A functional genomics study of extracellular protease production by Aspergillus niger

Braaksma, M.

CHAPTER 1

ASPERGILLUS AS A CELL FACTORY FOR PROTEIN PRODUCTION: CONTROLLING PROTEASE ACTIVITY IN FUNGAL PRODUCTION

Machtelt Braaksma and Peter J. Punt

This chapter has been published with minor modifications in:
The *Aspergilli*: Genomics, Medical Aspects, Biotechnology, and Research Methods (2008), pp. 441-455. Edited by G. H. Goldman & S. A. Osmani. Boca Raton, FL: CRC Press

INTRODUCTION

Since ancient times micro-organisms have been used in a variety of traditional food processes (e.g., the production of alcoholic beverages, cheese, and bread). Fungi are applied in cheese-making and in traditionally oriental food such as soy-sauce, tempeh, and sake. However, the presence and role of these micro-organisms was for most processes only identified in recent times. Fungi, like *Aspergillus oryzae* in the production of sake, were discovered to play a key role in food production by the excretion of enzymes. In 1894, the first microbial enzyme that was commercial produced appeared on the market, called "takadiastase"; it was in fact fungal amylase produced by *A. oryzae* (Gwynne & Devchand, 1992). Nowadays, a large number of fungal enzymes are commercially available and their application extends well beyond their traditional use in food processes. Glucoamylase, α-amylases, cellulase, lipase, and protease are only a few examples of enzymes produced by filamentous fungi that are commercially available. *Aspergillus* species, and particularly *A. niger* and *A. oryzae*, play a dominant role in the production of many of these enzymes (for a list of commercial enzymes see the Association of Manufacturers and Formulators of Enzyme Products (AMFEP)[1]).

For the last two decades, filamentous fungi have also been explored as hosts for the production of heterologous proteins. Because of their established use as production host of homologous proteins aspergilli are the obvious expression system for heterologous proteins. The Danish company Novozymes A/S was in 1988 the first on the market with a non-native fungal lipase (Lipolase) produced from a genetically modified *A. oryzae* strain (Nevalainen & Te'-o, 2003). Since then, several species of *Aspergillus* have been used to express a wide variety of foreign genes (see also the list of commercial enzymes of the AMFEP). However, the production of heterologous as well as homologous proteins is often limited by the high levels of proteases also produced by the fungal host organism. This review will focus on the role of protease activity in strain and process development. Both classical mutagenesis and gene disruption techniques have been applied to generate strains with reduced protease activity. And indeed, production levels improved significantly when using protease deficient strains (e.g., tissue plasminogen activator (t-PA) production with a protease deficient *A. niger* strain [Wiebe *et al.*, 2001]). Controlling the culture conditions can result in a further improvement of the heterologous protein production (e.g., green fluorescent protein (GFP) production with a protease deficient *A. niger* strain at controlled pH [O'Donnell *et al.*, 2001]). However, the production levels for

[1] http://www.amfep.org/list.html; August 24, 2010

heterologous proteins are in most cases one to two orders of magnitude lower than for homologous proteins.

With the availability of the complete genome sequence of several *Aspergillus* strains (e.g., *A. flavus*[2]; *A. fumigatus*[3] [Nierman *et al.*, 2005]; *A. niger*[4]; *A. oryzae*[5] [Machida *et al.*, 2005]; *A. nidulans*[6] [Galagan *et al.*, 2005] and *A. terreus*[6]), homology searches for genes involved in the proteolytic systems of these organisms resulted in a much higher number of genes encoding protease activity than previously known. For example, for *A. niger* approximately 200 genes involved in proteolytic degradation were found in the genome (Pel *et al*., 2007). In comparison, before the genome sequence of *A. niger* was known, an extensive analysis of the proteolytic system of *A. niger* led to the identification of only eight protease genes (van den Hombergh, 1996). Given this very large gene potential, actual protease production and its regulation is expected to be very complicated.

The understanding of the regulation of the proteolytic system of *Aspergillus* strains is still only in its infancy. The involvement of several wide-domain regulatory systems (carbon catabolite repression, nitrogen metabolite repression, pH regulation [van den Hombergh, 1996]) and probably sulfur metabolite repression (VanKuyk *et al.*, 2000) in the overall regulation of protease expression in *Aspergillus* is suggested. This review gives state of the art in the protease research field and provides an outlook on new research approaches.

STRAIN DEVELOPMENT

Classical methods to screen for protease mutants

Mutagenesis by means of X-ray or UV irradiation and chemicals mutagenesis were discovered in the first half of the past century. Hara *et al.* (1992) describe the successful attempts of Iguchi (1955-1956) to isolate a mutant strain producing higher levels of protease compared to the parent strain. After X-ray irradiation a large number of isolates were screened in a laborious and time-consuming effort for a hyperproducing mutant. The screening procedure was greatly improved by the

[2] http://www.aspergillusflavus.org/genomics/; August 25, 2010

[3] http://www.sanger.ac.uk/projects/A_fumigatus/; August 25, 2010

[4] http://genome.jgi-psf.org/Aspni5/Aspni5.home.html; August 24, 2010

[5] http://www.bio.nite.go.jp/dogan/project/view/AO; August 25, 2010

[6] http://www.broadinstitute.org/science/projects/fungal-genome-initiative; August 25, 2010

method developed by Sekine in 1969 which enabled the screening of a large number of isolates (Hara *et al.*, 1992). Around colonies grown on casein-containing medium a halo (clear zone) was formed of which the diameter has a significant correlation with the protease production (see Fig. 1).

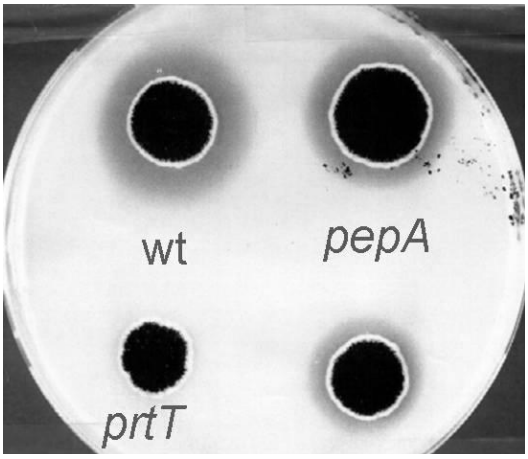

Fig. 1. Protease-deficient mutants (*pepA* and *prtT*) of *Aspergillus niger* show reduced degradation of casein compared to the wild type strain (WT).

These classical methods to generate and screen for mutants with altered levels of excreted protease are still successfully applied. Nowadays, mutagenesis of spores is most often conducted with ultraviolet light irradiation, which is the less-aggressive than irradiation with X-rays. This approach has been applied to isolate several protease-deficient mutants in different aspergilli, such as *A. niger* (Mattern *et al.*, 1992; van den Hombergh *et al.*, 1995) and *A. nidulans* (Katz *et al.*, 1996). Also mutagenesis with mutagens such as nitrosoguanidine has been described (Kolattukudy *et al.*, 1993; Moralejo *et al.*, 2000). After mutagenesis the spores are plated on milk or gelatin-casein medium. Mutants with low proteolytic activity are screened for reduced degradation of casein which results in a reduced or no halo on those plates. In this way, Mattern *et al.* (1992) isolated *A. niger* mutants with residual extracellular proteolytic activities that vary from 2% to 80% of the protease activity of the parental strain. Katz *et al.* (1996) describe *A. nidulans* mutants with tenfold reduced levels of extracellular protease compared to the parental strain.

Molecular genetic methods to construct protease mutants

Protease genes

Clearly, the random mutagenesis approach results in potent production hosts, but the genetic basis of these mutants remains unknown and may have unwanted pleiotropic effects on fungal fermentation performance (e.g., gene expression, growth rate). Therefore, with the development of molecular genetic tools also a more targeted approach to obtain protease-deficient mutants became available.

The general strategy for this approach is the so-called reverse genetics. By separating proteins produced in culture medium by SDS-PAGE or chromatography and subsequently testing for protease activity (as determined, e.g., by protease activity on skim milk agarose) of the different bands or fractions several proteases can be identified. By determining the (partial) amino acids sequence the protein oligonucleotide probes corresponding to these sequences can be designed. These oligonucleotides or PCR fragments generated by using similar oligonucleotides are subsequently used to screen genomic libraries to clone the corresponding protease genes. With the resulting clones a disruption vector for the protease gene can be constructed for actual gene disruption. The more recent availability of genome databases makes it also possible to use obtained amino acid sequences directly to clone the corresponding genes by genome mining using sequence comparison algorithms such as BLASTX. However, even with knowledge of the genome sequence, an activity screen (most preferably based on proteolytic activity against the protein one wants to produce) is still necessary to identify which of all the protease genes present in the fungal genome is actually new and most active and thus the desired target for gene disruption. Berka *et al.* (1990) were the first to describe the construction of gene replacement vectors for *Aspergillus*, which were used to specifically delete the chromosomal DNA of the protease gene encoding the major extracellular acid protease aspergillopepsin A (PEPA) in *A. awamori*. Disruption of this *pepA* gene reduced extracellular proteolytic activity compared to the wild type. Similar results were achieved by disruption of the aspergillopepsin A gene in *A. niger* (Mattern *et al.*, 1992). Probes containing part of the coding region of this *pepA* gene were also used to screen the genomic library of an *A. nidulans* strain (VanKuyk *et al.*, 2000). And although *A. nidulans* appears to lack detectable acid protease activity, a clone which hybridized with the *pepA* gene was obtained. This aspartic protease gene, which was designated *prtB*, was only expressed at a very low level. Furthermore, homologues of the *pepA* gene have been cloned from other *Aspergillus* species, such as *A. fumigatus* (Lee & Kolattukudy, 1995), *A. oryzae* (Berka *et al.*, 1993) and *A. satoi* (Shintani & Ichishima, 1994).

In non-acid producing aspergilli, such as *A. nidulans*, neutral or alkaline proteases are responsible for the major part of the extracellular protease activity. Disruption of the gene coding for the dominant extracellular serine protease in *A. nidulans* strain resulted, when cultured under various medium limitations, in reduced levels of proteolytic activity under all culture conditions (VanKuyk *et al.*, 2000). Controlled batch fermentations of an *A. sojae* strain with a disruption of an alkaline protease gene resulted in about 40% reduction of proteolytic activity in comparison to the wild type (Heerikhuisen *et al.*, 2005). Shake flasks cultures with *A. oryzae* expressing the heterologous protein endoglucanase showed enhanced stability of this protein when an alkaline protease gene of the host strain was disrupted (Lehmbeck, 2001).

Not in all cases disruption of a protease gene results in decreased protease activity. Disruption of the serine protease gene (*sep*) in *A. flavus* led to a compensatory increase in the expression and production of a metalloproteinase gene (*mep20*) (Ramesh & Kolattukudy, 1996). Both wild type and mutant degraded elastin at the same rate. The authors concluded that the expression of the genes encoding both proteases is controlled by a common regulatory system and that the fungus has a mechanism to sense the status of the extracellular proteolytic activities.

An alternative method for reduction of expression of a particular gene is the use of antisense RNA. This approach was applied in an *A. awamori* strain used to express the heterologous protein thaumatin (Moralejo *et al.*, 2002). Even though an insertion in the *pepA* gene had resulted in an inactive PEPA protein, thaumatin was still degraded. Another protease, aspergillopepsin B (PEPB; previously believed to be a pepstatin-insensitive aspartyl protease, but more recently established to be a member of the newly discovered family of glutamic proteases [Fujinaga *et al.*, 2004]), was identified as the most likely protease responsible for this degradation. Expression of *pepB* antisense RNA improved thaumatin production with 30%. Nevertheless, thaumatin was still degraded, indicating the antisense mRNA had only a partial silencing effect on *pepB* gene expression. Disruption of the *pepB* gene resulted in a significant further increase of the thaumatin production. However, an advantage of gene silencing with respect to gene disruption is that it can be used to suppress the expression of complete gene families. Zheng *et al.* (1998) describe that the expression of antisense RNA of the structural gene of carboxipeptidase in *A. oryzae* did not only decrease the activity of that carboypeptidase, but also of two other carboypeptidases.

Yet another approach to obtain strains with low protease levels is disruption of proteases that proteolytically activate other protease precursor proteins which require processing for their activation. Disruption of such a protease gene will have a

direct effect on the protease activity of one or more other proteases, as was described for *A. niger*. Disruption of the gene of an intracellular acid protease (PEPE) in *A. niger* did not only reduce the intracellular pepstatin-inhibitable aspartyl protease activity, but also intracellular serine protease and serine carboxypeptidase activities were significantly reduced in the Δ*pepE* strain (van den Hombergh *et al.*, 1997a). The transcription of these non-disrupted genes was not affected by the disruption of the single *pepE* gene. According to the authors this may indicate the presence of a cascade activation mechanism for several vacuolar proteases, triggered by the PEPE protein. A similar mechanism has been described for *Saccharomyces cerevisiae* (van den Hazel *et al.*, 1996).

In Table 1 described disruptions of protease genes in *Aspergillus* strains and the resulting residual proteolytic activities are summarized. In this table the construction of multiple disruptants can lead to further decrease of proteolytic activities. This was shown for a Δ*pepA*Δ*pepB*Δ*pepE* triple disruptant in *A. niger* (van den Hombergh *et al.*, 1997a) and disruption in *A. fumigatus* of both a gene encoding an extracellular serine alkaline protease and a gene encoding an extracellular metalloprotease (Jaton-Ogay *et al.*, 1994).

Protease regulators

Finally, a very efficient approach to generate strains with low protease levels is through disruption of genes that influence the expression of multiple protease genes. Two groups of regulatory genes have been described so far. In the first place, genes that encode specific regulators of protease genes; second, genes that encode wide-domain regulators. Interestingly, in the first group, to date, only one single gene has been identified both in fungi and yeast species. This gene is the *prtT* gene, as cloned from an UV-induced *A. niger* mutant (Punt *et al.*, 2008). This mutant was suggested to be a regulatory mutant as at least two proteases, including PEPA, were missing from the culture medium, while genetic data indicated the presence of a single semi-dominant mutation, not linked to the *pepA* gene (Mattern *et al.*, 1992). Recent analysis has indeed shown that the *prtT* gene is actually a regulatory gene encoding a member of the Zn-binuclear cluster family (Punt *et al.*, 2008). Interestingly, this gene is unique for *Aspergillus* species but actually absent in *A. nidulans*. With the disruption of the *prtT* gene in *A. niger* total protease activity was reduced to 20% of the wild type (Connelly & Brody, 2004).

Table 1. Effects on secreted protease activity of protease gene disruption strains in aspergilli

Species	Name disrupted gene	Residual extracellular protease activity *	Reference
	Extracellular serine protease (fam. S8)		
A. flavus	*sep*	100%	Ramesh *et al.*, 1996
A. fumigatus	*alp*	0-30%	Tang *et al.*, 1992; Monod *et al.*, 1993; Jaton-Ogay *et al.*, 1994
A. nidulans	*prtA*	10-50%	VanKuyk *et al.*, 2000
A. oryzae	*alp*	< WT	Lehmbeck, 2001
A. sojae	*alpA*	60%	Heerikhuisen *et al.*, 2005
	Vacuolar serine protease (fam. S8)		
A. oryzae	*pepC*	N/A	Christensen & Lehmbeck, 2000
	Extracellular aspartyl protease (fam. A1)		
A. awamori	*pepA*	<< WT	Berka *et al.*, 1990
A. fumigatus	*pep*	<< WT	Reichard *et al.*, 1997
A. niger	*pepA*	15-20%	Mattern *et al.*, 1992; van den Hombergh *et al.*, 1997a
	Vacuolar aspartyl protease (fam. A1)		
A. niger	*pepE*	~100%	van den Hombergh *et al.*, 1997a
A. oryzae	*pepE*	N/A	Christensen *et al.*, 2000
	Extracellular glutamic protease (fam. G1)		
A. awamori	*pepB*	< parent †	Moralejo *et al.*, 2002
A. niger	*pepB*	95%	van den Hombergh *et al.*, 1997a
	Extracellular metallo protease (fam. M35)		
A. nidulans	*pepI*	N/A	van den Hombergh & Visser, 1997b
	pepJ	N/A	van den Hombergh *et al.*, 1997b
A. oryzae	*npII*	< WT	Lehmbeck, 1999
	Extracellular metallo protease (fam. M36)		
A. fumigatus	*mep*	70%	Jaton-Ogay *et al.*, 1994
A. niger	*pepH*	< WT	van den Hombergh *et al.*, 1997b
A. oryzae	*npI*	N/A	Lehmbeck, 1999
	Multiple disruptants		
A. fumigatus	*alp, mep*	<< WT	Jaton-Ogay *et al.*, 1994
A. niger	*pepA, pepB*	10%	van den Hombergh *et al.*, 1997a
	pepA, pepE	~ Δ*pepA*	van den Hombergh *et al.*, 1997a
	pepB, pepE	~ Δ*pepB*	van den Hombergh *et al.*, 1997a
	pepA, pepB, pepE	<10%	van den Hombergh *et al.*, 1997a

* As determined with protease assays and expressed as percentage compared to the parent strain; N/A is data not available

† Parent strain is not the WT strain, but a classical *pepA*-deficient mutant

Besides regulatory genes specific for protease expression, wide-domain regulatory genes affect the expression of a broad spectrum of enzymes, including proteases, as a response to ambient pH (*pacC* gene), nitrogen source (*areA* gene) or carbon source (*creA* gene).

The *pacC* gene is expressed at alkaline pH and encodes a protein, which is able to activate the expression of other alkali-expressed genes and to prevent the expression of acid-expressed genes (Peñalva & Arst, Jr., 2002). In *A. nidulans* the expression of the major alkaline protease *prtA* gene is activated by PacC. However, disruption of the *pacC* coding region results in very poor growth, making this approach not very interesting to generate hosts for protein production (Tilburn *et al.*, 1995).

The gene *areA* is expressed in the absence of preferred nitrogen sources such as ammonium and encodes a protein that activates transcription of genes encoding enzymes (like proteases) involved in the utilizing of other resources (Ward *et al.*, 2005). Disruption of the *areA* gene in *A. oryzae* resulted in increased production of the heterologous protein chymosin due to reduced protease activity (Christensen & Hynes, 2000). Unfortunately, disruption of the *areA* gene in *A. niger* as well as *A. oryzae* also affected growth, even in culture medium with (low levels of) ammonium; this reduced growth was not noticed in *A. nidulans* (Christensen *et al.*, 1998; Lenouvel *et al.*, 2001).

The gene *creA* is expressed in the presence of preferred carbon sources such as glucose. The CreA protein represses the synthesis of enzymes (like proteases) involved in the catabolism of alternative carbon sources (Ruijter & Visser, 1997). However, attempts to disrupt the complete *creA* gene from *A. nidulans* resulted in lethal phenotypes (Dowzer & Kelly, 1991) or mutants with extremely severe effects on morphology (namely reduced growth rate and reduced conidiation) (Shroff *et al.*, 1997).

Altogether, the approach of using gene disruption of wide-domain regulatory genes seems unsuitable to generate proteases-deficient fungal host strains for protein production due to pleiotropic growth defects of this type of mutants. Specific mutation of these regulatory genes, alleviating the severe phenotypic effects of the complete knockout mutants could be used (Fraissinet-Tachet *et al.*, 1996). However, this approach relies on selection of specific spontaneous mutants making this approach not generally applicable.

The wide-domain regulatory mechanisms will be discussed in more detail later on in this chapter.

A novel and efficient method for isolation of protease-deficient fungi

Although both the classical screening approach and the gene-based approach have resulted in improved host strains, it is clear that both approaches have their limitations. The classical approach is very labor-intensive, whereas the disruption approach is limited by the availability of gene information. Therefore, we have developed a (direct) mutant selection approach, similar to those available for a number of other traits in filamentous fungi (*pyrG* [van Hartingsveldt *et al.*, 1987], *niaD* [Unkles *et al.*, 1989], *sC* [Buxton *et al.*, 1989]). This proprietary approach is based on a suicide substrate (SUI) to which protease mutants of fungi and yeasts are more resistant (SUI^R) than the parent strains (Punt *et al.*, unpublished results). The method can be used to select spontaneous mutants or mutants generated by mutagenesis by ultraviolet light irradiation. After a first round of selection the resulting mutants can be screened in a conventional milk halo screening. As shown in Table 2 the number of colonies resulting in a decreased halo formation is about 10% of the initial SUI^R strains even without UV-mutagenesis. In previous studies using milk halo screening after UV-mutagenesis only 0.1% of the surviving spores resulted in a reduced milk halo. With UV-mutagenesis prior to selection with the suicide substrate the efficiency of isolating protease-deficient mutants can be even further increased to over 50% (Punt *et al.*, unpublished result).

In Table 3 the analysis of a number of available and newly selected protease mutant strains is shown. Interestingly, also a mutant with a deficient intracellular protease gene (*pepE*), which results in no significant decrease of extracellular protease activity (van den Hombergh *et al.*, 1997a), can be selected with this method. From Table 3 it is also clear that, as is the case with virtually every method, not every type of protease mutant can be selected in this way. For example, a mutant lacking the major protease gene (*pepA*) in *A. niger*, which results in a residual extracellular protease activity of less than 20% (Mattern *et al.*, 1992; van den Hombergh *et al.*, 1997a), had no higher resistance against the suicide substrate than the wild type strain. Remarkably, with this approach also mutants with enhanced protease activity were selected (Punt *et al.*, unpublished results).

Table 2. Efficiency of isolation of protease-deficient mutants by spontaneous resistance to suicide substrate (SUI) compared to UV mutagenesis

Spontaneous resistance (SUI^R) of two *Aspergillus* species to suicide substrate *				
Strain	**No. of initial spores**	**No. of colonies SUI^R**	**Rescreen SUI^R**	**Reduced milk halo**
Aspergillus sp. section Nigri strain A	4×10^8	590	160/590	45/160
Aspergillus sp. section Nigri strain B	4×10^8	200	85/200	20/85
UV mutagenesis of *A. niger* † and *A. nidulans* ‡				
Strain	**No. of initial spores**	**Survival rate after UV mutagenesis**	**No. of spores screened for reduced milk halo**	**Reduced milk halo**
A. niger	5×10^4-1×10^5	10-20%	1×10^4	$7/1 \times 10^4$
A. nidulans	2.5×10^5-2.5×10^6	1-10%	2.5×10^4	$29/2.5 \times 10^4$

* Punt *et al.*, unpublished results
† Mattern *et al.*, 1992
‡ Katz *et al.*, 1996

Table 3. Protease mutants show higher resistance to the suicide substrate than WT strains *

Species	**SUI (mg/l)**						**Residual protease activity**	
	0	**100**	**200**	**300**	**400**	**500**	**intracellular**	**extracellular**
A. niger WT	+	+	-	-	-	-	100%	100%
A. niger pepA	+	+	-	-	-	-	100% †	15-20% †‡
A. niger pepE	+	+	+	+/-	-	-	30% †	~100% †
A. niger prtT	+	+	+	+	+/-	-	N/A	<5% ‡
A. niger prtT/phmA §	+	+	+	+	+	+/-	N/A	<5% *

* Punt *et al.*, unpublished results
† van den Hombergh *et al.*, 1997a
‡ Mattern *et al.*, 1992
§ The *A. niger prtT/phmA* mutant is a derivative of *A. niger prtT* that does not acidify its medium

FERMENTATION CONDITIONS

Strain improvement has proven to be a very useful tool for reducing the proteolytic degradation of especially heterologous proteins produced in the *Aspergillus* host strain. However, the large number of (extracellular) proteases able to degrade these heterologous proteins and the varying susceptibility of the produced heterologous proteins for the different proteases (Archer *et al.*, 1992; van den Hombergh *et al.*, 1995) makes one single (permanent) solution of the problem impossible. Therefore,

an additional way to improve heterologous protein production can be the development of fermentation conditions repressing protease production. Although numerous empirical approaches have been followed to address the protease issue, only very few systematic studies have been performed. From these studies three environmental parameters have emerged which have been studied in somewhat more detail, that is, ambient pH, carbon catabolite control and nitrogen metabolite control.

pH regulation

Ambient pH was shown to be an environmental parameter greatly influencing the expression of proteases. Controlled fermentations with *A. niger* at pH 4 or pH 5 resulted in a significant decrease of protease activity at higher pH. When cultured at pH 6, protease activity was even further decreased (Braaksma *et al.*, 2009). Culture pH was also suggested to be a key player during the production of recombinant GFP by *A. niger* and *A. sojae* (Gordon *et al.*, 2000; Heerikhuisen *et al.*, 2005). GFP excreted by the recombinant *A. niger* strain was rapidly degraded, whereas in *A. sojae* significant amounts of extracellular GFP could be detected. Acidification of the culture medium of *A. niger* was suggested to be the cause for proteolytic degradation of GFP, as under identical conditions *A. sojae* did not significantly acidify. Maintaining the pH at 6 during the production of GFP with *A. niger* resulted in a tenfold increase of GFP levels compared to a culture controlled at pH 3 (O'Donnell *et al.*, 2001). This increase was due to reduced degradation of GFP by proteases. Also, production of the human cytokine interleukin-6 (Il-6) in a protease deficient strain and a derivative of that strain, which did not acidify, resulted in improved yield and stability of Il-6 in the non-acidifying host strain (Punt *et al.*, 2002).

The genes encoding the two major extracellular proteases of *A. niger*, *pepA* and *pepB*, were not expressed under alkaline conditions (Jarai & Buxton, 1994). On the other hand, the transcript levels of the major alkaline protease gene *prtA* produced by *A. nidulans* was elevated under alkaline conditions (Tilburn *et al.*, 1995). This was, however, not confirmed by similar experiments conducted by Katz *et al.* (1996), where nitrogen starvation appeared to override the repression of *prtA* by low culture pH (VanKuyk *et al.*, 2000). From these results we conclude that ambient pH is a regulator of protease expression. In *A. nidulans* pH regulation is mediated mainly by seven genes, *pacC*, *palA*, *palB*, *palC*, *palF*, *palH*, and *palI*, where *pacC* plays the key role in the regulation of gene expression by ambient pH (Tilburn *et al.*, 1995). The products of the pal genes transduce a signal able to trigger the PacC into an active form. This active PacC is able to activate the expression of alkali-expressed genes (including *prtA*) and to inhibit the expression of acid-expressed genes (Peñalva *et al.*,

2002). Homologues of the *pacC* gene and the *pal* genes have been identified in other aspergilli, such as *A. niger* (MacCabe *et al.*, 1996), *A. fumigatus* (Bignell *et al.*, 2005) and *A. oryzae*, as well as all major groups of ascomycetes (Peñalva *et al.*, 2002). The involvement of pH control in extracellular protease production was further confirmed by analysis of protease expression in PacC mutants of *A. nidulans* and *A. niger* (Tilburn *et al.*, 1995; Fraissinet-Tachet *et al.*, 1996). However, the expression of three vacuolar proteases in *A. niger* is not regulated by PacC, which may also be the case with intracellular proteases of other aspergilli (Fraissinet-Tachet *et al.*, 1996).

Carbon catabolite control

Growth on glucose or other favored carbon sources prevents the synthesis of enzymes involved in the utilization of other substrates, such as polysaccharides (Ward *et al.*, 2005). This seems to apply for fungal extracellular proteases as well. Unfortunately, literature about the effect of carbon source on protease production by aspergilli is scattered and in addition often rather outdated. However, a few examples of the repressing effect of glucose and other carbon sources on the levels of excreted proteases have been described. When mycelia from *A. nidulans* were transferred to a medium without carbon source, extracellular proteases were abundantly produced. When mycelia were transferred to medium with glucose, lactose, galactose, or glycerol, protease production was severely repressed (Katz *et al.*, 2000). Similarly, transferring experiments with *A. oryzae* showed a strong decrease of protease production when mycelia were transferred to medium with casein and glucose compared to medium with casein only (Fukushima *et al.*, 1989).

The expression of the two extracellular proteases *pepA* and *pepB* of *A. niger* was studied in the presence of various carbon sources (Jarai & Buxton, 1994). When cells were transferred to medium supplemented with glucose, expression of both protease genes was repressed. In the presence of the less favorable carbon source glycerol the *pepA* gene was derepressed and in medium without carbon source *pepA* and *pepB* were both strongly derepressed. Thus, protease expression is clearly affected by glucose (or carbon catabolite) repression. Repression may be caused by various other carbon sources, but glucose is suspected to be the most repressive. The repressor protein CreA plays a major role in carbon repression. CreA inhibits transcription of many target genes by binding to specific sequences in the promoter of these genes (Ruijter & Visser, 1997). The gene encoding this protein has been identified in several *Aspergillus* species, such as *A. nidulans* (Dowzer & Kelly, 1989), *A. oryzae* (Kim *et al.*, 2001) and *A. niger* (Drysdale *et al.*, 1993). With Northern blot analysis, protease expression in *creA* mutants of *A. niger* gave clear evidence for the involvement of

carbon catabolite control (Fraissinet-Tachet *et al.*, 1996). Similarly, this was suggested by the fact that two of the isolated *A. nidulans* mutants, *xprF* and *xprG*, which carry a mutation in a hexokinase-like protein and an acid phosphatase, respectively, are thought to be involved in carbon catabolite repression and maybe also in nitrogen, sulfur, and phosphate regulation (Katz *et al.*, 2000; Katz *et al.*, 2006).

Nitrogen metabolite control

Similar as for the repression by glucose, the presence of preferred nitrogen sources such as ammonium suppress the production of enzymes, such as extracellular proteases, for utilizing other nitrogen sources (Ward *et al.*, 2005). For example, high concentrations of the preferred nitrogen source ammonium resulted in increased concentrations of bioactive tissue t-PA produced by *A. niger*, which was suspected to be due to less degradation of this heterologous protein (Wiebe *et al.*, 2001; Wiebe, 2003). Extracellular protease levels of *A. nidulans* were significant lower in a growth medium with ammonium compared to a nitrogen-free medium (VanKuyk *et al.*, 2000). The influence of nitrogen source on the expression of the *pepA* and *pepB* gene in *A. niger* was investigated by transferring cells to medium with and without ammonium. Cells grown with ammonia showed very low levels of both protease transcripts, whereas the levels of mRNA were much higher when cells were grown without ammonia (Jarai & Buxton, 1994).

The gene *areA* has been implicated in mediating the nitrogen metabolite control regulatory mechanism and it has been extensively studied in *A. nidulans* (Kudla *et al.*, 1990). The *areA* gene encodes a protein that activates transcription of many target genes by binding to specific sequences in the promoter of these genes. Homologues of this gene have also been identified in other *Aspergillus* species, such as *A. oryzae* (Christensen *et al.*, 1998) and *A. niger* (MacCabe *et al.*, 1998).

A study with an *A. niger* wild type strain and several different *areA* mutants (obtained by UV-mutagenesis and selection on chlorate plates) demonstrated that three intracellular protease genes were not controlled by AreA, because both wild type and *areA* mutants showed unaltered expression of these three genes (Fraissinet-Tachet *et al.*, 1996). The same study showed that three extracellular proteases were apparently regulated by AreA. However, the expression of the corresponding extracellular protease genes was not modulated in the same way in the different *areA* mutants, but depended on the combination of the protease gene and the particular *areA* mutation.

Sulfur and phosphorus metabolite repression

Already several decades ago the first studies on the effect of phosphorus and sulfur sources on protease expression in aspergilli were reported, but hardly any articles have been published on this subject since (Tomonaga *et al.* 1964; Cohen, 1972; Cohen, 1973; Cohen, 1981). Today, still little is known about sulfur and phosphorus metabolite repression in aspergilli and putative involvement in protease regulation. However, more recently a strong effect of sulfur limitation on the increase of protease activity for *A. nidulans* has been described (VanKuyk *et al.*, 2000). In addition, expression analyses of *prtA*, encoding the major extracellular protease in *A. nidulans*, showed a high transcript level when mycelia was transferred to sulfur-free medium (Katz *et al.*, 1996; Katz *et al.*, 1994).

Although the regulatory factors involved in sulfur metabolite repression are known (Natorff *et al.*, 1993; Natorff *et al.*, 2003), no information is available regarding protease gene expression. The regulatory factors involved in phosphorus metabolite repression are yet unknown. Identification of the role of these factors may help for a better understanding of the overall protease regulation.

Induction of protease by protein

The fact that in the presence of protein the production of proteases is stimulated has been applied for years in the production of extracellular proteases by the use of complex nitrogen and/or carbon sources (Singh & Vyas, 1977; Fukushima *et al.*, 1989; Srinivasan & Dhar, 1990; Singh *et al.*, 1994).

However, the opposite effect has also been described. Extracellular GFP could not be detected when the *A. niger* host strain was cultured on defined medium (Gordon *et al.*, 2000). When modified soya milk medium was used, fluorescence could be detected in the culture medium. The authors indicate that this was probably not due to a repressive effect of the soya milk protein, but due to the natural protease inhibitors that are present in the soya milk medium and the fact that the ambient pH can be maintained for longer at a value which limits protease induction than with defined medium. Another explanation is that the abundant availability of substrate for the proteases delayed the degradation of GFP.

The *A. niger pepA* and *pepB* protease genes were induced when mycelia was transferred to medium with elastin (Ruijter & Visser, 1997). Medium containing glucose next to elastin repressed expression of both proteases. Comparable

experiments by Jarai and Buxton (1994) showed a somewhat different picture, as *A. niger* expressed *pepA* and *pepB* in the presence of glucose if BSA was also present. When additional ammonia or urea was supplemented both protease genes were repressed. These results suggest that induction by the presence of extracellular protein plays only a secondary role in the regulation of extracellular proteases. As for the sulfur and phosphorus regulation mechanisms little is known about the mechanism of specific induction of protease gene expression by external addition of proteins. It is also possible that protein itself is not an inducer, but that the added protein or its peptide degradation products, being a complex carbon and nitrogen source all in one, play a role in the wide-domain regulation mechanisms of nitrogen metabolite and carbon catabolite control.

Bioprocess engineering

Affecting protease production by the means of bioprocess engineering has also proved to be a successful means of controlling extracellular protease activity. However, again very little has been published on the subject. Immobilization of the cells of *A. niger* to materials like a metal-coated pad or Celite beads reduced secretion of extracellular protease and increased the secretion of glucoamylase (Liu *et al.*, 1998; Papagianni *et al.*, 2002). Manipulating the morphology of *A. niger* by means of inoculum levels (concentration of spores) or inoculum type (vegetative or spores) was also shown to affect protease levels (Xu *et al.*, 2000; Papagianni & Moo-Young, 2002). Growth of the mycelium in the form of (large) pellets resulted in lower specific protease activities and increased protein production compared with a filamentous morphology. Morphology clearly affects protease secretion as well as protein production, but the exact mechanism needs further investigation (Grimm *et al.*, 2005).

The effect of the bioprocess parameters agitation intensity, dissolved oxygen tension as well as initial glucose and yeast extract concentration on protease and heterologous protein production has been studied in *A. niger* (Wang *et al.*, 2003). However, altogether these studies should be considered as exploratory, as no systematic analysis was performed.

SYSTEMS BIOLOGY APPROACH

Strain development and optimization of fermentation conditions have improved the production of (heterologous) proteins by aspergilli to a considerable extend. However, the problem of proteases has in most cases been approached by trial-and-error,

without taking the interaction between strain development and improvement of fermentation conditions in account (e.g., the best mutant may not be the best producer on the medium previously optimized for a precursor strain). Furthermore, the mechanism of induction and repression of protease production is far from completely understood. A more integrated approach is, therefore, desirable to come to a better understanding of the issue and from this to a solution that is also more generally applicable.

Recently developed techniques like (comparative) genomics, transcriptomics, proteomics, and metabolomics will very likely play a crucial role in understanding the proteolytic system of aspergilli. In addition to these –omics approaches we would also like to consider the role of the various physiological parameters involved in the fermentation process. These "physiomics" parameters such as pH, oxygenation, viscosity, agitation and so on add a further layer of data to be included in a full systems biology approach to study the proteolytic system of aspergilli.

Several articles reporting application of genomics techniques for research of *Aspergillus* strains have been published (e.g., Galagan *et al.*, 2005; Andersen *et al.*, 2008; Coutinho *et al.*, 2009). With the complete genome sequences of several *Aspergillus* strains open to the public (e.g., Machida *et al.*, 2005; Nierman *et al.*, 2005; Galagan *et al.*, 2005, Pel *et al.*, 2007) and more to be expected in the near future (for a recent overview, see Andersen & Nielsen, 2009), possibilities for studying these fungi on a systematic level are open for further research.

Transcriptomics is the most established of the genomics techniques. Several reviews discussing the results from these studies have already appeared (Breakspear & Momany, 2007; Andersen & Nielsen, 2009), illustrating the possibilities of this type of studies to elucidate complex biological processes in fungi.

The method for the identification of all proteins in complex mixtures is proteomic analysis. Initial approaches involved studying the proteins to be separated by one-dimensional (1D) SDS-PAGE. With the development of 2D gel electrophoresis, often coupled to mass spectrometry in order to identify the proteins, proteomic analysis has become a very powerful method for identification of proteins in complex mixtures. A few reviews on proteomics in filamentous fungi have been recently published (Carberry and Doyle, 2007; Kim *et al.*, 2007, Kim *et al.*, 2008).

One of the more recent functional genomics tools is metabolomics, the analysis of all intracellular and extracellular metabolites. Already in the mid-1990s a method to

extract intermediary metabolites from *A. niger* has been described by Ruijter and Visser (1996), and glycolytic intermediates were analyzed using an automated spectrophotometer. Since then, analytical platforms for metabolite detection have gone through major developments (van der Werf *et al.*, 2005; Koek *et al.*, 2006; Coulier *et al.*, 2006; Oldiges *et al.*, 2007). However, while based on these methods the potential for large-scale quantitative studies in aspergilli is present, relatively little has been published on metabolomics involving *Aspergillus* species (e.g., Frisvad *et al.*, 2008; Kouskoumvekaki *et al.*, 2008).

As is clear from the indicated studies, all functional genomics tools are still under development, with identification of expressed genes or proteins as the major challenge for transcriptomics and proteomics, respectively. However, for all genomics tools extracting relevant biological information from the overwhelming amount of data resulting from these tools is perhaps the biggest challenge. Focusing on the biggest changes in gene expression or protein or metabolite concentration does not automatically lead to the identification of the most important parameter in a biological process (van der Werf, 2004). The choice for a data pretreatment method and a data analysis method greatly affects the outcome (van den Berg *et al.*, 2006). The final goal will be to combine the results distilled from the high-throughput functional genomics methods with information from small-scale studies focusing on particular cellular functions and systems in order to construct a biological network of all protein and genetic interactions. A comprehensive collection of experimentally observed interactions has been put together for the best-studied eukaryote, the budding yeast *S. cerevisiae*, but it is suggested that there are probably many more interactions to be discovered (Reguly *et al.*, 2006). For *Aspergillus*, the study of complex biological networks, among which are also the proteolytic systems, is still in its infancy and will provide the scientific community with a huge challenge on the road to a more complete understanding of this type of organism.