 2. Growth and morphology of Δ*kexB* strain NW266 (panel A and C) and wild-type strain NW249 (panel B and D) at different pH. On the left are the colonies grown on agar plates of pH 5 (panel A and B) and pH 6 (panel C and D), on the right is the morphology of the hyphae of those colonies visualized by Nomarski microscopy.

A



B

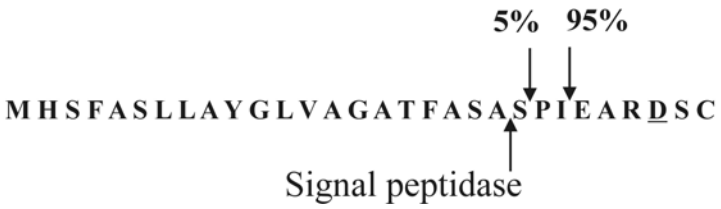


Fig. 3. A. Superimposed ion exchange chromatography elution profiles of endopolygalacturonase I (PGI; graph on the right) and endopolygalacturonase II (PGII; graph on the left) isolated from wild-type strain NW249 cultures (continuous line) and Δ*kexB* strain NW266 cultures (dashed line). B. The N-terminal sequence of PGII secreted by the Δ*kexB* strain. The underlined aspartic acid-residue is the N-terminal end of the mature protein. The presequence cleavage site and the N-terminal ends of the unprocessed PGII obtained by Edman degradation are indicated with arrows.

Construction and analysis of the ssKexB producing strains. For the purification of KexB we constructed a soluble secreted variant of KexB. By site directed mutagenesis a stop codon was introduced in the open reading frame between the proteolytic domain and the transmembrane domain, and the modified gene (Fig. 1a) was used to transform *A. niger* NW289 (Table 1). Transformants which showed a (partial) complementation of the typical Δ*kexB* hyper-branching phenotype at pH 6 were isolated and characterised. Selected transformants were grown for 24 hours in shake flasks after which the mycelium was harvested for Southern analysis and the culture filtrate

was used to determine the amount of KexB activity using Gly-Lys-Arg-MCA as substrate. To analyse the growth characteristics of the transformants they were grown on pH 4 and pH 6 agar plates, and after 3 days the morphology of the mycelium at the rim of the colonies was examined with Nomarski microscopy (Fig. 4a).

A

B

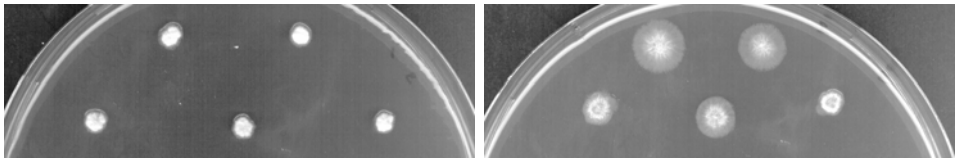


Fig. 4. A) Morphology of the hyphae of the soluble secreted KexB producing pIM4006 transformants. From left to right: strain NW266 (Δ*kexB*), transformant 1 containing approximately 2 copies and transformant 2 containing more than 10 copies of pIM4006. B) Screening of soluble secreted KexB producing pIM4008 transformants. On the left agar plate the transformants are grown on glucose, on the right plate on xylose. Both plates are buffered at pH 6. The top left colony on both plates is transformant NW289::pIM4008#1.

Transformants that have more copies of *sskexB* in the genome also have a higher kexin-like activity in the culture fluid and in addition show a higher degree of complementation of the hyper-branching morphology, indicating that the soluble secreted variant is active. To increase the expression still further, the weak endogenous *kexB* promoter was replaced with the strong *xlnA* promoter from *Aspergillus tubingensis*. This promoter is induced during growth on xylose (de Graaff et al., 1994). The modified construct (Fig. 1b) was used to transform strain NW289. These transformants were analysed on glucose and xylose plates buffered at pH 6

to identify those transformants that were selectively complemented on xylose (Fig. 4b). Grown on xylose transformant NW289::pIM4008#1 had a wild-type morphology. Southern analysis and kexin activity measurements in the culture filtrate of shake flask cultures clearly demonstrated that this transformant produced the highest levels of ssKexB.

Purification of ssKexB and characterization of its substrate specificity concerning dibasic and monobasic processing sites. Transformant NW289::pIM4008#1 was grown in shake flasks for 24 hours and the culture filtrate was bound to a Streamline-DEAE matrix, eluted and subsequently purified in two steps using anion exchange columns, respectively Source-Q and Resource-Q. To separate KexB from other proteins different pH's were used for the two columns. The resulting fractions were analysed by measuring the kexin-like activity on Gly-Lys-Arg-MCA (Table 2). The activity correlated with a single protein peak on the chromatogram (data not shown), and also on SDS-Page a single protein band was visible (Fig. 5). N-terminal sequencing confirmed the identity of the protein.

Table 2. Purification of soluble secreted KexB from *A. niger* NW289::pIM4008#1.

Step	Volume (ml)	Activity (nkat)	Total Protein (mg)	Spec. act. (nkat.mg^{-1})	Purification (-fold)	Yield (%)
Culture filtrate	600	334	41.30	8	1	100
Streamline-DEAE	70	197	6.65	30	3.8	59
Source-30Q (pH 6.0)	22	134	3.74	36	4.5	40
Resource-Q (pH 7.25)	1	18	0.38	49	6.1	5.4

The purified enzyme was used to determine the specific activities on a number of monobasic substrates (Table 3). Of these substrates the monobasic sequences Q-A-R, V-P-R and F-S-R were cleaved respectively 25%, 45% and 40% as effectively as the most preferred dibasic L-K-R sequence.

Table 3. Soluble secreted KexB substrate specificity

MCA derivative	Specific activity (nkat/mg)
Boc-L-K-R-MCA	99.5 ± 12.6
Boc-G-K-R-MCA	37.9 ± 1.11
Boc-L-S-T-R-MCA	10.6 ± 3.29
Boc-E-K-K-MCA	N.A.*
Boc-Q-G-R-MCA	0.4 ± 0.1
Boc-L-G-R-MCA	0.5 ± 0.1
Boc-Q-A-R-MCA	19.1 ± 1.40
Boc-V-P-R-MCA	39.3 ± 4.11
Boc-F-S-R-MCA	37.2 ± 6.16

*N.A. = No Activity

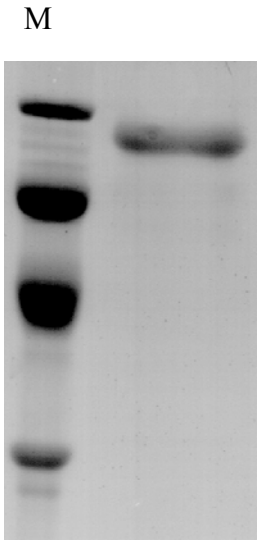


Fig. 5. SDS-Page of active soluble secreted KexB eluted from the Resource-Q column. Lane M, molecular weight marker (from top to bottom: 92, 67, 45, 29). The sample on the right is 5 times concentrated.

Discussion

We have previously reported that the disruption of *A. niger kexB* results in growth retardation and hyper-branching if the fungus is grown on agar plates, while no effect is seen in shake flasks (Jalving et al., 2000). *A. niger* acidifies the medium of shake flasks faster than agar plate medium.

We therefore analysed whether the external pH influences the morphology of the *kexB*-disrupted strain. There are two possible explanations for the demonstrated pH dependent growth. Either lack of KexB processing, results in this pH dependent growth, or at a lower pH other proteases might be induced that are able to complement some of the *kexB* activities. In yeast aspartic proteases can complement some of the Kex2 processing (Olsen et al., 1999). At this stage we do not know whether aspartic proteases play a role in this reversion to wild-type morphology at pH 5 and lower.

The dramatic change in morphology at pH 6 leaves no doubt about the profound effect the *kexB* deficiency has on the processing of *A. niger* proteins. Potential targets for KexB are secreted proteins such as the *A. niger* pectinases. Many of these proteins, including polygalacturonase I (PGI), contain a dibasic cleavage site. Bussink et al. (1991) determined the N-terminal sequence of secreted mature PGI, which starts with the first amino acid following the dibasic cleavage site in the unprocessed protein. The N-terminal sequence of mature *A. niger* PGII, which has also been experimentally verified (Bussink et al., 1990), starts with the first amino acid following a monobasic cleavage site in the unprocessed protein. In yeast, propeptides with a monobasic processing site are not removed by Kex2 but by the aspartic proteases of the yapsin family (Olsen et al., 1999). PGII was therefore thought to be processed by a yapsin-like maturase. However, no direct proof of KexB processing or processing by aspartic proteases of natural substrates has been shown in *A. niger* until now. As expected the dibasic processing site of PGI produced by the *kexB* disrupted strain was not processed. The surprising result was that also PGII is not matured in the *kexB* disrupted strains. Active KexB is therefore required for correct proteolytic processing of both PGI and PGII. A distinct possibility exists that PGI or PGII is processed by another endoprotease, which needs to be activated by KexB, but most likely, PGI and PGII are directly processed by KexB. An analysis of fungal genomes available via the NCBI genome website for the existence of genes encoding yapsin orthologs in the genomes of filamentous fungi did not produce any significant hit.

To investigate the KexB activity on monobasic cleavage sites, we needed to purify the KexB protease. Proteins with a transmembrane domain, such as the kexin-like proteases are difficult to produce. In yeast and mammalian systems this problem was solved by the construction of soluble secreted variants of these maturases (Brenner and Fuller, 1992; Hatsuzawa et al., 1992). We used a similar construct to complement a *kexB-disrupted* strain. Since the soluble secreted variant is not efficiently retained in the late Golgi, only high expression levels are able to complement the phenotype of

the *kexB* deletion. In addition the *kexB* promoter appeared to be too weak to produce enough KexB and had to be replaced by the *xlnA* promoter. The estimated size of the produced ssKexB protein, about 85 kDa, is larger than 61.8 kDa, the calculated size of ssKexB. The difference is most likely due to glycosylation.

The difference in substrate specificity of the ssKexB protease for the two dibasic MCA derivatives with L-K-R and G-K-L as recognition site is identical to the previously (Jalving et al., 2000) measured difference in kexin-like activity measured for both substrates using membrane fractions of the wild-type strain. Furthermore for wild-type membrane fractions we measured a comparable proteolytic activity on Q-A-R-MCA as measured for ssKexB. This clearly indicates that the proteolytic activity of the purified ssKexB protease is identical to the kexin-like activity measured in the wild-type membrane fractions. We were able to show significant activity of the purified ssKexB protease on a range of MCA derivatives with a monobasic cleavage sites and expect that the unmodified KexB protease has similar substrate specificity. Although we could successfully cleavage artificial substrates *in vitro* using ssKexB, we were not able to process purified pro-PGI and pro-PGII *in vitro* with ssKexB, which might be due to differences in glycosylation of the endopolygalacturonases produced by the Δ*kexB* strain and the wild-type strain.

In summary, deletion of *kexB* results in a pH dependent hyper-branching phenotype. Furthermore KexB is capable of monobasic specific cleavage as we have shown by substrate analysis of ssKexB, and it is necessary for the monobasic specific maturation of PGII. The ssKexB protease could be used to finalize processing of partially processed proteins of industrial interest, or for *in vitro* removal of homologous protein carriers used to protect heterologous proteins produced in fungi. We did not detect any aberrant processing as result of increased production of PGI or PGII in the wild-type production strain, indicating that the wild-type expression of *kexB* normally is sufficient for maturation of large-scale production of proteins containing either a monobasic or dibasic cleavage site.

References

Baker, D., Shiau, A. K., and Agard, D. A. (1993). The role of pro regions in protein folding. *Curr. Opin. Cell Biol.* **5: 966-970.**

Benen, J. A. E., Kester, H. C. M., and Visser, J. (1999). Kinetic characterization of *Aspergillus niger* N400 endopolygalacturonases I, II and C. *Eur. J. Biochem.* **259: 577-585.**

Blobel, G. (2000). Protein Targeting (Nobel Lecture). *Chembiochem* **1: 86-102.**

Brenner, C., and Fuller, R. S. (1992). Structural and enzymatic characterization of a purified prohormone-processing enzyme: secreted, soluble Kex2 protease. *Proc. Natl. Acad. Sci. USA* **89: 922-926.**

Bussink, H. J. D., Brouwer, K. B., de Graaff, L. H., Kester, H. C. M., and Visser, J. (1991). Identification and characterization of a second polygalacturonase gene of *Aspergillus niger*. *Curr. Genet.* **20: 301-307.**

Bussink, H. J. D., Kester, H. C. M., and Visser, J. (1990). Molecular cloning, nucleotide sequence and expression of the gene encoding prepro-polygalacturonase II of *Aspergillus niger*. *FEBS Lett.* **273: 127-130.**

de Graaff, L., van den Broek, H., and Visser, J. (1988). Isolation and expression of the *Aspergillus nidulans* pyruvate kinase gene. Curr. Genet. **13: 315-321.**

de Graaff, L. H., van den Broeck, H. C.,van Ooijen, A. J., and Visser, J. (1994). Regulation of the xylanase-encoding xlnA gene of Aspergillus tubigensis. *Mol. Microbiol.* **12: 479-490.**

Fuller, R. S., Brake, A., and Thorner, J. (1989). Yeast prohormone processing enzyme (Kex2 gene product) is a Ca^{2+}-dependent serine protease. *Proc. Natl. Acad. Sci. USA* **86: 1434-1438.**

Goosen, T., Bloemheuvel, G., Gysler, C., de Bie, D. A., van den Broek, H. W. J., and Swart, K. (1987). Transformation of *Aspergillus niger* using the homologous orotidine-5-phosphate decarboxylase gene. *Curr. Genet.* **11: 499-503.**

Hatsuzawa, K., Nagahama, M., Takahashi, S., Takada, K., Murakami, K., and Nakayama, K. (1992). Purification and characterization of furin, a Kex2-like processing endoprotease, produced in Chinese hamster ovary cells. *J. Biol. Chem.* **267: 16094-16099.**

Jalving, R., van de Vondervoort, P. J. I., Visser, J., and Schaap, P. J. (2000). Characterization of the kexin-like maturase of *Aspergillus niger*. *Appl. Environ. Microbiol.* **66: 363-368.**

Kusters - van Someren, M. A., Harmsen, J. A. M., Kester, H. C. M., and Visser, J. (1991). Structure of the *Aspergillus niger pelA* gene and its

expression in *Aspergillus niger* and *Aspergillus nidulans*. *Curr. Genet.* **20: 293-299.**

Kwon, B. K., Han, K. H., Han, K. Y., Yu, S. M., Hwang, S. G., Jeon, B. H., Han, D. M., and Kim, W. S. (2001). Molecular cloning of kpcA gene encoding a Kex2p-like endoprotease from Aspergillus nidulans. *Mol. Cells* **12: 142-147.**

Ledgerwood, E. C., George, P. M., Peach, R. J., and Brennan, S. O. (1995). Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2 protease. *Biochem. J.* **308 (Pt 1): 321-325.**

Molloy, S. S., Anderson, E. D., Jean, F., and Thomas, G. (1999). Bi-cycling the furin pathway: from TGN localization to pathogen activation and embryogenesis. *Trends Cell Biol.* **9: 28-35.**

Olsen, V., Cawley, N. X., Brandt, J., Egel-Mitani, M., and Loh, Y. P. (1999). Identification and characterization of Saccharomyces cerevisiae yapsin 3, a new member of the yapsin family of aspartic proteases encoded by the YPS3 gene. *Biochem. J.* **339: 407-411.**

Pontecorvo G., **Roper, J. A.**, **Hemmons, L. J., MacDonald, K. D., and Bufton, A. W. J. (1953).** The genetics of *Aspergillus nidulans*. *Adv. Genet.* **5: 141-238.**

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

Seidah, N. G., and Chrétien, M. (1999). Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. *Brain Res.* **848: 45-62.**

Spencer, J. A., Jeenes, D. J., MacKenzie, D. A., Haynie, D. T. and Archer, D. B. (1998). Determinants of the fidelity of processing glucoamylase-lysozyme fusions by Aspergillus niger. *Eur. J. Biochem.* **258: 107-112.**

Zhou, A., Webb, G., Zhu, X., and Steiner, D. F. (1999). Proteolytic processing in the secretory pathway. *J. Biol. Chem.* **274: 20745-20748**.

Chapter 4

Characterisation of *Aspergillus niger dapB,* encoding a novel fungal type IV dipeptidyl aminopeptidase

Jalving, R., J. Godefrooij, W. J. ter Veen, A. J. J. van Ooyen, and P. J. Schaap

Submitted for publication

Characterisation of *Aspergillus niger dapB,* encoding a novel fungal type IV dipeptidyl aminopeptidase

Ruud Jalving, Jeroen Godefrooij, Wim J. ter Veen, Albert J. J. van Ooyen, and Peter J. Schaap

Wageningen University and Research Centre, Laboratory of Microbiology, Section Fungal Genomics, Wageningen, The Netherlands

Abstract

We have cloned the *Aspergillus niger dapB* gene. Analysis of the nucleic acid and the deduced protein sequence indicates that the gene encodes a type IV dipeptidyl aminopeptidase (DPP IV). Based upon the deduced protein sequence we predicted the presence of a transmembrane domain. Furthermorre, *dapB* overexpressing transformants displayed an increase in intracellular DPP IV activity. Thus we are the first to report the characterisation of a dipeptidyl aminopeptidase from a filamentous fungus with a transmembrane domain. Using the *dapB* nucleic acid sequence as query, we were able to identify 14 DPP IV encoding genes in the publicly available genomic databases. Phylogenetic analysis shows that in yeasts two gene clusters exist, of which the genes encode DPP IV proteases with a transmembrane domain. With this study we demonstrate that similar to the yeasts in filamentous fungi also two DPP IV encoding gene clusters exist. However, in filamentous fungi, only 1 of the 2 clusters contains genes encoding DPP IV proteases with a transmembrane domain. The second gene cluster present in filamentous fungi contains DPPIV genes encoding extracellular DPPIV proteases. The *dapB* protease gene belongs to the first mentioned cluster. We propose that DapB plays a role in proteolytic maturation of enzymes, produced by *Aspergillus niger*.

Introduction

Proteins destined for secretion contain information necessary for correct modification, routing and folding (Baker et al., 1993). Much of this information is present in the propeptide, which is removed from the protein upon maturation. This proteolytic maturation is also used as mechanism to activate enzymes. The yeast α-mating type factor is an example of a secreted protein which needs this kind of proteolytic activation. In addition to a propeptide, the α-factor precursors contain either two or four copies of the end product. The yeast endoprotease KEX2 removes the propeptide and separates the different α-factor proteins. However after KEX2 processing the α-factor is not yet mature (Anna-Arriola et al., 1994). Either two or three X-Ala dipeptides (X is Asn, Asp, Glu, or Val) still need to be removed from the N-terminus of the resulting products, which is done by a dipeptidyl aminopeptidase.

Dipeptidyl aminopeptidases are proteases which cleave a dipeptidyl moiety from the N-terminus of a polypeptide (Kreil, 1990). They have been classified according to their substrate preference. Dipeptidyl aminopeptidases of type IV (DPP IV) cleave specifically behind an Ala or Pro residue, removing thereby N-terminal X-Ala or X-Pro dipeptides. Due to its unique structural features proline residues often protect proteins from degradation and in some cases even play a role in the regulation of their activation. DPP IV belongs to a limited set of specialised proteases that are capable of cleaving proline-adjacent bonds (Cunningham and O'Connor, 1997).

DPP IV proteases from a variety of organisms such as yeasts, insects, frogs and humans have been characterised and were found to perform many different functions. Dipeptidyl aminopeptidase A (DPAP A), a Golgi localized protease is responsible for α-factor processing in the yeast *Saccharomyces cerevisiae* and plays an important role in the maturation of proteins. The second DPP IV found in *S. cerevisiae*, designated DPAP B, is a vacuolar membrane bound protease (Roberts et al., 1989), of which the *in vivo* function is unknown. DPAP B is processed in the endoplasmic reticulum and Golgi before it reaches the vacuole. Overexpression of DPAP B in yeast mutants that lack DPAP A resulted in complementation of the α-factor maturation deficiency.

In filamentous fungi only one extracellular DPP IV, designated dipeptidyl-peptidase IV, has been characterised in several *Aspergillus* sp. The *A. fumigatus* protein is able to bind to collagen and could therefore play a role in the process of lung invasion by this pathogenic fungus (Beauvais et

al., 1997a; Beauvais et al., 1997b), while the *A. oryzae* extracellular DPP IV is thought to be involved in the degradation of wheat gluten during koji fermentation (Doumas et al., 1998). Activity measurements indicate that this extracellular DPP IV is absent in *A. niger*, but an intracellular DPP IV activity could be detected.

Furthermore in the *A. nidulans* (Dyer et al., 2003) and *A. fumigatus* (Pöggeler, 2002) genomes mating type genes, including α-factor homologues, were identified. The *A. niger* KEX2 ortholog, KexB, has been characterised by Jalving et al. (2000) and also prolidase, a proline specific dipeptidase, has been identified in both *A. nidulans* by Jalving et al. (2002) and *A. niger* (Jalving et al., unpublished results). Prolidase cleaves Xaa-Pro dipeptides, such as the dipeptides produced by DPP IV proteases. Prolidase, however, is located in the cytosol.

In this study we have cloned and characterised a novel type IV dipeptidyl aminopeptidase encoding gene, designated *dapB*, in a filamentous fungus. We propose that DapB is involved in protein maturation.

Materials and Methods

Strains, plasmids and growth conditions. *A. niger* strain NW219 [cspA1, pyrA6, leuA1, nicA1] was used for transformation essentially as described by Kusters - van Someren et al. (1991). For Southern analysis, Northern analysis and analysis of DPP IV activity in the cell free extracts the transformants were grown overnight in 70 ml complete medium (Pontecorvo et al., 1953) with 1% glucose (w/v) as carbon source, in 250-ml Erlenmeyer flasks in an Innova incubator at 250 rpm and 37°C.

Escherichia coli strains LE392 and DH5α (Promega) were used for phage propagation and plasmid transformation, respectively. For cloning purposes plasmid pUC19 (GibcoBRL) and phagemid pBluescript (Stratagene) were used. Plasmid pIM4007 was constructed as described in Fig. 2a.

Molecular biological techniques. Standard DNA manipulations were carried out essentially as described by Sambrook et al. (1989). DNA fragments were sequenced using the Thermo Sequenase fluorescent-labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham Pharmacia Biotech, Uppsala, Sweden) and the ALF automated sequencer (Amersham Pharmacia Biotech). For Southern analysis of genomic DNA, total DNA from *Aspergillus* strains was isolated as described by de Graaff et al. (1988). For Northern analysis RNA was extracted with Trizol reagent. As a loading control, Northern blots were hybridised with *A. niger* DNA encoding 18S

ribosomal RNA. All hybridisations were done in standard hybridisation buffer (Sambrook et al., 1989).

The degenerate forward primer GTNTAYACIGARMGNTAYATG and degenerate reverse primer TGRAARTGNACRTTRTCRTC, were used to obtain the 144 bp fragment. The PCR was performed using an annealing temperature of 45°C. The complete *dapB* gene was obtained by screening of a λ EMBL4 genomic library of *A. niger* with this PCR fragment.

Trace walking. For BLAST analysis against *Podospora anserina* sequences of the NCBI trace archive, a locally installed version of the NCBI BLAST tools (Altschul et al., 1990) was used. (ftp://ftp.ncbi.nlm.nih.gov/blast/executables). We used the windows based BLAST executables of the 2.2.6 release. The *Neurospora crassa* XM_331713 sequence was used to obtain DPP IV related sequences by TBLASTX. The identified traces were subsequently retrieved from the Trace archive and assembled in the Staden Package (Staden et al., 1998). The resulting consensus sequence was exported as FASTA sequence and used for local BLASTN analysis of the *P. anserina* sequences. Additional trace files were retrieved from the trace archive and added to the assembly in the Staden Package. The consensus was used for the next BLASTN analysis and these steps were repeated until the genes of interest were complete.

Comparative molecular analysis. For structural analysis of the *dapB* gene and the translated protein, the BLAST search tools were used to search public databases. SOSUI (Hirokawa et al., 1998) was used to identify putative transmembrane regions. Signal sequences for secretion were identified with SignalP (Nielsen et al., 1997).

For the phylogenetic analysis the DapB protein sequence was used to query publicly available databases by BLASTP. Protein sequences producing e-values $\leq e^{-100}$ were retrieved as well as the corresponding nucleic acid sequences. The protein sequences were used separately to query the yeast and fungal genomic databases by TBLASTN. Genomic sequences which produced BLAST e-values $\leq e^{-100}$ in any of these queries, and which did not contain sequence identical to one of the existing DPP IV sequences were retrieved, and the putative DPP IV was annotated using the BLAST homology. The obtained and deduced protein sequences were aligned using ClustalX (Thompson et al., 1994). A bootstrap analysis (500 replicates) was performed on the resulting alignments using the PHYLIP package version 3.57c (Felsenstein, 1995) . ProtDist, part of PHYLIP, was used with the standard Dayhoff PAM 001 matrix. Fitch, also part of PHYLIP, was used

for the actual tree generation. Treeview (Page, 1996) was used to view the resulting phylogram.

Overexpression and DPP IV activity measurements. After overnight growth on complete medium (CM) (Pontecorvo et al., 1953), cultures of DapB overproducing transformants were harvested by filtration on nylon gauze. The mycelium was rapidly frozen with liquid nitrogen and stored at -80°C.

To obtain cell free extracts harvested mycelium was ground and resuspended in 1 ml extraction buffer (0.1 M Hepes/Tris pH7.0). After 30 seconds on ice the extracts were centrifuged for 15 minutes at 14000 rpm, and the supernatant was collected. DPP IV activity was measured in a 500 µl assay mix (including 25 µl of enzyme solution) containing 0.3 mM of Ala-Pro-pNA (Bachem) and 0.2 M Hepes/Tris pH 7.0. The assay mix was incubated at 37°C and the increase in hydrolysed pNA product formed was followed by measuring the A_{400} in a spectrophotometer.

The Nucleotide sequence reported in this paper has been deposited in the EMBL database under the Accession No. AJ278532.

Results and Discussion

Cloning and analysis of *dapB*. Based on homologous regions between type IV dipeptidyl aminopeptidases of yeast, rat and human two degenerate primers were designed (see materials and methods). With these primers a 144 bp fragment was isolated by PCR, using genomic DNA of *A. niger* as template. The 144 bp fragment was used as a probe to screen a genomic library of *A. niger* in bacteriophage λ. Two positive clones were identified and subsequently purified. Southern analysis of DNA from these phages with the 144 bp fragment resulted in the identification of a 6 kb *Eco*RI fragment. Sequence analysis of this *Eco*RI fragment revealed the presence of an ORF with only 500 bp promoter sequence. Therefore a *Sal*I and an *Eco*RI-*Sst*I restriction product of the 6 kb *Eco*RI fragment were used in a second Southern analysis, identifying three adjacent *Sal*I fragments of 1.3, 1.5 and 2.6 kb, respectively.

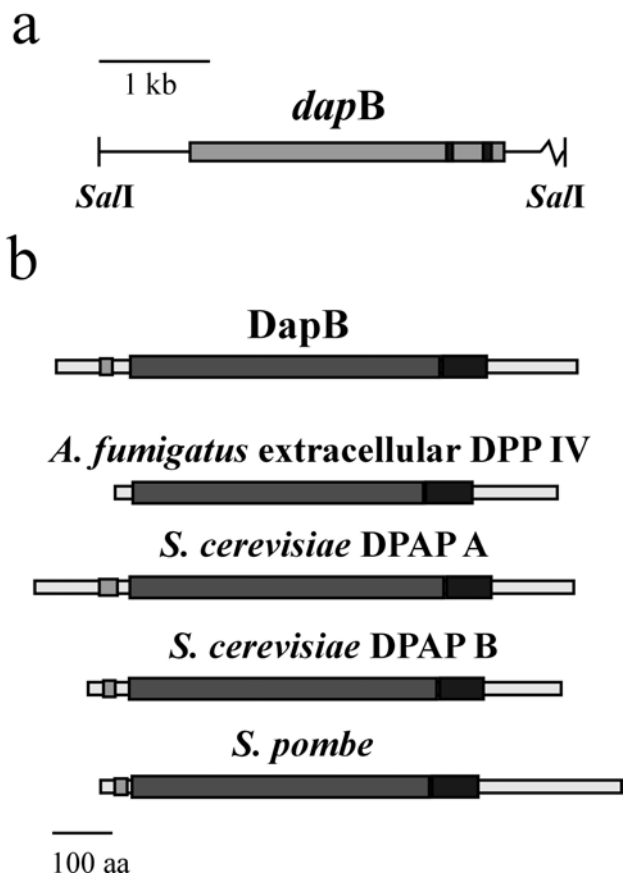


Fig. 1. A) The *dapB* genomic region. The *dapB* gene is indicated with the grey box, the two black boxes represent both introns. B) Structural scheme of the DapB protein and comparison with four other type IV dipeptidyl peptidases. From left to right the three boxes indicate the transmembrane domain, the dipeptidyl peptidase IV N-terminal region (pfam00930.8) and the peptidase S9 domain (pfam00326.8).

The *Sal*I fragments were isolated and sequenced. Analysis of the resulting 3989 bp sequence revealed a 2706 bp ORF starting at bp 817 (Fig. 1a). The ORF is interrupted by 2 introns positioned at bp 3134-3192 and 3463-3544. Both introns were confirmed by comparison of the genomic sequence to *A. niger* EST sequences, which were obtained by BLASTN analysis using the publicly available EST databases. A BLASTX analysis of the sequence was performed using publicly available protein sequence databases and similarity was found with several members of the DPP IV protease family. The highest degree of identity (69 % over 775 amino acids)

was found with the hypothetical protein AN2946.2 of *Aspergillus nidulans.* The next best hits were found with hypothetical proteins of other filamentous fungi. The best hit with a characterised protein was found with the vacuolar *S. cerevisiae* DPP IV protease, DPAP B (39% identity over 674 amino acids). Identities of 37% and 31% were found with the extracellular DPP IV of *A. oryzae* and DPAP A, the Golgi located DPP IV, of *S. cerevisiae,* respectively. The *in vivo* function of DPAP A is proteolytic maturation of proteins. The extracellular *A. oryzae* DPP IV is believed to play a role in degradation of proteins for consumption by the fungus. The *in vivo* function of the other DPP IV proteases is not known. Therefore the results of the BLAST analysis indicate that DapB is a member of the DPP IV protease family but it does not provide any insights into the functional role of the protease in *A. niger.*

The translated amino acid sequence was analysed for the presence of possible structural domains (Fig. 1b). SOSUI identified a putative transmembrane domain, starting at amino acid 77 and ending at amino acid 99. RPS-BLAST analysis resulted in recognition of two typical protease domains, the dipeptidyl peptidase IV (DPP IV) N-terminal region (pfam00930.8 (http://pfam.wustl.edu/index.html)) and the prolyl oligo peptidase domain (pfam00326.8). Therefore, the DapB topology resembles the topology of a type II membrane protein, with an N-terminus located in the cytosol, one membrane spanning region, and the remaining functional part located in the lumen. The functional part of the protein was compared to the functional part of other DPP IV's resulting in the identification of a putative catalytic triad, which consist of Serine 739, Aspartic acid 816 and Histidine 849. The DWVYEEE sequence, which is highly conserved among the different DPP IV's (Abbott et al., 1999) is also present in the DAPB sequence (residues 305-311). This motif contains two glutamic acid residues that are necessary for proteolytic activity. Therefore DapB has all the characteristics of a type IV dipeptidyl peptidase, and it has a putative membrane spanning domain.

Construction of a DapB overproducing strain. To confirm the DPP IV activity of DapB by overexpression, plasmid pIM4007 was constructed containing the complete *dapB* gene, 810 bp upstream and about 900 bp downstream sequence (Fig. 2a). *A. niger* transformants generated with this construct, were selected for the restoration of uridine auxothropy. Subsequently, 5 selected transformants and a *pyrA* complemented control strain were grown overnight and the mycelium was harvested to obtain genomic DNA, RNA and cell free extracts. Southern analysis of the

genomic DNA (Fig. 2b) and Northern analysis (Fig. 3a) were performed to determine if these extra *dapB* copies were present in the genome and if they were actively transcribed in the mycelium of the transformants. In addition to the Northern and Southern analysis the transformants were analysed for the DPP IV activity in the cell free extracts (Fig. 3b). The 5 to 7 times increase in DPP IV activity in the cell free extracts of the DapB multicopy transformants is clearly significant and shows that increased *dapB* mRNA levels lead to more DPP IV activity. No significant DPP IV activity could be measured in the culture medium of any of the transformants (data not shown), making the assignment of the transmembrane domain more likely.

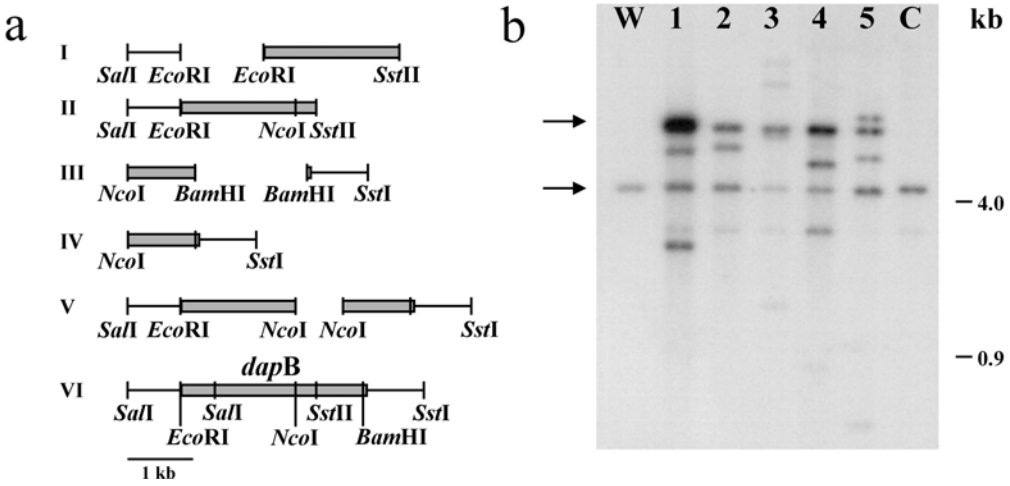


Fig. 2. A) Construction scheme of pIM4007. The *dapB* gene is indicated with the grey box. The following construction steps are indicated: I. Cloning of the *Eco*RI-*Sst*II upstream *dapB* fragment behind the *Sal*I-*Eco*RI *dapB* promoter fragment resulting in II, the upstream part of the construct. III. Cloning of the *Bam*HI-*Sst*I terminator region behind the the *Nco*I-*Bam*HI downstream *dapB* fragment resulting in IV, the downstream part of the construct. V Cloning of the downstream part of the construct (IV) behind the *Sal*I-*Nco*I fragment of the upstream part of the construct (II), resulting in VI, the *dapB* overexpression construct.
B) Southern analysis of *dapB* multicopy transformants. Genomic DNA of the transformants (numbers are indicated above the lanes), of the negative control being A. niger transformed with *pyrA* plasmid (C) and the wildtype (W) strain was digested with *Xho*I/*Hin*DIII. The molecular weights are indicated. The upper arrow indicates tandem insertions of the construct; the lower arrow indicates the endogenous *dapB* gene. As probe a 1.2 kb *Eco*RI-*Pst*I fragment of the *dapB* gene was used.

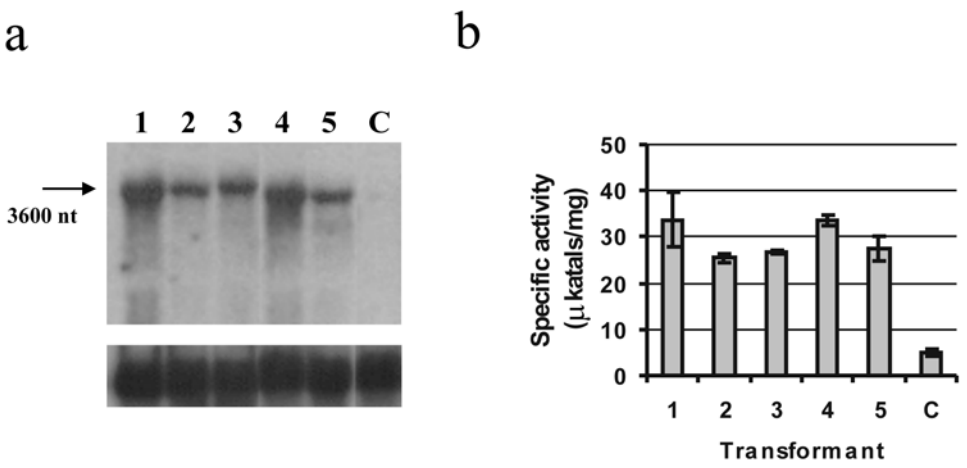


Fig. 3. Analysis of the *dapB* transformants A) Northern analysis of *dapB* multicopy transformants. Transformant numbers are indicated above the lanes. The arrow indicates the position of the *dapB* RNA. A 1.2 kb *Eco*RI/*Pst*I fragment of the *dapB* gene was used as probe. As a loading control, the membrane was rehybridized with DNA encoding the 18S ribosomal RNA (lower panel). B) Specific DPP IV activity in the cell free extracts of the transformants. The numbers below the bars correspond to the transformant numbers, transformant C is *A .niger* transformed with the *pyrA* construct. Ala-Pro-pNA was used as substrate.

Phylogenetic analysis of DPP IV proteases in ascomycetes. To clarify the phylogenetic relationship of DapB with other DPP IV proteases, we analysed publicly available DNA and protein sequence databases by BLAST analysis. As a result we identified 14 new DPP IV-like genes in the publicly available but yet unannotated genomic sequences of *Aspergillus nidulans*, *Coccidioides posadasii, Gibberella zeae, Saccharomyces bayanus*, *Saccharomyces castellii, Saccharomyces kluyveri, Saccharomyces mikatae* and *Saccharomyces paradoxus*. In addition to the previously characterised *Aspergillus fumigatus* extracelullar DPP IV, we obtained the sequence encoding a second DPP IV-like protease in this fungus. This sequence was also previously reported by Pöggeler as putative *A. fumigatus* DPAP A homologue. At the time of our analysis the genomic sequence of *Podospora anserina* was not yet available. However the sequence chromatograms were available via the NCBI Trace archive. We downloaded the entire collection of *P. anserina* sequence files and by local BLAST analysis and trace walking (see Materials and Methods) we were able to identify and assemble two DPP IV-like genes, present in this fungus. Introns were identified in the *A. fumigatus, A. nidulans, C. posadasii, G. zeae* and *P. anserina* genes by

BLAST homology with either the *A. niger dapB* DNA sequence or the DNA sequence encoding the extracellular DPP IV of *A. fumigatus*. The position of the introns in the *A. nidulans* and *N. crassa* genes were confirmed by EST sequence comparison. We identified one additional intron in the annotated DPAP A-like *A. nidulans* gene (BK001296), resulting in both a larger ORF and protein sequence. Interestingly, both the DNA sequences of the putative secreted DPP IV of *G. zeae* and *N. crassa* lack an intron at the start of the gene, which is present in the other extracellular DPP IV encoding genes. The protein sequences, 28 in total, were deduced from the DNA sequences or retrieved from the protein databases and subsequently analysed for the presence of putative domains (Fig. 1b) and aligned with the *A. niger* DapB sequence. The resulting alignment was used for a bootstrap analysis (Fig. 4). Sequences with a low bootstrap value were removed from the analysis, without affecting the structure of the phylogram. The remaining sequences used in the phylogenetic analysis are in Table 1. The resulting phylogram contains four clusters of protein sequences, which contain all at least one characterised protein (underlined). Clusters 1 and 2 contain the Golgi located maturase DPAP A and the vacuolar DPAP B and similar yeast protein sequences. All the protein sequences in cluster 3 are from filamentous fungi. They all resemble the domain structure of DapB and a putative transmembrane domain was identified at a similar position. This is a novel cluster of DPP IV proteases in filamentous fungi. Cluster 4 contains sequences of filamentous fungi similar to the extracellular DPP IV of *A. fumigatus.* SOSUI could detect a transmembrane domain in proteases of this cluster, however it is located in the first 25 amino acid residues and according to SignalP it resembles a signal sequence, which will be removed from the protein before secretion. These results indicate that in most of the analysed filamentous fungi one extracellular DPP IV protease exist and one intracellular DPP IV, containing a transmembrane domain.

We established that in both *A. niger* (results not shown) and *N. crassa* only one DPP IV gene is present in the genome. In both cases it encodes the DapB-like intracellular DPP IV protease. In *N. crassa*, as in several *Aspergilli* the α-factor homologue has been identified (Bobrowicz et al., 2002). It contains the same DPP IV cleavable dipeptide stretches as its yeast counterparts. Furthermore it has been shown that inactivation of this pheromone leads to a disturbed mating-type phenotype in *N. crassa*. If processing of α-factor in *N. crassa* is as essential as in yeast, the *N. crassa* DPP IV protease is likely to be responsible for this removal. We propose that the DPP IV proteases in the DapB cluster (cluster 3) are involved in protein maturation.

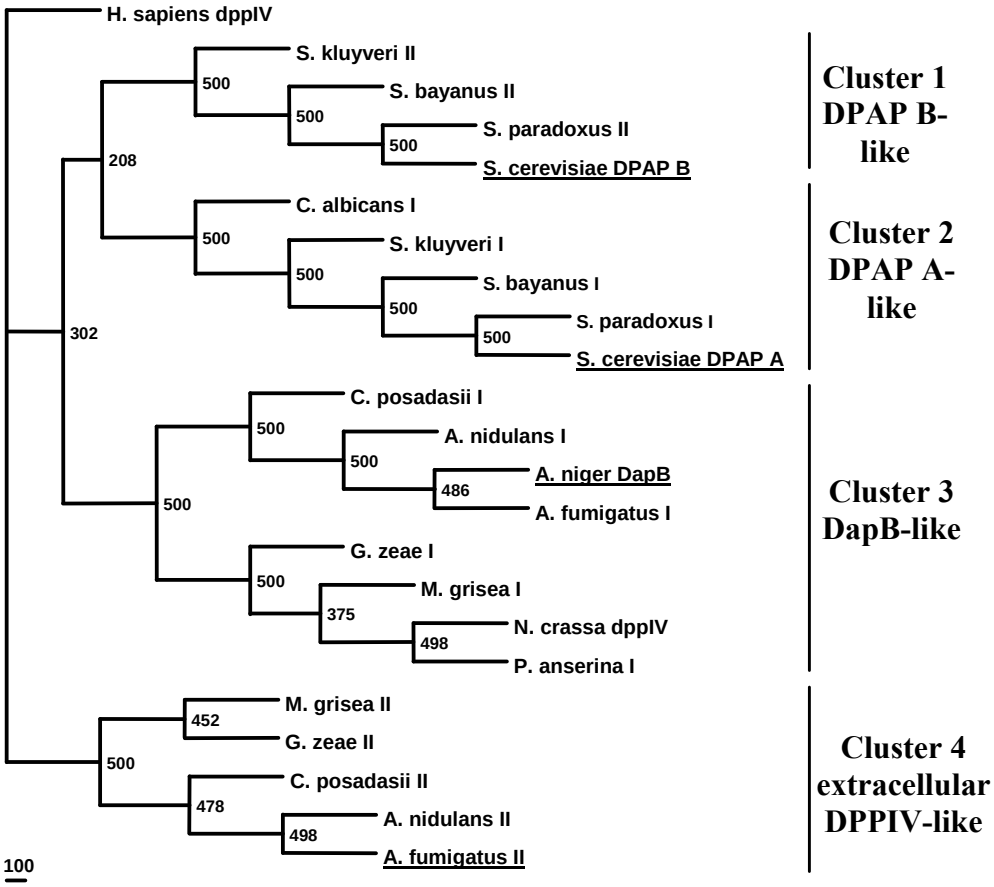


Fig. 4. Phylogenetic analysis of yeast and fungal type IV dipeptidyl peptidases. The analysis contains the following DPP IV subgroups: yeast DPAP B-like (cluster 1), yeast DPAP A-like (cluster 2), DapB-like of the filamentous fungi (cluster 3), and the extracellular DPP IV of filamentous fungi (cluster 4). The human DPP IV was taken as outgroup and does not fit in one of these four categories. The bootstrap values are indicated at each node. The branch lengths indicate the evolutionary distance.

Table 1. Sequences used for phylogenetic analysis.

Organism		Protein identifier	contig (coding basepairs)
Aspergillus fumigatus	I		TIGR_5085|contig:5237 (2102460- 2100090; 2100033-2099764; 2099714- 2099599)
	II	AAC34310	
Aspergillus nidulans	I		AACD01000051 (401192-403520 ; 403589-403858 ; 403940-404058)
	II	EAA58460	
Aspergillus niger	DapB	CAC41019	
Candida albicans	I	orf6.4953.prot	
Coccidioides posadasii	I		TIGR_222929|contig:2046 (587428-589789; 589857-590127; 590191-590309)
	II		TIGR_222929|contig:2010 (628474-628468 ;628374-626050)
Gibberella zeae	I		AACM01000231 (17853-20472; 20524-20636)
	II		AACM01000457 (23575-25914)
Magnaporthe grisea	I	EAA53468	
	II	EAA52861	
Neurospora crassa	DPPIV	EAA36410	
Podospora anserina	I[a]		
Saccharomyces cerevisiae	DPAP A	P33894	
	DPAP B	P18962	
Saccharomyces bayanus	I		AACA01000231 (11642-14434)
	II		AACA01000410 (4540-6996)
Saccharomyces kluyveri	I		AACE01000037 (4699-7341)
	II		AACE01000285 (4461-1993)
Saccharomyces paradoxus	I		AABY01000246 (7389-4594)
	II		AABY01000007 (68079-70535)
Homo sapiens[b]	DPPIV	P27487	

[a] sequence obtained via tracewalking (see M&M); [b] sequence used as outgroup

We can conclude that the *A. niger* DPP IV protease identified and analysed in this study is an intracellular DPP IV protease. Furthermore, by analysing partially sequenced genomes we could identify two distinct DPP IV-like protease clusters in the yeasts and two distinct DPP IV-like protease clusters in filamentous fungi. Of the filamentous fungal DPP IV-like proteases, only the DapB-like protein sequences contain a putative transmembrane domain. A possible DPP IV-like maturase would need a membrane spanning domain to remain in the secretory pathway, where it can perform its maturation function. We propose that DapB of *A. niger* plays a role in the proteolytic maturation of enzymes produced by this fungus.

Acknowledgements

This work was supported by the Dutch Technology Foundation (STW) grant WBI .4100.

References

Abbott, C. A, McCaughan G. W., and Gorrell, M. D. (1999). Two highly conserved glutamic acid residues in the predicted β propeller domain of dipeptidyl peptidase IV are required for its enzyme activity. *FEBS Lett.* **458: 278-284.**

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J (1990). Basic local alignment search tool. *J. Mol. Biol.* **215: 403-410.**

Anna-Arriola, S. S, and Herskowitz, I. (1994). Isolation and DNA sequence of the *STE13* gene encoding dipeptidyl aminopeptidase. *Yeast* **10: 801-810.**

Baker, D., Shiau, A. K., and Agard, D. A. (1993). The role of pro regions in protein folding. *Curr. Opin. Cell Biol.* **5: 966-970.**

Beauvais, A., Monod, M., Debeaupuis, J. P., Diaquin, M., Kobayashi, H., and Latgé, J. P. (1997a). Biochemical and antigenic characterization of a new dipeptidyl-peptidase isolated from *Aspergillus fumigatus*. *J. Biol. Chem.* **272: 6238-6244.**

Beauvais, A., Monod, M., Wyniger, J., Debeaupuis, J. P., Grouzmann, E., Brakch, N., Svab, J., Hovanessian, A. G., and Latgé, J. P. (1997b). Dipeptidyl-peptidase IV secreted by *Aspergillus fumigatus*, a fungus pathogenic to humans. *Infect. Immun.* **65: 3042-3047.**

Bobrowicz, P., Pawlak, R., Correa, A., Bell-Pedersen, D., and Ebbole, D. J. (2002). The *Neurospora crassa* pheromone precursor genes are

regulated by mating type locus and circadian clock. *Mol. Microbiol.* **45: 795-804.**

Cunningham, D. F., and O'Connor, B. (1997). Proline specific peptidases. *Biochim. Biophys. Acta* **1343: 160-186.**

Dyer, P. S., Paoletti, M., and Archer, D. B. (2003). Genomics reveals sexual secrets of *Aspergillus*. *Microbiology* **149: 2301-2303.**

de Graaff, L., van den Broek, H., and Visser, J. (1988). Isolation and expression of the *Aspergillus nidulans* pyruvate kinase gene. *Curr. Genet.* **13: 315-321.**

Doumas, A., van den Broek, P., Affolter, M., and Monod, M. (1998). Characterization of the prolyl dipeptidyl peptidase gene (*dpplV*) from the koji mold *Aspergillus oryzae Appl. Environ. Microbiol.* **64: 4809-4815.**

Felsenstein, J. (1995). PHYLIP: Phylogeny Inference Package. Version 3.57c. University of Washington, Seattle, WA.

Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14: 378-379.**

Jalving, R., Bron, P., Kester, H. C. M., Visser, J., and Schaap, P. J. (2002). Cloning of a prolidase gene from *Aspergillus nidulans* and characterisation of its product. *Mol. Genet. Genomics* **267: 218-222.**

Jalving, R., van de Vondervoort, P. J. I., Visser, J., and Schaap, P. J. (2000). Characterization of the kexin-like maturase of *Aspergillus niger*. *Appl. Environ. Microbiol.* **66: 363-368.**

Kreil, G. (1990). Processing of precursors by dipeptidylaminopeptidases : a case of molecular ticketing. *Trends Biochem. Sci.* **15: 23-26.**

Kusters-van Someren, M. A., Harmsen, J. A. M., Kester, H. C. M., and Visser, J. (1991). Structure of the *Aspergillus niger pelA* gene and its expression in *Aspergillus niger* and *Aspergillus nidulans*. *Curr. Genet.* **20: 293-299.**

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10: 1-6.**

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12: 357-358.**

Pöggeler, S. (2002). Genomic evidence for mating abilities in the asexual pathogen *Aspergillus fumigatus*. *Curr. Genet.* **42: 153-160.**

Pontecorvo, G., Roper, J. A., Hemmons, L. J., MacDonald, K. D., and Bufton, A. W. J. (1953). The genetics of *Aspergillus nidulans*. *Adv. Genet.* **5: 141-238.**

Roberts, C. J., Pohlig, G., Rothman, J. H., and Stevens, T. H. (1989). Structure, biosynthesis, and localisation of dipeptidyl aminopeptidase B, an integral membrane glycoprotein of the yeast vacuole. *J. Cell Biol.* **108: 1363-1373.**

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Staden, R., Beal, K. F., and Bonfield, J. K. (1998). The Staden Package. In: Computer Methods in Molecular Biology. The Humana Press Inc. Totowa.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22: 4673-4680.**

Chapter 5

Cloning of a prolidase gene from *Aspergillus nidulans* and characterisation of its product

Jalving, R., P. Bron , H. C. M. Kester, J. Visser, and P. J. Schaap
(2002)
Molecular Genetics and Genomics **267**: 218-222

Cloning of a prolidase gene from *Aspergillus nidulans* and charcterisation of its product

R. Jalving P. Bron H. C. M. Kester J. Visser P. J. Schaap

Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands

Abstract

Using EST sequence information available from the filamentous fungus *Aspergillus nidulans* as a starting point we have cloned the prolidase-encoding gene, designated *pepP*. Introduction of multiple copies of this gene into the *A. nidulans* genome leads to overexpression of an intracellular prolidase activity. Prolidase was subsequently purified and characterised from an overexpressing strain. The enzyme activity is dependent on manganese as a cofactor, is specific for dipeptides and hydrolyses only dipeptides with a C-terminal proline residue. Although these proline dipeptides are released both intracellularly and extracellularly, prolidase activity was detected only intracellularly.

Introduction

Aspergilli express a large array of proteolytic enzymes (see van den Homberg et al., 1997 and references therein), which allow these fungi to utilise a large variety of protein substrates as nutrients. The major endo- and exopeptidases involved normally do not hydrolyse bonds adjacent to proline residues and proline is therefore thought to protect against excessive enzymatic degradation of proteins (Vanhoof et al., 1995). Instead proline-adjacent bonds are hydrolysed by specific proteases, whose activity depends on the isomeric state of the proline residue and its position in the peptide chain. Post-proline amide-bonds can be cleaved by a prolyl oligopeptidase (EC 3.4.21.26), a prolyl carboxypeptidase (EC 3.4.16.2) and by a proline-specific type IV dipeptidylpeptidase (EC 3.4.14.5) (Cunningham and O'Connor, 1997), all of which are serine peptidases. Both pre- and post-proline bonds can be cleaved by dedicated metallopeptidases. Thus, aminopeptidase P (EC 3.4.11.9) is a manganese-dependent oligopeptidase that cleaves the imide bond between an N-terminal amino acid and a penultimate proline from peptides of any length. For cleavage to occur, the

N-terminal amino acid must have a free amino group, the penultimate residue must be proline and the scissile bond must be in the *trans* configuration (Cunningham and O'Connor, 1997 and references therein). A structurally closely related enzyme is prolidase (EC 3.4.13.9). Prolidase, however, hydrolyses only dipeptides with a C-terminal proline residue. Like aminopeptidase P, the N-terminal amino acid of the substrate must have a free amino group and the imide bond must be in the *trans* configuration (Cunningham and O'Connor, 1997).

Recently a novel secreted type IV dipeptidyl peptidase was identified in various *Aspergillus* spp (Beauvais et al., 1997; Doumas et al., 1998) including *A. nidulans* (Doumas et al., 1998; Jalving and Schaap, unpublished results). Expression of this protease results in an increase in X-Pro dipeptides in the medium. Since we could not detect an extracellular prolidase under these conditions, it seems likely that after uptake of these X-Pro dipeptides by the fungus, an intracellular prolidase activity is needed for their further utilisation. Using sequence data available in the *A. nidulans* EST database as a starting point, we set to clone the *A. nidulans* prolidase gene and to characterise the encoded activity.

Materials and methods

Strains, transformation, DNA and RNA techniques. The *A. nidulans* strain WG312 (*pabaA1*, *pyrG89*, *fwA1*, *uaY9*) (Ballance and Turner, 1985) was used for transformation essentially as previously described (Kusters – van Someren et al., 1991). The *Escherichia coli* strains LE392 and DH5α (Promega) were used for phage propagation and plasmid transformation, respectively. Standard DNA and RNA manipulations were carried out essentially as described by Sambrook et al., 1989.

For Southern analysis of genomic DNA, total DNA from *Aspergillus* strains was isolated as previously described (de Graaff et al., 1988). For Northern analysis strains were cultured for 17 h in 50 ml complete medium (Pontecorvo et al., 1953) with 1 % glucose (w/v) as carbon source in 250-ml Erlenmeyer flasks in an Innova incubator at 250 rpm and 37 °C. RNA was extracted with Trizol reagent. As a loading control, Northern blots were hybridised with *A. niger* DNA encoding 18S ribosomal RNA.

Cloning and sequencing of *pepP*. A 658-bp genomic fragment was amplified from *A. nidulans* genomic DNA, using the primers: 5'-ACTCTTCACTATGGGAAGAAC-3' (forward) and 5'-GTCCTCGATACGAACTCCACC-3' (reverse) and cloned in vector pGEM-T (Promega) Genomic sequences of the *pepP* gene were obtained by

screening an *A. nidulans* genomic library in the bacteriophage Charon 4A (Zimmerman et al., 1980) with this fragment, using standard methods (Sambrook et al., 1989). The isolated 2.6-kb *Eco*RI fragment was subcloned in pUC19 and sequenced on both strands. Sequence analysis was performed using the sequence analysis software package PC/Gene (IntelliGenetics).

Enzyme purification. The PepP transformant (see below) was grown in complete medium for 24 h and the mycelium was harvested. Mycelium frozen in liquid nitrogen was ground using a Waring blender. Buffer A (20 mM TRIS-HCl pH 7.5, 1 mM $MnCl_2$, 0.1 mM EDTA) was added (12 ml/g) and cell debris was removed by centrifugation (10 min, 10,000 x *g*). The cell extract was brought to 50% saturation with $(NH_4)_2SO_4$, and centrifuged (20 min, 10,000 x *g*). The resulting supernatant was brought to 90% saturation and centrifuged (20 min, 10,000 x *g*). The pellet was dissolved in 20 ml of Buffer A, and loaded onto a Source 30Q anion exchange chromatography column (Amersham Pharmacia Biotech). The bound protein was eluted in 5-ml fractions using a 150-ml linear gradient from 0 to 0.6 M NaCl in Buffer A. Active fractions were reloaded on the same column and eluted in 5-ml fractions using a 150-ml linear gradient from 0.1 to 0.4 M NaCl in Buffer A.

Enzyme assays and biochemical characterisation. A standard reaction mixture (210 µl), containing 50 mM bisTRIS-HCl (pH 7.0), 0.05 mM $MnCl_2$, 3.5 mM dipeptide substrate (Sigma) and 2 µl enzyme solution, was incubated at 37 °C for 30 min. After 0, 5, 15, and 30 min 50 µl samples were withdrawn from the reaction mixture and mixed with 200 µl of 1 M sodium acetate (pH 2.8). The amount of the imino acids released was determined with ninhydrin (Troll and Lindsley, 1955), the amino acids with *o*-phthalaldehyde (Roth, 1971). Protein concentrations were determined with the bicinchoninic acid method (Sigma).

The optimal pH for enzymatic activity was determined using 50 mM bisTRIS buffer at pH values ranging from 6.0 to 7.5, 50 mM HEPES buffer from 7.0 to 8.0 and 50 mM TRIS buffer from 7.5 to 9. All buffers contained 0.05 mM $MnCl_2$.

The thermal stability of prolidase was tested by pre-incubation of the purified enzyme (0.02 mg/ml) for 60 min in 50 mM HEPES, pH 7.0, 0.05 mM $MnCl_2$, at 0, 30, 40, 50 and 60 °C, before the standard enzyme assay.

To determine the metal ion requirements for activity, the purified enzyme was incubated for 4 h with 5 mM EDTA, followed by dialysis of the enzyme preparation against 20 mM HEPES pH 7.5. The EDTA-treated enzyme was subsequently incubated with 0.5 mM Mg^{2+}, Mn^{2+}, Zn^{2+}, Co^{2+} and Ca^{2+} for 4 hr and then subjected to the standard assay procedure.

The molecular mass of the native enzyme was determined by size-exclusion chromatography using a Bio-Gel Sec 40 XL column (Bio-Rad) in series with a Bio-Gel Sec 30 XL column at a flow-rate of 0.5 ml/min. The following reference proteins were used: bovine serum albumin (67 kDa), *Aspergillus niger* hexokinase (100 kDa), yeast alcohol dehydrogenase (145 kDa), and beef liver catalase (240 kDa).

The nucleotide sequence reported in this paper has been deposited in the EMBL database under the accession No. AJ296646

Results and discussion

Cloning and analysis of the prolidase-encoding gene of *Aspergillus nidulans.* Strain WG096 was cultured for 24 h in shake flasks, and the culture filtrate and cell extracts were assayed for prolidase activity. The dipeptides tested (Met-Pro, Ala-Pro and Gly-Pro) could all be hydrolysed by the cell extract, but not by the culture filtrate. Using "prolidase" as keyword, the *Aspergillus nidulans* EST database (http://www.genome.ou.edu/asper.htm) was searched for cDNA sequences potentially specifying amino acid sequences of an intracellular prolidase. A number of nearly identical sequences were found which all contained about 750-bp of the 3' portion of the ORF of a putative prolidase gene. The EST sequences were aligned and the resulting consensus sequence was used to design primers for PCR. From genomic DNA of *A. nidulans* a 658-bp fragment was amplified and sequenced. With this fragment a hybridising 2.6-kb *Eco*RI fragment was recovered from an *A. nidulans* genomic library and sequenced. The fragment contained the complete ORF of the putative prolidase gene (designated *pepP*) together with upstream (513-bp) and downstream (408-bp) regions.

The ORF is 496 amino acids long and is interrupted by two introns, both confirmed by comparison of the genomic sequence with the prolidase cDNA sequences available in the *A. nidulans* EST database. The inferred protein sequence was scanned for known sub-cellular sorting and transmembranesequences and none were found. This suggests that, like all other prolidases characterised so far, the *Aspergillus* prolidase resides in the cytosol. Using the BLAST algorithm (Altschul et al., 1997) similar sequences were retrieved from the nucleotide and protein databases. No fungal prolidase sequences have previously been characterised. However, the *Aspergillus* protease is most similar to ORF YFH6, a putative peptidase from the yeast *Saccharomyces cerevisiae*. High similarity scores were

obtained with the prolidases characterised from human and mouse (Endo et al., 1989; Ishii, et al. 1996). Furthermore, a significant similarity was found with aminopeptidase P characterised from *E. coli* (Wilce et al., 1998; Yoshimoto et al., 1989). No significant sequence similarity was found with the putative aminopeptidase P of *A. nidulans* (Jalving and Schaap, unpublished results). A dendrogram depicting the sequence relationships of these proteins with the *A. nidulans* protein was constructed (Fig. 1) using the Clustal X program (Thompson et al., 1997). The dendrogram, displayed with the program TreeView (Page, 1996), shows the strong clustering of *pepP* with yeast ORF YFH6 and shows that some bacterial aminopeptidase P amino acid sequences share a significant sequence similarity with the eukaryotic prolidases. Characterised aminopeptidase P sequences from eukaryotes form a separate group.

Construction of an *A. nidulans* strain overexpressing *pepP* and purification of the fungal prolidase. The molecular data obtained strongly suggested that we had cloned the *A. nidulans* prolidase-encoding gene. To confirm this, and to facilitate purification, a prolidase-overexpressing strain was constructed.

The plasmid containing the 2.6-kb *Eco*RI fragment including *pepP* was introduced into *A. nidulans* WG312 by co-transformation with plasmid pGW635 carrying the *A. niger pyrA* gene (Goosen et al., 1987). Southern analysis was used to screen sixteen $PyrA^+$ strains for multiple integration of the *Eco*RI fragment into the genome (results not shown). Transformants Tr10, Tr11 and Tr15 were selected and cell extracts were tested for increased prolidase activity using the dipeptides Met-Pro and Ala-Pro as substrates. These peptides were the preferred substrates for a partially purified prolidase from *Neurospora crassa* (Johnson and Brown, 1974). Compared to the wild-type strain all three transformed strains showed a considerably higher rate of hydrolysis of both substrates (Table 1). Northern analysis showed that in the three transformed strains the steady-state level of *pepP* mRNA are increased by at least fivefold. The highest mRNA level was observed in Tr11 (Fig. 2A).

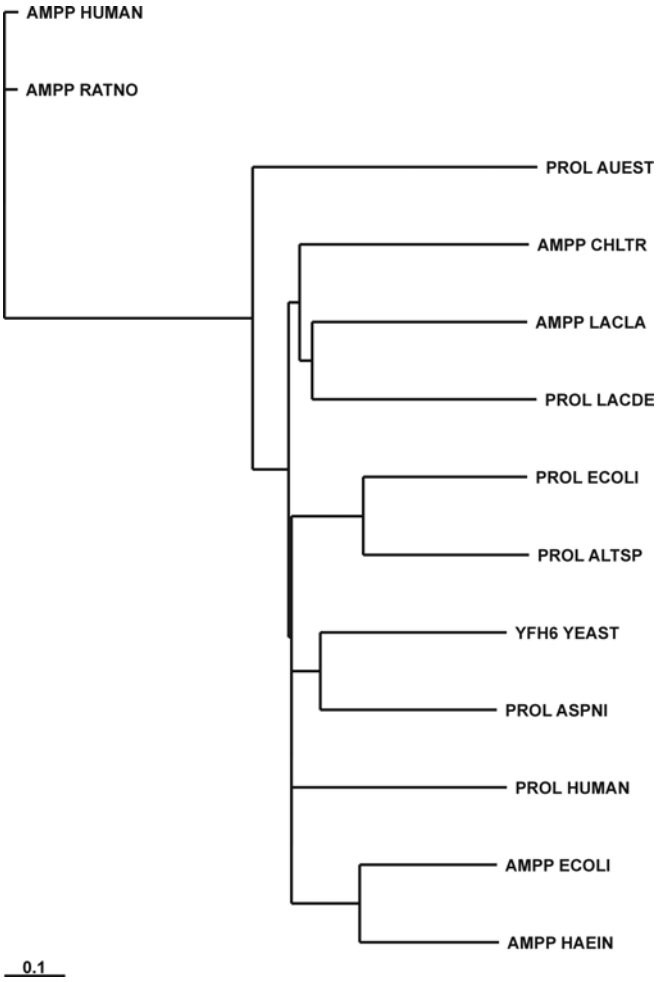


Fig. 1. Dendrogram of sequence relationships between prolidases and proline specific aminopeptidases. The *A. nidulans* prolidase (PROL ASPNI) was compared with prolidases of *Escherichia coli* (X54687.1); PROL ECOLI, *Aureobacterium esteraromaticum* (BAA77794); PROL AUEST, *Alteromonas* sp.JD6.5 (AAB05590); PROL ALTSP, *Homo sapiens* (A32454); PROL HUMAN, *Lactobacillus delbrueckii* (Z93944.1); PROL LACDE, aminopeptidases of *Chlamydia trachomatis* (AE001328.1); AMPP CHLTR, *Escherichia coli* (P15034); AMPP ECOLI, *Haemophilus influenzae* (P44881); AMPP HAEIN, *Lactococcus lactis* (Y08842.1); AMPP LACLA, *Homo sapiens* (NM_006523.1); AMPP HUMAN, *Ratus norvegicus* (AF038591.1); AMPP RATNO and the hypothetical protease YFH6 YEAST of *Saccharomyces cerevisiae* (P43590). Database accession numbers are given in parenthesis.

Table 1. Prolidase activity in cell extracts of *A. nidulans* strains.

Strain	Substrate[a]	
	Ala-Pro	Met-Pro
WG312	0.005	0.008
WG312::pyrA[b]	0.003	0.010
Tr10	0.015	0.082
Tr11	0.110	0.583
Tr15	0.043	0.245

[a]Activities are in nkat / mg protein and are the means of at least two independent assays. [b]Recipient strain WG312 transformed with pGW635.

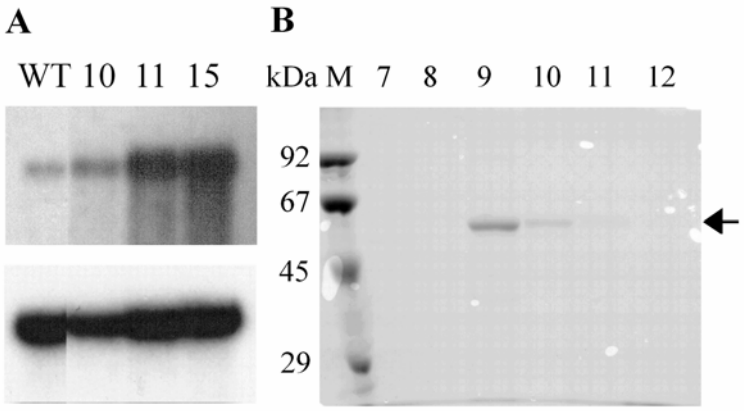


Fig. 2. A. Northern analysis of the *pepP* multicopy transformants. Transformant numbers are indicated above the lanes. WT: strain WG312. As a loading control, the membrane was rehybridised with DNA encoding the 18S ribosomal RNA (lower panel). B. Analysis of active fractions by SDS-PAGE. Fractions (7-12) eluted from the final Source-30Q chromatography step were analysed. Fraction numbers are indicated above the lanes M: Protein molecular weight marker The arrow indicates PepP. The gel was stained with Coomassie Brilliant blue R250.

A cell extract of Tr11 was used to purify the enzyme by ammonium sulphate precipitation, and by repetitive anion-exchange chromatography (Table 2). The purification procedure was monitored using a standard assay with using Met-Pro as the substrate. The final preparation of the prolidase was analysed by SDS PAGE (Fig. 2B). The highest prolidase activity was found in fraction 9 of the Source 30Q elution profile. In agreement with the molecular mass of 55,290 Daltons calculated from the ORF, the molecular mass of the purified protein under denatured conditions was estimated to be

58 kDa. The molecular mass of the native protein was determined by size-exclusion chromatography and was estimated to be 125 kDa. This indicates that, like most reported prolidases (Cunningham and O'Connor, 1997), the native enzyme is a homodimer.

Table 2. Purification of prolidase from *A. nidulans* Tr11.

Fraction	Volume (ml)	Activity (nkat)	Total Protein (mg)	Specific activity (nkat/mg)	Purification factor	Yield (%)
Cell extract	280	33.1	1475	0.022	1	100
$(NH_4)_2SO_4$ precipitation (50-90%)	25	3.9	193	0.020	0.89	11.7
Source-30Q (1)	15	3.9	46	0.085	3.82	11.9
Source-30Q (2)	5	4.0	17	0.23	10.12	11.9

Activity and substrate specificity of the fungal prolidase. The enzyme is active between pH 6.5 and pH 8.5. Highest activity was observed at pH 7 (results not shown). PepP is stable for 60 min at 40 °C. After a 60 min incubation at 50 °C the activity was reduced to 75% of the original level, while 25% remained after a 60-min incubation at 60 °C. A 4-h incubation of PepP with 5 mM EDTA completely eliminated the enzyme activity. The activity of the EDTA-treated enzyme could be restored by addition of Mn^{2+}. Other divalent cations tested were unable to restore the activity.

Besides the dipeptides Met-Pro and Ala-Pro, the enzyme is capable of hydrolysing Phe-Pro, Leu-Pro and Val-Pro. For these five substrates K_M and V_{max} values were determined (Table 3). The highest rate of hydrolysis was obtained with Phe-Pro. The enzyme showed a relatively low rate of hydrolysis of Gly-Pro (15% of the activity on Phe-Pro). A perhaps similar prolidase partially purified from *N. crassa* displayed highest activity towards Met-Pro. This *Neurospora* prolidase also hydrolysed Ala-Pro and Val-Pro efficiently and the least preferred substrate was also Gly-Pro (Johnson and Brown, 1974). The rate of hydrolysis of Phe-Pro has not been reported for the *N. crassa* enzyme.

Table 3. Kinetic parameters for the hydrolysis of dipeptides by prolidase.

Substrate[a]	V_{max} (nkat/mg)	K_M (mM)
Phe-Pro	63	16.4
Ala-Pro	18	10.2
Leu-Pro	15	14.1
Met-Pro	15	38.0
Val-Pro	3.2	6.2

[a]Kinetic parameters were determined from Lineweaver-Burk plots generated using several fixed concentrations of substrate (0, 2, 4, 5, 8, 10, 20, 30 and 40 μM). Each value used to generate the plots was the mean of at least two independent measurements.

When Leu-Pro was used as a substrate, an inhibition of enzyme activity was observed at substrate concentrations higher than 7 mM (Fig. 3). No such inhibitory effect was observed with the other dipepeptides tested, using similar assay conditions. This kind of substrate inhibition specifically by Leu-Pro has also been reported for the prolidase from two *Lactobacillus* spp. (Booth et al., 1990; Fernández-Esplá et al., 1997). Prolidases are only capable of hydrolysing an imido bond and, in agreement with this, the enzyme did not hydrolyse the dipeptides Pro-Ala and Ala-Ala. Moreover, the enzyme could not hydrolyse the imido bond in the tripeptide Ala-Pro-Gly, demonstrating its strict specificity for dipeptides.

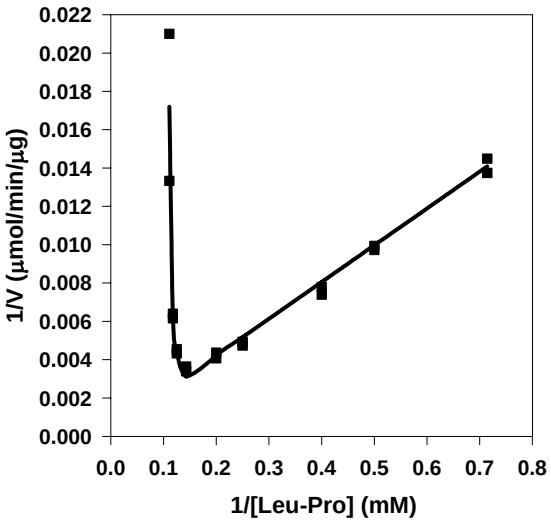


Fig. 3. Lineweaver-Burk plot for prolidase activity with Leu-Pro as substrate.

The results presented here confirm that we have cloned an *A. nidulans* gene that encodes an intracellular proline-specific dipeptidase. The enzyme was purified from a cell extract and characterised. It is not clear to what extent the peptidase contributes to the release of proline residues from protein substrates *in vivo*. No secreted prolidase activity could be detected in an *A. nidulans* culture expressing an extracellular type IV dipeptidyl aminopeptidase activity. In the absence of secreted prolidase activity, uptake of proline dipeptides for further utilisation is implied. It is, however, not known how dipeptides are imported into the fungal cell. In the yeast *S. cerevisiae* a highly regulated peptide transporter is found, which is specific for dipeptides (Turner et al., 2000). Whether this is also the case in filamentous fungi remains to be seen.

Acknowledgements

The authors thank Y. Müller for technical assistance. This work was supported by the Dutch Technology Foundation (STW) grant WBI .4100.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25: 3389-3402.**

Ballance, D. J., and Turner, G. (1985). Development of high frequency transforming vector of *Aspergillus niger*. *Gene* **36: 321-331.**

Beauvais, A., Monod, M., Wyniger, J., Debeaupuis, J. P., Grouzmann, E., Brakch, N., Svab, J., Hovanessian, A. G., and Latgé, J. P. (1997). Dipeptidyl-peptidase IV secreted by *Aspergillus fumigatus*, a fungus pathogenic to humans. *Infect. Immun.* **65: 3042-3047.**

Booth, M., Jennings, P. V., Ni Fhaolain, I., and O'Cuinn, G. (1990). Prolidase activity of *Lactococcus lactis* spp. *cremoris* AM2: purification and characterization. *J. Dairy Res.* **57: 245-254.**

Cunningham, D. F., and O'Connor B. (1997). Proline specific peptidases. *Biochim. Biophys. Acta* **1343: 160-186.**

de Graaff, L., van den Broek, H., and Visser, J. (1988). Isolation and expression of the *Aspergillus nidulans* pyruvate kinase gene. *Curr. Genet.* **13: 315-321.**

Doumas, A., van den Broek, P., Affolter, M., and Monod, M. (1998). Characterization of the prolyl dipeptidyl peptidase gene (*dppIV*) from the koji mold *Aspergillus oryzae*. *Appl. Environ. Microbiol.* **64: 4809-4815.**

Endo, F., Tanoue, A., Nakai, H., Hata, A., Indo, Y., Titani, K., and Matsuda, I. (1989). Primary structure and gene localization of human prolidase. *J. Biol. Chem.* **264: 4476-4481.**

Fernández-Esplá, M. D., Martín-Hernández, M. C., and Fox, P. F. (1997). Purification and characterization of a prolidase from *Lactobacillus casei* subsp. *casei* IFPL 731. *Appl. Environ. Microbiol.* **63: 314-316.**

Goosen, T., Bloemheuvel, G., Gysler, C., de Bie, D. A., van den Broek, H. W. J., and Swart, K. (1987). Transformation of *Aspergillus niger* using the homologous orotidine-5-phosphate decarboxylase gene. *Curr. Genet.* **11: 499-503.**

van den Hombergh, J. P. T. W., van de Vondervoort, P. J. I., Fraissinet-Tachet, L., and Visser, J. (1997). *Aspergillus* as a host for heterologous protein production: the problem of proteases. *Trends Biotechnol.* **15: 256-263.**

Ishii, T., Tsujino, S., Matsunobu, S., Endo, F., Sato, K., and Sakuragawa, N. (1996). Cloning of mouse prolidase cDNA: predominant expression of prolidase mRNA in kidney. *Biochim. Biophys. Acta* **1308: 15-16.**

Johnson, G. L., and Brown, J. L. (1974). Partial purification and characterization of two peptidases from *Neurospora crassa*. *Biochim. Biophys. Acta* **370: 530-540.**

Kusters-van Someren, M. A., Harmsen, J. A. M., Kester, H. C. M., and Visser, J. (1991). Structure of the *Aspergillus niger pel*A gene and its expression in *Aspergillus niger* and *Aspergillus nidulans*. *Curr. Genet.* **20: 293-299.**

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12: 357-358.**

Pontecorvo, G., Roper, J. A., Hemmons, L. J., MacDonald, K. D., Bufton, A. W. J. (1953). The genetics of *Aspergillus nidulans*. *Adv. Genet.* **5: 141-238.**

Roth, M. (1971). Fluorescence reactions for amino acids. *Anal. Chem.* **43: 880-882.**

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

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1994). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25: 4876-4882.**

Troll, W., and Lindsley, J. (1955). A photometric method for the determination of proline. *J. Biol. Chem.* **215: 655-660.**
Turner, G. C., Du, F., and Varshavsky, A. (2000). Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway. *Nature* **405: 579-583.**
Vanhoof, G., Goossens, F., De Meester, I., Hendriks, D., and Scharpé, S. (1995). Proline motifs in peptides and their biological processing. *FASEB J.* **9: 736-744.**
Wilce, M. C., Bond, C. S., Dixon, N. E., Freeman, H. C., Guss, J. M., Lilley, P. E., and Wilce, J. A. (1998). Structure and mechanism of a proline-specific aminopeptidase from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **95: 3472-3477.**
Yoshimoto, T., Tone, H., Honda, T., Osatomi, K., Kobayashi, R., and Tsuru, D. (1989). Sequencing and high expression of aminopeptidase P gene from *Escherichia coli* HB101. *J. Biochem.* **105: 412-416.**
Zimmerman, C. R., Orr, W. C., Leclerc, R. F., Barnard, E. C., Timberlake, W. E. (1980). Molecular cloning and selection of genes regulated in *Aspergillus* development. *Cell* **21: 709-715.**

Chapter 6

General discussion

Many filamentous fungi are well known for their saprophytic life style. They can degrade a large variety of different and often complex substrates because they produce a wide spectrum of enzymes. Several of these enzymes are produced at large quantities. For the production of all these enzymes the secretory pathway of filamentous fungi has to be very efficient. Nevertheless, our knowledge of the secretory pathway of filamentous fungi is still limited.

The secretory pathway of the yeast *Saccharomyces cerevisiae* functions as model for that from filamentous fungi. Until recently, the identity of the proteins and enzymes that are part of the secretory pathway of these fungi remained elusive. Orthologs of yeast chaperones, foldases and GTPases have now been identified in filamentous fungi (Conesa et al., 2002). However, at the start of our study the proteases responsible for processing of enzymes produced by filamentous fungi were still unknown. Research conducted on proteases secreted by filamentous fungi showed that they differ a lot with regard to proteases from yeast. For example the protease family Aspergillopepsin I (EC 3.4.23.18), to which the most active extracellular protease belongs, does not have any yeast members. Therefore, characterization of proteases that play a role in protein processing in the secretory pathway is highly desirable to arrive at a good model.

The research described in the previous chapters was performed to identify and characterize the proteases involved in the processing of secreted enzymes produced by the filamentous fungus *Aspergillus niger*. The strategy that we employed for the characterization of these proteases was to identify orthologs of proteases known to play a processing role in yeast. Because both KEX2 and DPAP A have been reported to play an important role in the processing of several natural yeast substrates, the study of *A. niger* orthologs of these two proteases became our main objective.

KEX2 is an endoprotease involved in removal of propeptides, and its activity was shown to be necessary for processing of α-factor and killer toxin. A similar proteolytic activity had been detected in both *Aspergillus nidulans* (Contreras et al., 1991) and *A. niger* (Broekhuijsen et al., 1993), using fusion constructs.

The yeast DPAP A protease is a Golgi located exoprotease which removes dipeptides from the N-terminus (Anna-Arriola and Herskowitz, 1994). It was shown to play a role in the processing of killer toxin, α-factor pheromone and KEX2. DPAP A recognizes dipeptidyl residues with either an alanine or proline residue at the penultimate position. DPAP B is a vacuolar homologue of DPAP A with unknown function. Both proteases are

dipeptidyl aminopeptidases of type IV (DPP IV). So far only one DPP IV had been identified in filamentous fungi. This protease, extracellular DPP IV, characterized in both *Aspergillus fumigatus* (Beauvais et al., 1997a, b) and *Aspergillus oryzae* (Doumas et al., 1998) is thought to play a role in substrate catabolism, not in protein processing.

We have successfully cloned the *kexB* gene and characterized KexB, the *A. niger* kexin protease (chapter 2). Using isolated membrane fractions, containing KexB, we showed that *in vitro* KexB recognizes dibasic cleavage sites (chapter 2) and using a secreted soluble KexB, constructed by genetic engineering, we showed that *in vitro* KexB also recognizes monobasic cleavage sites (chapter 3). We have demonstrated that a fusion protein consisting of glucoamylase, a linker peptide with a kexin cleavage site and human interleukin-6 produced in *A. niger* depends on the Golgi located endoprotease KexB for correct processing (chapter 2). Furthermore we have shown that KexB is solely responsible for the proteolytic processing of endopolygalacturonase II, a naturally produced enzyme of *A. niger* (chapter 3).

Disruption of the *kexB* gene resulted in an *A. niger* strain with a hyperbranching morphology at neutral and basic pH. We have shown that at acidic pH the deficient fungus grows like the wild-type strain (chapter 3). Mizutani et al. (2004) noted a similar hyperbranching morphology upon disruption of the *kexB* gene of *A. oryzae*. They were able to complement the *kexB* deficient morphology by growing the disrupted strain under conditions of high osmolarity. The hyperbranching *kexB* deficient strain had a significant higher expression of the *mpkA* gene. This gene encodes a putative mitogen-activated protein kinase involved in cell integrity signaling. Furthermore, or maybe as a result, several other genes encoding cell-wall biogenesis related enzymes were expressed at a higher level in the *kexB* deficient strain compared to wild-type.

It is still unknown how the *kexB* deficiency results in the hyperbranching morphology. In yeast it was demonstrated that the monobasic specific yapsin proteases are (at least partially) capable of complementing the *kex2* deficiency (Olsen et al., 1999). It might be that in a similar fashion the hyperbranching morphology is complemented by the action of an additional protease which is expressed at acidic pH in *A. niger*. We did not find evidence for the presence of yapsin-like proteases in *A. niger*. The presence of additional endoproteases involved in proteolytic processing remains an interesting topic for further research.

We have successfully cloned *dapB*, encoding an *A. niger* homologue of the Golgi located yeast DPAP A exoprotease (chapter 4). We could assign a putative transmembrane domain in the protein sequence of DapB. By constructing *dapB* overexpression strains we proved that the gene encodes a dipeptidyl aminopeptidase of type IV (DPP IV). We were able to measure a 5 to 7 fold increase of DPP IV activity in cell free extracts of the overexpression strains compared to a control strain with wild type DPP IV activity. Therefore, DapB is very likely an intracellularly located protease. We did not succeed in the purification of this protease. We could also not resolve whether DapB is an orthologue of the Golgi located yeast DPAP A protease or the vacuolar located yeast DPAP B protease. We do not know whether DapB plays a role in the proteolytic processing of secreted enzymes. It might well be that a processing protease like the yeast DPAP A does not exist in filamentous fungi. Both α-factor and killer toxin are not produced by *A. niger* and from our analysis of the amino acid sequence of KexB we could determine that there are no N-terminal dipeptides to be removed from the N-terminus of the *A. niger* kexin.

With the *A. nidulans* EST library becoming available we obtained a partial DNA sequence of an *A. nidulans* prolidase transcript. Prolidase is a proline specific dipeptidase, and its substrate resembles the dipeptide liberated by DPP IV proteases. Using the partial sequence as starting point we have successfully cloned the *A. nidulans* prolidase gene, designated *pepP*, and characterised the encoded product (Chapter 5). Using the purified dipeptidase we could determine that, like prolidases of other organisms, PepP cleaves dipeptides with a carboxy-terminal proline residue.

As a result of its unique cyclic structure proline residues introduce conformational constraints on the protein backbone (Vanhoof et al., 1995). These constraints imposed by proline residues render the protein less susceptible to proteolytic degradation. For example, many of the mammalian cytokines and growth factors have proline residues in the protein termini, making them less vulnerable for proteolysis. Because of their resistance against aspecific proteolysis, proline residues can play a role in regulation of enzyme activity.

The following proline specific protease have now been cloned and characterized in filamentous fungi: DapB (chapter 4), extracellular DPPIV, PepP (chapter 5), and prolyl aminopeptidase A (Basten, 2004). The identification and characterization of additional proline specific proteases

remains and interesting target. Information of their specificity simplifies the analysis of the effect of proline residues on protease susceptibility.

In summary, our aim was to clone the genes encoding the *Aspergillus* orthologs of the yeast KEX2 endoprotease and DPAP A exoprotease, both located in the Golgi, and characterize both *Aspergillus* proteases. We successfully cloned the *kexB* gene, and characterized KexB, the *A. niger* ortholog of yeast KEX2. While yeast KEX2 has specificity for dibasic substrates we found that *A. niger* KexB can process both dibasic and monobasic substrates. We have cloned *dapB*, encoding an *A. niger* homolog of the DPAPA protease, however we do not know if it is also its orthologue. What we now do know is that in yeast two dipeptidyl aminopeptidase exist, DPAPA and DPAP B. DPAP A is located in the Golgi and plays an important role in proteolytic processing, while DPAP B is located in the vacuole and its function is unknown. In filamentous fungi also two dipeptidyl aminopeptidases exist, however only one of these protease (DapB) has a putative transmembrane domain and might be involved in protein processing, the other dipeptidyl aminopeptidase is located extracellular. In addition we have cloned *pepP* and characterized the *A. nidulans* prolidase, a proline specific dipeptidase. Our study presents an important contribution to the elucidation of processing proteases that are active in the secretory pathway of industrially relevant filamentous fungi.

References

Anna-Arriola, S. S., and Herskowitz, I. (1994). Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase. *Yeast* **10: 801-810.**

Basten, D. E. J. W. (2004). Aminopeptidases from *Aspergillus niger. Thesis.* Wageningen University, Wageningen, The Netherlands.

Beauvais, A., Monod, M., Debeaupuis, J. P., Diaquin, M., Kobayashi, H., and Latgé, J. P. (1997a) Biochemical and antigenic characterization of a new dipeptidyl-peptidase isolated from *Aspergillus fumigatus*. *J. Biol. Chem.* **272: 6238-6244.**

Beauvais, A., Monod, M., Wyniger, J., Debeaupuis, J. P., Grouzmann, E., Brakch, N., Svab, J., Hovanessian, A. G., and Latgé, J. P. (1997b). Dipeptidyl-peptidase IV secreted by *Aspergillus fumigatus*, a fungus pathogenic to humans. *Infect. Immun.* **65: 3042-3047.**

Broekhuijsen, M. P., Mattern, I. E., Contreras, R., Kinghorn, J. R., and van den Hondel, C. A. (1993). Secretion of heterologous proteins by *Aspergillus niger*: production of active human interleukin-6 in a protease-deficient mutant by KEX2-like processing of a glucoamylase-hIL6 fusion protein. *J. Biotechnol.* **31: 135-145.**

Conesa, A., Punt, P. J., van Luijk, N., and van den Hondel, C. A. M. J. J. (2001). The secretion pathway in filamentous fungi: A biotechnological view. *Fungal Genet. Biol.* **33: 155-171.**

Contreras, R., Carrez, D., Kinghorn, J. R., van den Hondel, C. A., and Fiers, W. (1991). Efficient KEX2-like processing of a glucoamylase-interleukin-6 fusion protein by *Aspergillus nidulans* and secretion of mature interleukin-6. *Biotechnology* **9: 378-381.**

Doumas, A., van den Broek, P. Affolter, M., and Monod. M. (1998). Characterization of the prolyl dipeptidyl peptidase gene (*dpplV*) from the koji mold *Aspergillus oryzae*. *Appl. Environ. Microbiol.* **64:4809-4815.**

Mitzutani, O., Nojima, A., Yamamoto, M., Furukawa, K., Fujioka, T., Yamagata, Y., Abe, K., and Nakajima, T. (2004). Disordered Cell Integrity Signaling Caused by Disruption of the *kexB* Gene in *Aspergillus oryzae*. *Eukaryotic Cell* **3: 1036-1048.**

Olsen, V., Cawley, N. X., Brandt, J.. Egel-Mitani, M., and Loh, Y. P. (1999). Identification and characterization of *Saccharomyces cerevisiae* yapsin 3, a new member of the yapsin family of aspartic proteases encoded by the *YPS3* gene. *Biochem. J.* **339: 407-411.**

Vanhoof, G., Goossens, F., de Meester, I., Hendriks, D., and Scharpé, S. (1995). Proline motifs in peptides and their biological processing. *FASEB J.* **9: 736-744.**

Summary

Samenvatting

Summary

A number of filamentous fungi are saprophytes and they secrete a wide spectrum of enzymes to degrade their complex substrates. Many secreted proteins enter the secretory pathway as proproteins and need some form of proteolytic processing before they obtain their mature active state. As described in this thesis we have cloned the genes encoding two proteases involved in such proteolytic processing events occurring in the secretory pathway of *Aspergillus niger*. In addition, we have characterized both proteases.

We have cloned *kexB* and characterized the kexin maturase. KexB matures by hydrolysation of propeptides from proproteins and plays an important role in the processing of proteins produced in the secretory pathway of *A. niger*. We found that *A. niger* strains with a disrupted *kexB* gene have a hyper-branching morphology at pH 6 (Chapter 3). This morphology is pH dependent, because at pH 4 there are no apparent morphological differences between the *kexB* disrupted strain and wild-type *A. niger*. KexB is a modular protein (Chapter 2). In addition to a subtilisin-like protease domain, which in yeast functions in the late Golgi, KexB contains a transmembrane domain and putative cytoplasmic Golgi retention signal. To enable a biochemical characterization of KexB a truncated soluble secreted form of the protease was constructed, by removal of the C-terminal transmembrane domain and Golgi retention signal. (Chapter 3). Using fluorogenic substrates we were able to show that this soluble form of KexB can efficiently process both dibasic and monobasic amino acid cleavage sites. This is in remarkable contrast to Kex2, the KexB homolog of *Saccharomyces cerevisiae*, which only cleaves dibasic amino acid cleavage sites efficiently. Furthermore we have demonstrated that KexB is responsible for the *in vivo* processing of endopolygalacturonase I, and is a likely candidate for processing of endopolygalacturonase II. Both enzymes are naturally secreted by *A. niger* and enter the secretory pathway as proproteins.

We have also cloned *dapB* and characterized *A. niger* DapB, a possible ortholog of *S. cerevisiae* DPAP A (Chapter 4). DPAP A is a Golgi located exoprotease, which plays a role in the processing and activation of secreted proteins. DPAP A is a type IV dipeptidyl aminopeptidase (DPP IV). Phylogenetic analysis showed that DapB is part of a novel subfamily of DPP IV proteases in filamentous fungi with a transmembrane domain. DPP

IV proteases remove dipeptides at the N-terminus with an alanine or proline residue at the penultimate position. We were able to confirm the DPP IV-like activity of DapB by overexpression of the protease, resulting in 5 to 7 times higher DPPIV like activity on artificial substrates in cell extracts from DapB transformants. Although we could predict the presence of a transmembrane domain, we were not able to determine the location of the mature protease. Currently we do not know whether DapB plays a role in protein maturation.

DapB is a proline specific protease. Currently not much is known about the role of proline residues in the protection and regulation (delay of activation) of enzymes produced by filamentous fungi. However in other organisms proline is often found at the termini of secreted enzymes to protect against exoprotease activity. To make a study into the role of proline in protein degradation and regulation feasible we have started with the identification and characterization of proline specific proteases

In addition to the *A. niger kexB* and *dapB* genes, we have cloned *pepP* and characterized the proline specific dipeptidase (PepP) of *Aspergilllus nidulans* (Chapter 5). Publicly available EST sequences of *A. nidulans* led to the identification and cloning of the prolidase encoding gene of this fungus. A prolidase overexpressing *A. nidulans* strain was constructed. The protease purified from such an overexpressing strain was used for a biochemical characterization of the protein. The protein is a manganese^{2+} dependent metallo protease and *in vitro* cleavage studies showed that it only hydrolyses dipeptides with a C-terminal proline residue. DapB, the *A. niger* DPP IV protease releases such dipeptides from the N-terminus of a polypeptide chain.

In summary, we have cloned the *kexB* gene and characterized the *A. niger* endoprotease KexB. KexB removes the propeptide of proteins traveling through the secretory pathway. In addition, we have cloned *dapB* and characterized the *A. niger* exoprotease DapB, a possible ortholog of *S. cerevisiae* DPAP A. DPAP A also plays a role in the maturation of proteins. Furthermore, the *A. nidulans* prolidase gene has been cloned and PepP characterized. PepP is a proline specific dipeptidase, its substrates are dipeptides like those released by DapB. With our research we have gained insight into proteolytic processing in the secretory pathway of *Aspergillus niger*, and more specifically we have obtained detailed knowledge about propeptide processing of proproteins destined for secretion.

Samenvatting

Verscheidene filamenteuze schimmelsoorten behoren tot de saprofieten. Ze secreteren een breed spectrum aan enzymen om hun complexe substraten af te breken. Van de te secreteren eiwitten zijn er veel die de secretieroute binnen komen als pro-eiwit en vervolgens een vorm van proteolytische bewerking nodig hebben voordat ze hun mature actieve vorm bereiken. Zoals beschreven in dit proefschrift hebben we de genen gekloneerd van twee proteases die een rol spelen bij deze proteolytische bewerking in de secretieroute van *Aspergillus niger*, en deze proteases verder gekarakteriseerd.

We hebben het *kexB* gen gekloneerd en de kexin maturase gekarakteriseerd. KexB matureert pro-eiwitten door de propeptide te scheiden middels hydrolysatie en speelt een belangrijke rol bij de processing van eiwitten die geproduceerd worden in de secretieroute van *A. niger*. We hebben ontdekt dat *A. niger* stammen met een uitgeschakeld *kexB* gen een hypervertakkende morfologie hebben bij pH 6 (Hoofdstuk 3). Deze morfologie is pH afhankelijk, omdat er bij pH 4 geen duidelijke morfologische verschillen zijn tussen de stam met het uitgeschakelde *kexB* gen en de wildtype *A. niger* stam. KexB is een modulair eiwit (Hoofdstuk 2). Naast een subtilase gelijkend protease domein, dat in gist zijn functie uitvoert in het Golgi, bevat KexB een transmembraan domein en een verondersteld cytoplasmatisch Golgi retentie signaal. Om een biochemische karakterisatie van KexB mogelijk te maken werd een verkorte oplosbare en gesecreteerde vorm van de protease geconstrueerd door het C-terminale transmembraan domein en het Golgi retentie signaal te verwijderen (Hoofdstuk 3). Met behulp van fluorogene substraten konden we aantonen dat deze oplosbare vorm van KexB op efficiente wijze zowel dibasische en monobasische motieven kan klieven. Dit staat in schril contrast tot de specificiteit van Kex2, the KexB homoloog van *Saccharomyces cerevisiae*, die alleen dibasische aminozuur motieven efficient kan klieven. We hebben daarnaast laten zien dat KexB verantwoordelijk is voor de *in vivo* processing van endopolygalacturonase I, en waarschijnlijk verantwoordelijk is voor de processing van endopolygalacturonase II. Beide enzymen worden van nature gesecreteerd door Aspergillus niger, en komen de secretieroute binnen als pro-eiwitten.

We hebben ook het gen voor *A. niger* DapB, een mogelijke ortholoog van S. cerevisiae DPAP A (hoofdstuk 4), gekloneerd en de

protease gekarakteriseerd. DPAP A is een Golgi gelocaliseerde exoprotease, die een rol speelt in de processing en activatie van gesecreteerde eiwitten. DPAP A is een type 4 dipeptidyl aminopeptidase (DPP IV). Uit phylogenetische analyse bleek dat DapB deel uitmaakt van een nog niet eerder beschreven subfamile van DPP IV proteases in filamenteuze schimmels met een transmembraan domein. DPP IV proteases verwijderen dipeptides van de N-terminus met een alanine of proline residue als voorlaatste positie. We konden de DPP IV aktiviteit van DapB bevestigen door overexpressie van de protease, wat resulteerde in 5 tot 7 keer hogere DPPIV aktiviteit op artificiele substraten in cel extracten van de DapB transformanten.Hoewel we de aanwezigheid van een trans-membraan domein konden voorspellen, waren we niet in staat om de locatie van het mature eiwit te bepalen. We weten momenteel niet of DapB een rol speelt bij eiwit maturatie.

DapB is een proline specifieke protease. Momenteel is er niet veel bekend over de rol van proline residuen in de bescherming en regulatie (vertragingn van activatie) van enzymen die geproduceerd worden door filamenteuze schimmels. Echter in andere organismen wordt proline vaak gevonden aan de uiteinden van gesecreteerde enzymen om ze te beschermen voor exoprotease activiteit. Om onderzoek naar de rol van proline in eiwit degradatie en regulatie mogelijk te maken zijn we gestart met de identificatie en karakterisatie van proline specifieke proteases.

Naast *A. niger kexB* en *dapB*, hebben we het gen dat codeert voor de proline specifieke dipeptidase (PepP) van *Aspergillus nidulans* gekloneerd en de protease gekarakteriseerd (Hoofdstuk 5). Algemeen beschikbare EST sequenties van *A. nidulans* leidden tot de identificatie en klonering van het prolidase coderende gen van deze schimmel. Vervolgens is er een prolidase overproducerende *A. nidulans* stam gemaakt. PepP gezuiverd van deze overproducerende stam is vervolgens gebruikt voor een biochemische karakterisatie van de protease. PepP is een mangaan afhankelijke metalloprotease, en *in vitro* processing experimenten hebben aangetoont dat het alleen dipeptides hydroliseert met een C-terminale proline residue. DapB, de *A. niger* DPPIV protease haalt zulke dipeptides van de N-terminus van eiwitketens af.

Samengevat, hebben we A. niger *kexB* gekloneerd en de KexB protease gekarakteriseerd. KexB verwijderd de propeptide van eiwitten die door de secretieroute reizen. Daarnaast hebben we *A. niger dapB* gekloneerd, en DapB, een mogelijke ortholoog van S. cerevisiae DPAP A, gekarakteriseerd. DPAP A speelt ook een rol in de maturatie van eiwitten. Tevens hebben we *A. nidulans pepP* gekloneerd. De proline specifieke

dipeptidase PepP hebben we nader gekarakteriseerd. Zijn substaten zijn dipeptides zoals die door DapB vrij gemaakt worden. Met ons onderzoek hebben we inzicht verworven in de proteolytische processing in de secretieroute van *A. niger*, en meer specifiek hebben we gedetaileerd kennis verworven over propeptide processing van pro-eiwitten op weg naar hun secretie.

Curriculum vitae

Ruud Jalving werd op 31 december 1972 geboren te Buinen (Drenthe). Hij behaalde in 1990 zijn HAVO diploma aan het Ubbo Emmius Lyceum in Stadskanaal. Twee jaar later voltooide hij bij het Ubbo Emmius Lyceum de bovenbouw van het VWO. In 1992 begon hij zijn studie Biologie aan de Rijks*universiteit* van Groningen. Deze studie sloot hij in 1997 succesvol af met het behalen van zijn diploma. Daarna begon hij halverwege juli 1997 als onderzoeker in opleiding bij de sectie moleculaire genetica van industriële micro-organismen van de Wageningen Universiteit. De resultaten van dat onderzoek staan beschreven in dit proefschrift. Op 1 september 2002 begon hij als postdoc bij de faculteit Diergeneeskunde van de Universiteit van Utrecht. Dat project was erop gericht om de genetische variatie in het kippengenoom in kaart te brengen en liep tot eind december 2004.

Publications

R. Jalving, P. J. I. van de Vondervoort, J. Visser and P.J. Schaap. (2000). Characterization of the kexin-like maturase of *Aspergillus niger*. Appl. Environ. Microbiol. **66:** 363-368.

R. Jalving, P. Bron, H. C. M. Kester, J. Visser and P. J. Schaap. (2002). Cloning of the prolidase gene from *Aspergillus nidulans* and characterisation of its product. Mol. Gen. Genom. **267:** 218-222.

L. G. Lugones, J. F. de Jong, O. M. H. de Vries, R. Jalving, J. Dijksterhuis and H. A. B. Wösten. (2004). The SC15 protein of *Schizophyllum commune* mediates formation of aerial hyphae and attachment in the absence of the SC3 hydrophobin. Mol. Microbiol. **53:** 707-716.

R. Jalving, R. van 't Slot and B. A. van Oost. (2004). Chicken single nucleotide polymorphism identification and selection for genetic mapping. Poult. Sci. **83:** 1925-1931.

R. Jalving, J. Godefrooij, W. J. ter Veen, A. J. J. van Ooyen and P. J. Schaap. Characterisation of *Aspergillus niger dapB*, a novel fungal type IV dipeptidyl aminopeptidase. **Submitted for publication.**

A. Mete, R. Jalving, B. A. van Oost, J. E. van Dijk, and J. J. M. Marx. Intestinal over-expression of iron transporters induces iron overload in birds in captivity. **Submitted for publication.**

A. C. Wiersma, R. Jalving and B. A. van Oost. Identification of single nucleotide polymorphisms and haplotype construction in the canine *COL4A3* and *COL4A4* genes. **Submitted for publication**

Nawoord

Nu het boekje eindelijk voltooid is wil ik toch even stil staan bij iedereen die mij geholpen en gesteund heeft gedurende de periode dat ik aan mijn promotieonderzoek en aan mijn proefschrift heb gewerkt.

Als eerste wil ik een woord van dank richten tot mijn promotor en co-promotor. Ab, aan jouw was de ondankbare taak weggelegd om een serie ontredderde AIO's door hun laatste periode heen te helpen. Ik wil jouw bedanken voor het feit dat je als promotor op wilde treden nadat ik in mijn vierde jaar geen promotor en co-promotor meer had. Ik heb je inbreng zeer gewaardeerd. Peter, ik wil jouw bedanken voor jouw bijdrage als dagelijkse begeleider en de laatste jaren als co-promotor. Jouw verfrissende ideeën en kritische kanttekeningen waren erg belangrijk voor mij en mijn onderzoek.

Bernard, ik wil jou bedanken voor jouw geduld en begrip de afgelopen twee jaar. Daardoor kon mijn proefschrift ook daadwerkelijk tot een goed einde gebracht worden.

Ik wil mijn labgenoten van lab 1004 bedanken voor de gezellige werksfeer en de onderlinge solidariteit die er altijd was. Vooral met de vaste medebewoners van het lab Danielle, Yvonne, en later ook Kim was het erg prettig samenwerken.

Daarnaast wil ik graag al mijn collega's bij MGIM bedanken voor de steun waar ik steeds op kon rekenen. In het bijzonder wil ik Peter van de Vondervoort en Harry Kester bedanken voor hun bijdrage aan dit boekje. Bjarne en Marco, het mensahappen bij Unitas was een vaste en zeer prettige traditie tijdens drukke dagen.

Behalve werk was er natuurlijk de vereniging Marijkeweg. De vaste groep bezoekers van de Marijkebar was een zeer hechte, vriendelijke en loyale groep. Zij maakten de bar tot de meest gezellige huiskamer in Nederland. Ik heb er vele vrienden en kennissen opgedaan. Door jullie is mijn verblijf in Wageningen een fantastische ervaring geworden.

Ik wil mijn famillie bedanken voor de aandacht die ze de afgelopen jaren getoond hebben voor het verloop van mijn onderzoek. Mijn ouders wil ik bedanken voor hun wijze raad in moeilijke tijden en hun niet aflatende steun en waardering.

Margot, jij was de afgelopen jaren een onmisbare steun voor mij. Graag wil ik jouw bedanken voor de getoonde begrip in de moeilijke momenten. Nu het boekje af is kunnen we samen leuke dingen gaan doen.

The research described in this thesis was carried out at the department Microbiology, section Fungal Genomics, formerly known as the section Molecular Genetics of Industrial Micro-organisms, Wageningen University. The project was funded by the technology foundation (STW) of the Netherlands Organisation for Scientific Research (NWO).