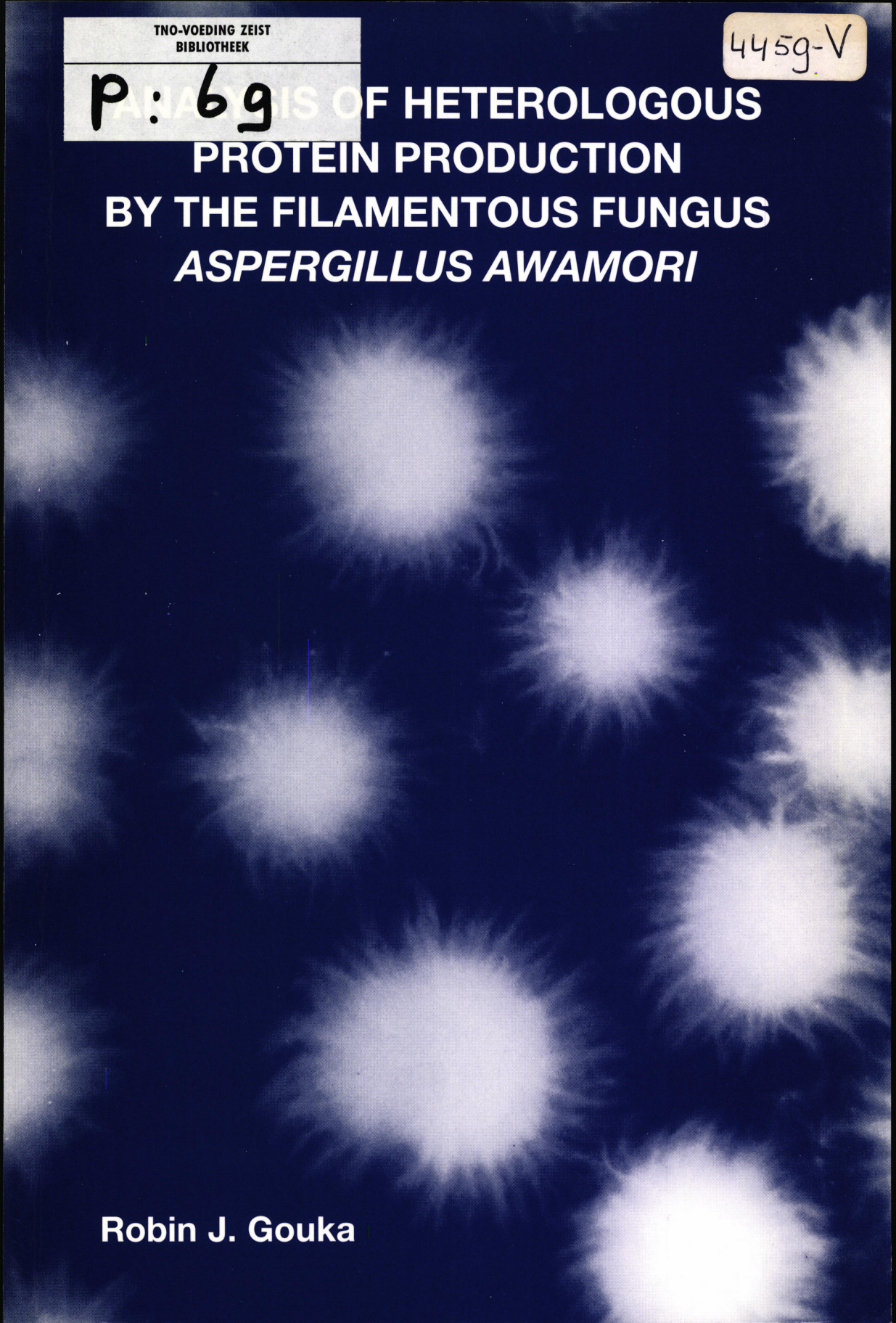
ANALYSIS OF HETEROLOGOUS PROTEIN PRODUCTION BY THE FILAMENTOUS FUNGUS *ASPERGILLUS AWAMORI*



Robin J. Gouka

ANALYSIS OF HETEROLOGOUS PROTEIN PRODUCTION BY THE FILAMENTOUS FUNGUS *ASPERGILLUS AWAMORI*

Analyse van heterologe eiwit produktie

door de filamenteuze schimmel *Aspergillus awamori*

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit Utrecht

op gezag van de Rector Magnificus, Prof. Dr. J.A. van Ginkel,

ingevolge het besluit van het College van Decanen

in het openbaar te verdedigen

op dinsdag 28 mei 1996 des middags te 13.30 uur

door

Robertus Johannes Gouka

geboren op 29 maart 1965 te Schiedam

Promotor : Prof. Dr. Ir. C.T. Verrips

(Verbonden aan de Faculteit Biologie
van de Universiteit Utrecht)

Co-promotor : Dr. C.A.M.J.J. van den Hondel

(Verbonden aan TNO, Instituut Voeding,
Afdeling Moleculaire Genetica en Gentechnologie, Rijswijk)

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Gouka, Robertus Johannes

Analysis of heterologous protein production by the filamentous fungus *Aspergillus awamori* / Robertus Johannes Gouka. - Utrecht : Universiteit Utrecht, Faculteit Biologie

Thesis Universiteit Utrecht. - With ref. - With summary in Dutch.

ISBN 90-393-1104-8

Subject headings: *Aspergillus awamori* / heterologous protein production.

Dit onderzoek is mede tot stand gekomen dankzij financiële ondersteuning van Unilever Research Laboratorium, Gist-brocades, DSM, Solvay-Duphar en het Ministerie van Economische Zaken.

Aan mijn ouders

Voor Jackie en Linda

Fotografie : M.J.M. Boermans

Illustratie omslag : Pelletvormig mycelium van *Aspergillus awamori*

Fotografie omslag : J.B.A. Gouka-Bergen Henegouwen

Gedrukt door: Drukkerij Haveka B.V., Alblasserdam.

Het onderzoek beschreven in dit proefschrift is uitgevoerd binnen de afdeling Moleculaire Genetica en Gentechnologie van het Instituut Voeding van de Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (TNO) te Rijswijk. Uitgave van dit proefschrift is financieel ondersteund door een bijdrage van TNO-Voeding en de faculteit biologie van de Universiteit Utrecht

CONTENTS

CHAPTER 1

I. INTRODUCTION

Filamentous fungi are able to occupy many different environments like soil, organic waste or plant cell material. The fact that filamentous fungi can secrete a broad range of hydrolysing and oxidizing enzymes, enables them to use a wide variety of organic compounds as source for their nutrients. This characteristic probably underlies their successful life style. The enzymes they secrete can degrade large complex biopolymers such as celluloses, hemicelluloses and starch.

Already more than a century ago, it became evident that fungi did not only secrete a wide variety of enzymes, but also were able to produce them in large quantities. In addition, fungi are able to secrete a great diversity of primary (e.g. organic acids) and secondary metabolites (e.g. antibiotics). These features have resulted in an increasing interest in studying filamentous fungi in the laboratory and using them in industrial processes. As a consequence, over the years, classical genetic studies have resulted in a lot of knowledge about their genetics and their biochemical pathways. To date, a number of filamentous fungal strains and their products have been used in food and food processing industry and have resulted in a so called GRAS status (Generally Regarded As Safe) of some of their products.

The development of molecular biological techniques has opened new ways to use filamentous fungi for the production of homologous and heterologous (especially mammalian) proteins. In particular species of the genera *Aspergillus* and *Trichoderma* are being used extensively for large scale production of proteins. Initially, work on heterologous protein production in *Aspergilli* was started with *A. nidulans*, the species in which (molecular-)genetics and biochemical pathways have been studied since the early 1940's (see Pontecorvo et al. 1953). During the last decade, other species, mainly *A. niger*, *A. awamori* and *A. oryzae*, have been used more frequently since with these strains higher levels of recombinant proteins can be obtained.

This chapter will be focused on the use of filamentous fungi for the production of heterologous, especially non-fungal, proteins and will deal with the tools, the prospects, and improvement strategies for overproduction of these proteins.

II. EXPRESSION SYSTEMS FOR HETEROLOGOUS GENE EXPRESSION

The development of fungal recombinant DNA techniques (in particular that of DNA-mediated transformation) has greatly stimulated the use of fungi for the production of proteins. In fact, molecular-genetic techniques have become an important tool to construct protein overproducing strains and have partly replaced classical strain improvement strategies (mutagenesis). Unlike traditional strain improvement techniques, molecular-genetics uses techniques that highly increase specificity in generating protein overproducing strains, so reducing the time necessary for screening of putative overproducing strains. The most commonly used strategy to generate overproducing strains is to isolate the gene encoding the protein of interest, clone it in a suitable vector and introduce the resulting expression vector via a gene-transfer (transformation) procedure in multiple copies into the genome of the fungus. Expression vectors comprise an expression cassette, containing the gene encoding the protein of interest provided with appropriate transcription/translation(/secretion) control sequences. Using this strategy a large number of homologous and heterologous proteins have already been successfully produced in *Aspergillus* (reviewed by a.o. Van den Hondel et al. 1991; Gwynne and Devchand 1992; Verdoes et al. 1995) and *Trichoderma* species (Keränen et al. 1995).

A. Gene-transfer systems

Since the first successful transformation procedures for *Aspergillus* species have been published in the early 1980's (Ballance et al. 1983, Tilburn et al. 1983), an increasing number of gene-transfer (transformation) systems for a variety of fungi have been developed (reviewed by Ballance 1991). The majority of the transformation systems are based on either complementation of auxotrophic mutant strains with the corresponding wild-type gene or dominant selection marker genes (often bacterial), which confer resistance to a fungicide (reviewed by Van den Hondel and Punt 1991; Turner 1994). In contrast to yeasts and prokaryotes, where vectors can be maintained autonomously, the introduction of a vector in a fungus generally results in the integration into the host genome (Turner 1994). As a consequence of genomic integration, the vector is usually present in a low copy number. In a low percentage also strains containing multiple copies of the vector can be obtained, and especially the use of specialized cosmid-based vectors can result in strains containing up to 200 copies of a vector (Verdoes et al. 1995). A few years ago, also an autonomously replicating vector for filamentous fungi has been described (Gems et al. 1991). However, due

to gross instability, transformants obtained with this vector cannot be used for the overproduction of proteins.

The advantage of dominant selection markers is that they are applicable to wild-type strains, if sensitive to the compound (e.g. hygromycin resistance, Punt and van den Hondel 1993). However, in general, the lack of homology between the marker gene and the genome of the fungus results in random integration.

In case of auxotrophic selection markers, isolation of mutant strains is required. Mostly, mutations are introduced randomly by UV mutagenesis, followed by screening for the correct mutants. For some mutants, however, specific selection systems are available which facilitate the isolation of specific mutants. Examples are the selection of nitrate reductase (*niaD*) mutants on chlorate (Gouka et al. 1991), acetyl-CoA synthetase (*facA*) mutants on fluoro-acetate (Gouka et al. 1993) and orotidine-5'-monophosphate-decarboxylase (designated *pyrA*, *pyrG*, *pyr4* or *ura3*) mutants on 5-fluoro-orotic acid (5-FOA). In particular the use of 5-FOA for selection of mutants is convenient since only one additional type of mutants (orotate pyrophosphoribosyl transferase, designated *pyrF* or *ura5*) is obtained, whereas with the other selection systems multiple mutations can lead to resistance. A second advantage of the usage of *pyrG* and *pyrF* mutants is that both genes have been isolated and transformation systems based on both genes have been developed. The relative ease of selection and isolation of *pyrG* mutant strains has resulted in the development of transformation systems based on the *pyrG* gene for many fungi (Table I). For a large number of these mutant fungi, it was shown that a heterologous selection marker gene can also be used to obtain transformants. However, the integration of these non-homologous vectors is, by nature, ectopic. If site-specific integration is required, e.g. for comparison of expression levels or requirement of defined strains, only a homologous gene can be used. Nevertheless, also with a homologous gene the integration type depends on the fungal species used (Table I). The frequency of homologous integration could drastically be increased by the usage of a vector containing a mutated *A. niger pyrG* gene as selection marker. However, the transformation frequencies obtained with a mutated *pyrG* gene are considerably lower (Van Gorcom et al. 1988). This was probably the reason that a similar system for site-specific integration was not successful in a different strain of *A. niger* (Fowler et al. 1990).

Table I. Homologous transformation systems based on the gene encoding orotidine-5'-monophosphate decarboxylase (*pyrG*, *pyr4*, *ura3*)

Organism	Transf.freq. per μg DNA	n[1]	Integration event[2] I	II	III	Reference
A. nidulans	2000	10	60	30	10	Oakley et al. 1987
A. niger	50	5	80	20	0	Goossen et al. 1987
	40	16	40	50	10	Van Hartingsveldt 1987
	>100	48	0	100[3]		Van Gorcom et al. 1988
	10[4]	32	45	55[3]		Van Gorcom et al. 1988
	<<[5]					Fowler et al. 1990
A. oryzae	20	12	65	10	25	De Ruiter-Jacobs 1989
A. parasiticus	30-50	10	20	20	60	Skory et al. 1990
M. circinnelloides	600-800[6]	15				Benito et al. 1992
N. crassa	1-5[7]					Buxton and Radford 1983
P. chrysosporium	200	9	0	0	100	Akileswaran et al. 1993
T. reesei	12,000	7	15	0	85[8]	Gruber et al. 1990
	> 10,000	9	20	10	70[8]	Bergès et al. 1991

1) Number of transformants tested by Southern blot analysis

2) Percentage of: I, homologous integration of one or more copies; II, gene-replacement/conversion; III, ectopic integration

3) No data available about distribution of percentage of type II and III integrations

4) Vector contains a *pyrG* gene with a mutation

5) Vector with a promoterless *pyrG* gene (from a different *A. niger* strain) to promote homologous integration. The transformation frequencies were too low to obtain the desired transformants

6) Vector is based on an autonomously replicating plasmid, no integration occurs

7) Direct selection of transformants could not be carried out because the *pyr4* allele was too leaky to permit selection. Transformants were obtained by double selection with the *qa-2*$^+$ gene

8) Also autonomously replicating vector DNA observed

B. Expression vectors

The vectors used for transformation are usually plasmids originally isolated from bacteria, containing a bacterial origin of replication and a bacterial marker. In these vectors a marker for selection of fungal transformants has been introduced. Furthermore, the vector contains an expression cassette comprising the gene of interest provided with efficient fungal expression signals, i.e. a transcription control region and a transcription termination sequence. The transcription control region (further referred to as tcr) is defined as the DNA sequences required to initiate gene transcription plus those required to regulate the rate at which initiation occurs. These include the promoter, where the general transcription factors and the polymerase assemble plus all of the regulatory sequences to which gene regulatory proteins bind to control the rate of these assembly processes at the promoter. For secretion of the gene product a signal sequence is included, usually that of a well-secreted protein although secretion signals of the heterologous gene have also been used successfully. The selection marker and expression cassette might also be present on separate vectors and introduced in the fungus by cotransformation.

B1. Transcription control regions

In general, the *Aspergillus* tcrs used for the expression of heterologous genes are derived from Aspergillus genes that are highly expressed (reviewed by Davies 1994). These tcrs originate from inducible or constitutively expressed genes involved in carbon-source utilization pathways, as in particular several of these genes are highly expressed.

In the late 1980's most of the work on heterologous gene expression in filamentous fungi was started by three companies, who developed each their own fungal expression system. Genencor started with the *A. niger glaA* expression system to produce bovine chymosin. Although initial production levels of chymosin did not exceed 3 mg l^{-1} (Cullen et al. 1987), the system was optimized by a series of different approaches (Ward et al. 1990; Dunn-Coleman et al. 1991) to yield more than 1 g l^{-1} of chymosin. Both the *A. niger* and the *A. awamori glaA* tcr has since then been utilized to produce a great variety of other heterologous proteins like human interferon α2 and *Cellulomonas fimi* endoglucanase (Gwynne et al. 1987), hen egg white lysozyme (Archer et al. 1990b), porcine pancreatic phospholipase A2 (Roberts et al. 1992), human interleukin-6 (Carrez et al. 1990, Contreras et al. 1991) and human lactoferrin (Ward et al. 1995).

The Allelix group started with the *A. nidulans* ethanol-inducible alcohol dehydrogenase (*alcA*) expression system (Gwynne et al. 1987, Devchand and Gwynne 1991). In strains

comprising multiple copies of the *alcA* tcr it was found that titration of a regulatory gene (*alcR*) occurred. The *alcR* gene was cloned and reintroduced in multiple copies into the genome and this multiple copy strain is now used as a host for expression of genes controlled by the *alcA* tcr (Gwynne et al. 1989). The *alcA* expression system was also used to produce human lactoferrin in *A. nidulans* (Ward et al. 1992a).

The group at Zymogenetics started with the constitutive *A. nidulans* triosephosphate isomerase (*tpiA*) tcr (Upshall et al. 1987) to produce human tissue plasminogen activator. After a comparative study of three different fungal tcrs, that of the aldehyde dehydrogenase (*adhA)* of *A. niger* gave the highest production levels (13-300 μg l^{-1}) and was later used to produce human granulocyte-macrophage colony-stimulating factor (hGMCSF) in *A. nidulans* in levels of 1 mg l^{-1} (Upshall et al. 1991).

Other strong fungal tcrs that have been used by other research groups are derived from the *A. oryzae* α-amylase (*amy*) or TAKA-amylase gene (Boel et al. 1987; Christensen et al. 1988; Boel et al. 1989; Ward et al. 1992b), mainly used by Novo Nordisk, and the constitutive *A. nidulans* glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) gene (Archer et al. 1990b; Contreras et al. 1991; Punt et al. 1991) developed in the TNO laboratory at Rijswijk.

From an industrial point of view, favourable transcription control regions are those that can be efficiently induced (i) to high levels of expression and (ii) with a cheap substrate. Several of the transcription control regions listed above, and also that of the frequently used cellobiohydrolase (*cbhI*) gene of *Trichoderma reesei* (a.o. Harkki et al. 1989), fulfil these requirements. The substrates are usually high molecular carbohydrates (biopolymers) that are abundantly present in nature e.g. starch and cellulose. In the natural environment in which filamentous fungi live, these (complex) substrates are readily available. To live, filamentous fungi produce and secrete the necessary polymer degrading enzymes after the induction of expression of the corresponding genes. Plant cell walls contain three major polymers: cellulose, hemicellulose and lignin. After cellulose, hemicelluloses (including glucans, mannans, arabinans, galactans and xylans) are the most abundant biopolymers synthesized. They can make up as much as 35% of the dry weight of plants (Joseleau et al. 1992). Therefore, transcription control regions of enzymes involved in the degradation of hemicelluloses might be good candidates for the development of efficient expression systems.

Recently, in the laboratory of TNO, the *A. awamori exlA* gene, encoding 1,4-ß-endoxylanase A (EXL) was isolated. Endoxylanases cleave internal ß-1,4-glycosidic bonds within the backbone of xylan, a hemicellulose (recently reviewed by Thomson 1993). The majority of

xylans, like most hemicelluloses, are branched polymers, usually carrying a variable number of substituents (Biely 1985). The nature of the various substituents present in the branched polymers depend on the plants cell from which the xylan originates. Typical residues are 1,3-linked α-L-arabinose and α-1,2-linked 4-O-methyl glucuronic acid residues, other substituents are acetic acids and phenolic residues such as coumaric and ferulic acid (Joseleau 1992). Cleavage of xylan by 1,4-ß-endoxylanases results in a release of xylo-oligosaccharides, the end products depend on the type of 1,4-ß-endoxylanase. Kormelink et al. (1993) isolated three different 1,4-ß-endoxylanases (I, II and III) from *A. awamori*. Only one type (I) released xylose, xylobiose and xylotriose, whereas the other two types (II and III) mainly released xylobiose and xylotriose. Complete hydrolysis of oligosaccharides to xylose is achieved by a series of additional enzymes, including an exo-acting enzyme, ß-xylosidase and a series of accessory debranching enzymes including α-L-arabinofuranosidase, α-D-glucuronidase, acetyl esterase and ferulic and coumaric esterase (Biely 1985; Coughlan and Hazlewood 1993).

Regulation of induction of xylanases is still poorly understood. Xylan itself is too large to be transported into the cells and is therefore not considered to be the physiological inducer. Evidence for constitutive expression of certain xylanases has been obtained in *Trichoderma reesei* (Herzog et al. 1992) and *Streptomyces cyaneus* (Wang et al. 1992), suggesting that a constitutive level is required to generate low-molecular mass inducers from xylan. Besides xylan, xylobiose and xylose have been shown to induce the synthesis of xylanases in a range of microorganisms like *A. terreus* (Hrmova et al. 1991), *A. sydowii* (Ghosh and Nanda 1994) and *A. tubigensis* (De Graaff et al. 1994). Interestingly, xylose is not an inducer of 1,4-ß-endoxylanases in *T. reesei* (Hrmova et al. 1986) and *A. nidulans* (Fernandez-Espinar et al. 1992; Pinaga et al. 1994).

In some studies xylo-oligosaccharides, non-metabolized analogues of xylobiose (such as ß-methyl-xyloside), and/or positional isomers of xylobiose have been shown to induce 1,4-ß-endoxylanases (Bailey and Poutanen 1989; Hrmova et al. 1991; Ghosh and Nanda 1994; De Graaff et al. 1994). Studies carried out in the TNO laboratory in Rijswijk have indicated that the transcription control region of the *A. awamori* 1,4-ß-endoxylanase A gene (*exlA*) can be used efficiently for the overexpression of its gene product 1,4-ß-endoxylanase A (Hessing et al. 1994) after induction with xylan. The usefulness of this transcription control region for the production of heterologous proteins was studied both in shake flask cultivations and in continuous culture fermentations (Chapter 3, 6).

B2. Transcription terminators

Gene expression is usually regulated at the level of transcription. Besides initiation and elongation of mRNA transcription, also termination of mRNA transcription and processing of the 3' end of the transcripts may be crucial steps in gene regulation. Termination of transcription involves processing of the 3' end and addition of a poly(A) tail, which is found in almost all eukaryotic messenger RNAs. In the nuclei of mammalian and yeast cells, which have been studied with respect to transcription termination most in detail, poly(A) tails are added post-transcriptionally in a two-step reaction: endonucleolytic cleavage of the primary transcript is rapidly followed by polyadenylation of the cleavage product (Christofori and Keller 1988; Sheets and Wickens 1989; Wickens 1990, Proudfoot 1991; Wahle 1991; Bienroth et al. 1993). In mammalian cells, cleavage as well as polyadenylation depends on two sequence elements (Wahle and Keller 1992): a hexanucleotide sequence (AAUAAA), present 10-30 nucleotides upstream of the cleavage site (Proudfoot 1991), and a less conserved GU-rich sequence downstream of the cleavage site. This GU-rich sequence has been proposed to determine the efficiency of polyadenylation (reviewed by a.o. Wickens 1990; Proudfoot 1991). In contrast to the highly conserved mammalian AAUAAA sequence, the polyadenylation signal in yeasts does not seem to be very conserved. In fact, most genes of *S. cerevisiae* lack the AAUAAA processing signal. In yeast three elements, of which two are diffuse AU-rich sequences, have been distinguished (Irniger et al. 1991; Russo et al. 1991; Russo et al. 1993; Guo and Sherman 1995): (i) an upstream element (such as TATATA, TAG...TATGTA and TTTTTATA) that enhances the efficiency of utilization of the downstream element, (ii) a downstream element (such as TTAAGAAC, AAGAA and AATAAA) which defines the position of the polyadenylation site and (iii) the actual site of polyadenylation, which is located 3' of the downstream element, often directly after a cytidine.

In filamentous fungi, the mechanisms and requirements of transcription termination, 3' end processing and polyadenylation have not been investigated experimentally. In some genes a hexanucleotide AAUAAA sequence has been found in the 3' DNA sequence, although in many others it is absent (Gurr et al. 1988; Bruchez et al. 1993), indicating that also in filamentous fungi other motifs or combinations of motifs might be necessary to act as polyadenylation signal. When a hexanucleotide AAUAAA is present, the functionality of it has never been proven experimentally. Nevertheless, in general, filamentous fungal polyadenylation signals are probably more yeast-like, i.e. diffuse AU-rich sequence(s).

The transcription termination sequences that are used in fungal expression cassettes may correspond to the origin of the transcription control region such as the *A. niger glaA*

transcription termination region (a.o. Cullen et al. 1987; Archer et al. 1990b; Van Hartingsveldt et al. 1990; Ward et al. 1990; Ward et al. 1995). In addition, some other fungal transcription termination regions are used which do not correspond to the origin of the transcription control region e.g. the *A. nidulans trpC* transcription termination region (Van den Hondel et al. 1991).

B3. Signal sequences

Secretion of proteins into the medium is initiated by the presence of a short hydrophobic signal peptide at the amino-terminal end of the protein (Kikuchi and Ikehara 1991; Dalbey and Von Heijne 1992). This signal peptide targets the nascent polypeptide chain to the endoplasmic reticulum (ER). Although little is known about the secretory pathway in filamentous fungi, a lot of data are available from yeast and mammalian systems (reviewed by a.o. Rapoport 1990; Rapoport 1992; Larriba 1993; Gilmore 1993). In general, it is believed that secretion of filamentous fungal proteins does not differ to a large extend from these systems.

As signal peptides are the signal for translocation into the ER, they are not only found in secreted proteins but also in proteins that have cellular compartments or the plasma membrane as their destination (Gilmore 1993). Signal peptides are cotranslationally recognized and bound by a signal recognition particle (SRP), which transports the ribosome/mRNA/peptide complex to the endoplasmic reticulum. The complex is targeted to the ER-membrane by an interaction with a membrane-bound receptor, the SRP receptor or docking protein. In the ER, the signal peptide is removed cotranslationally by the signal peptidase complex, an integral component of the ER membrane. The process of translocation is still unclear. One of the current models suggests that proteins are transported through a hydrophillic tunnel that is assembled from transmembrane proteins (Gilmore et al. 1993). Evidence has been obtained that an integral ER-membrane protein complex consisting of several membrane proteins (including sec61p, sec62p and sec63p) is indeed involved in translocation (Rapoport 1992).

Although signal peptides show a great diversity in both primary sequence and amino acid composition, a general design appears to be common from bacteria to mammals. Signal peptides are usually short (14-22 amino acids) and have a conserved primary three-domain structure (Von Heijne, 1986), consisting of an amino-terminal positively charged region, followed by a central hydrophobic part and a more polar carboxy-terminal part. This common design probably underlies the effective use of signal sequences of heterologous proteins for their production and secretion in filamentous fungi and yeast (e.g. bovine chymosin (Cullen

et al. 1987), human tPA (Upshall et al. 1991), hen egg white lysozyme (Archer et al. 1990b), plant thaumatin (Hahm and Batt 1990). However, most fungal expression cassettes contain a fungal signal sequence that corresponds to the transcription control region sequence (e.g. *glaA*; Ward et al. 1990; Carrez et al. 1990; Van Hartingsveldt et al. 1990).

B4. Prosequences

A large number of mature proteins are initially synthesized with a N-terminal extension, the propeptide. In most cases, the exact amino acid sequence of propeptides has not been determined experimentally, since these non-processed proproteins are difficult to obtain unless a processing mutant strain is available. Nevertheless, the cleavage site between a signal peptide and a propeptide can be determined theoretically with reasonable fidelity. From these data, it can be deduced that propeptides are highly heterogeneous, varying from very small peptides e.g. GLA (6 AA; Nunberg et al. 1984) and EXL (11 AA; Hessing et al. 1994), up to relatively long peptides e.g. PEPA (49 AA, Berka et al, 1990) and PEPD (101 AA, Jarai et al. 1994a).

Since a considerable amount of proteins destined for the secretion pathway from mammalian cells and some yeast proteins contain a propeptide, most experimental studies on the function of propeptides have been carried out in these organisms. From these data it is clear that propeptides can perform a number of different functions. Firstly, propeptides might contribute to the efficiency of cotranslational translocation of the protein across the ER-membrane. Secondly, propeptides might contribute to cotranslational proteolytic processing of the polypeptide. For example, deletion of the propeptide (Ala-Leu-Val-Arg-Arg) from human preproapolipoprotein A-I redirected cotranslational processing by the signal peptidase to a site located between the 2nd and the 3rd residue of the mature protein, whereas it did not affect translocation (Folz and Gordon 1986). Thirdly, they might act as intracellular targeting signal for routing to a specific cellular compartment. In some yeast proteins, e.g. carboxypeptidase Y, amino acid residues in the propeptide have been determined that participate in vacuolar targeting (a.o. Rothman et al. 1989). Fourthly, in some proproteins like proproteases, the propeptide keeps the protein inactive until it reaches its site of action (e.g. prochymosin).

Removal of the propeptide from the mature protein, in mammalian, yeast as well as in fungal cells, occurs in general by processing by a specific endopeptidase, usually after the two positively charged amino acid residues Arg-Arg or Lys-Arg (reviewed by Steiner et al. 1992). However, also other amino acid combinations, containing at least one basic amino

acid, have been found to be processed (e.g. EXL; Ser-Arg, Hessing et al. 1994). Comparison of all possible dipeptide distributions along a large number of fungal cytoplasmic and secreted proteins revealed that the secreted fungal proteins are biased against certain combinations, including Lys-Arg, Arg-Arg and Arg-Lys (Calmels et al. 1991). These results suggest that the absence of these doublets in mature, endogenous secreted proteins is to protect them from proteolytic cleavage. As dibasic cleavage is thought to occur in the Golgi, the internal dibasic peptide sequence in cytoplasmic proteins will not be attacked by this processing. In this respect it might be worthwhile to investigate whether the presence of dibasic doublets in foreign proteins accounts for the low efficiency of their secretion by filamentous fungi. For example, in yeast, Elliot et al. (1989) showed that cleavage at dibasic sites was responsible for the release of truncated human erythropoietin.

Little experimental data are available on the function of fungal propeptides. Deletion of the propeptide from *Fusarium solani pisi* cutinase resulted in increased levels of secreted cutinase by *A. awamori*, compared to a strain containing cutinase with its propeptide (Van Gemeren et al, in press). Van Hartingsveldt et al (1990) found intracellular accumulation of prochymosin in strains containing a fusion of prochymosin with the preprosequence in *A. niger*. When the *glaA* prosequence was deleted intracellular accumulation was not observed. The mRNA levels were identical, indicating that in this case the effect was exclusively at the secretion level.

For the production of foreign proteins in filamentous fungi, in general, the authentic heterologous propeptide is included in the fungal expression cassette when the heterologous protein contains one (e.g. prochymosin, Ward et al. 1990; prophospholipase A, Roberts et al. 1992).

B5. Ready to use expression vectors

To date, several groups have constructed expression vectors which contain useful restriction sites at positions to clone a heterologous gene and to produce the gene product, either intra- or extracellularly. In the TNO laboratory in Rijswijk, the pAN52-vector series have been developed, based on the *A. nidulans gpdA* or *A. niger glaA* transcription control region and the *A. nidulans trpC* transcription termination region (Van den Hondel. 1991). These vectors have either a unique cloning site at the translation initiation codon to connect the gene with the transcription control region for intracellular production of the protein, or a cloning site downstream of the *glaA* pre(pro)sequence for extracellular protein production. Similarly, the pAN56-vector series was developed for construction of glucoamylase gene-fusions (Fig. 1; see also section IV). These vectors are based on the *A. niger glaA* or *gpdA*

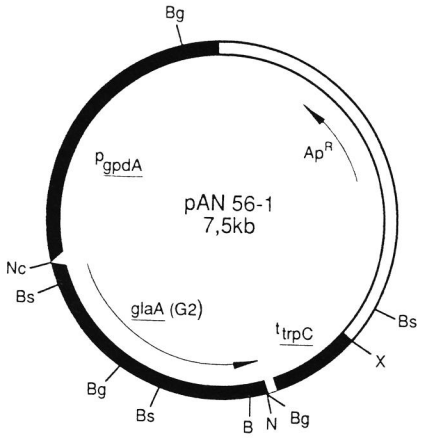


Figure 1 Example of a ready to use expression vector for the production of heterologous proteins. The vector is explained in the text.

transcription control region, followed by the *glaA* gene, and the *A. nidulans trpC* transcription termination region (Punt et al. unpublished data) with the latter two separated by unique cloning sites.

III. PRODUCTION OF EXTRACELLULAR PROTEINS

As already detailed above, especially *Aspergillus* and *Trichoderma* species have a number of favourable characteristics which makes them attractive hosts for large scale production of proteins. Besides their capacity to secrete substantial amounts of protein into the medium (reviewed by a.o. Fowler and Berka 1991; Van den Hondel et al. 1991; Verdoes et al. 1995), also the well-established and safe use of the enzymes they produce in food and food processing industry and their long history in large scale fermentation processes has contributed to the use of these species in food and food processing industry and has resulted in a GRAS status for some of the products derived from species like *A. niger* and *A. oryzae*.

At present, a number of enzymes, mainly native secreted *Aspergillus* enzymes, are produced commercially in large scale industrial fermentation processes (Table II; Berka et al. 1992; Bodie et al. 1994). Major applications are the manufacture of foods and beverages. The increasing amount of data on heterologous protein production in papers and patents emphasizes the effort that is being put into the use of filamentous fungi for the commercial

Table II. **Enzymes produced by *Aspergilli* in large scale fermentation processes**

Enzyme	Applications
Fungal	
Glucoamylase	Saccharification of starch for glucose syrup and alcohol
α-Amylase	Saccharification of starch, production of bread, maltose syrups, fermented oriental products
Pectinolytic enzymes	Production of wine and fruit juice (increased juice yield and clarification)
Proteases	Fermentation of wheat, soybeans and rice in the production of Asian foods and beverages, processing of dough, beer brewing
Catalase	Production of gluconic acid, removal of H_2O_2 and/or generation of O_2 in foods and beverages
Glucose Oxidase	Determination of glucose concentration in body fluids, component of other diagnostic enzyme kits, removal of glucose or oxygen from foods and beverages, determination of gluconic acid
Cellulases	Source of carbon, energy and specialty chemicals, treatment of brewing and soybean waste, generation of flavours and aromas, maceration of fruits and vegetables, enzymatic modification of textiles
Lipase*	Dairy industry for enhanced flavour development (milk, butter, cheese), yoghurt production, detergent
Hemicellulase (Xylanase)*	Brewing industry to reduce wort viscosity, baking industry to decrease dough strength and increase loaf volume, poultry feeds to aid digestion, paper and pulp industry
Phytase*	Animal feed supplement to increase the availability of dietary minerals and decrease phosphate waste
ß-Galactosidase	Food and dairy industry
Asparaginase	Pharmaceutical industry
Non-fungal	
bovine chymosin*	Milk clotting during cheese manufacture
lipase*	Detergent, food and dairy industry

* Recombinant strains

production of heterologous proteins. However, from Table II it is clear that the commercial production of heterologous proteins is still in its infancy and more basic knowledge is needed to improve production levels for many proteins now under investigation.

A. Production of fungal proteins

Production of fungal proteins, either of homologous or heterologous origin, by filamentous fungi is usually very efficient and production levels of grams per litre are within reach. To date, *Aspergillus* and *Trichoderma* strains have been reported which are able to secrete more than 30 grams per litre of homologous proteins in fermentation processes (e.g. cellulases, Durand et al, 1988 or glucoamylases, Finkelstein et al, 1989). As an example of a heterologous fungal protein, *Rhizomucor miehei* aspartic proteinase was secreted at levels up to 3 grams per litre by an *A. oryzae* transformant in controlled fermentations (Christensen et al. 1988). Production of fungal proteins has been reviewed extensively by many authors and is not discussed here further in detail (Van den Hondel et al. 1991; Gwynne and Devchand 1992; Verdoes et al. 1995).

B. Production of non-fungal proteins

Production levels of most non-fungal proteins, either of mammalian, bacterial, avian or plant origin, are low as compared to homologous proteins and reach levels that, with some exceptions, do not exceed a few tens of milligrams per litre of culture medium (for reviews see a.o. Jeenes et al. 1991; Van den Hondel et al. 1991; Gwynne and Devchand 1992).

In literature (see reviews above), a large number of factors that might influence the production levels of non-fungal proteins have been described. These data indicate that production can be limited at any level, beit transcription, translation, secretion and/or extracellular degradation. Limitations at the transcriptional level can be caused by low stability of the mRNA. Since usually highly inducible fungal transcription control regions are used, in general the mRNA levels of heterologous genes, especially in multi copy strains, are not considered to be limiting. Therefore, in a large number of studies mRNA levels are not reported or have not been determined and experimental data concerning the influence of mRNA stability on the levels of secreted protein remain sparse.

Limitations at the (post-)translational level can be inefficient translation initiation and elongation, translocation, folding, transport, processing or secretion of the protein (Ward et al. 1990; Turnbull et al. 1989; Tsuchiya et al. 1992; Broekhuijsen et al. 1993). Additionally, even when proteins are finally secreted efficiently, a major problem for heterologous proteins

is their degradation by extracellular proteases. *Aspergilli* can secrete a diversity of extracellular proteases (Mattern et al. 1992; Van den Hombergh et al. 1995) and proteases have been shown to be responsible for the degradation of a large number of heterologous proteins like human interleukin-6 (Broekhuijsen et al. 1993), bovine chymosin (Van Hartingsveldt et al. 1990) or porcine pancreatic phospholipase A (Roberts et al. 1992).

From the data available it is clear that the limiting factor(s) often depends on the protein to be expressed and/or the fungal host strain. Unfortunately, thus far systematic studies have not been carried out to investigate this in detail.

IV. APPROACHES FOR IMPROVEMENT OF HETEROLOGOUS (NON-FUNGAL) PROTEIN PRODUCTION

Based on the limitations observed for the production of non-fungal proteins (see section III B), several strategies have been developed to improve protein yields (reviewed by Archer et al. 1994). Most of these strategies are similar as for fungal proteins (Verdoes et al. 1995) and include (i) the introduction of a large number of gene copies, (ii) the use of strong fungal transcription control regions and efficient secretion signals, (iii) the construction and use of protease deficient host strains, (iv) development of an optimal medium and (v) gene fusions with a gene encoding part or whole of a well-secreted protein (see IVA).

A. Gene fusion strategies

Especially the application of a gene fusion strategy has been successful for the production of non-fungal proteins and has resulted in a change in type of expression cassettes in the last few years. Where originally the coding region of a non-fungal gene was fused to efficient fungal expression signals, now the heterologous gene is fused to the 3' end of a highly expressed gene, of which the gene product is efficiently secreted. For this gene fusion strategy mainly the *A. niger* or *A. awamori* glucoamylase gene was used. Fusion has resulted in higher levels of secreted bovine prochymosin (Ward et al. 1990), porcine pancreatic phospholipase A2 (Roberts et al. 1992), human interleukin-6 (Contreras et al, 1991; Broekhuijsen et al, 1993), hen egg white lysozyme (Jeenes et al. 1993), and human lactoferrin (Ward et al. 1995). The increase varies per protein from 5- to 1000-fold, resulting in protein levels varying from 5 to 250 mg l^{-1} (Table III). Even higher levels, up to 1 - 2 g l^{-1}, were obtained for chymosin (Dunn-Coleman et al. 1991) and lactoferrin (Ward et al. 1995) when high level production strains were subjected to several rounds of mutagenesis.

Besides the *glaA* gene-fusions, fusions with the *Trichoderma reesei* cellobiohydrolase I gene (Nyyssönen et al. 1995) improved the production levels of antibody fragments with more than 150-fold. Also a fusion of *A. awamori* α-amylase (both *amyA* and *amyB*) with prochymosin was reported to be successful (Korman et al. 1990). Shibuya et al. (1992) constructed an homologous *A. oryzae* TAKA-amylase/glucoamylase gene fusion to produce a fusion protein, up to levels of 0.5 g l^{-1}, with both α-amylase and glucoamylase activities.

Interestingly, the domain structure of both glucoamylase and cellobiohydrolase can be divided in three parts: a N-terminal catalytic domain, a C-terminal starch or cellulose binding domain, and a flexible O-glycosylated linker region which separates both N- and C-terminal domains. The C-terminal starch binding domain can be efficiently replaced by the heterologous protein (Broekhuijsen et al. 1993; Jeenes et al. 1993; Ward et al. 1995; Nyyssönen et al. 1995), although also fusions to full length GLA have been successful (Ward et al. 1990; Contreras et al. 1991; Roberts et al. 1992). It has been suggested that the positive effect of the fusion is probably because the linker region permits independent folding of the catalytic and substrate binding domains. Although glucoamylase from *A. oryzae* lacks the O-glycosylated linker region, Tsuchiya et al. (1994) showed that fusions with bovine prochymosin could also be used to produce chymosin (Table III). However, the chymosin levels obtained with strains containing a fusion with the total *glaA* gene did not exceed the levels of a strain in which prochymosin was fused to the *glaA* transcription control region. When a truncated *glaA* gene was used, lacking the starch binding domain, chymosin levels improved 5-fold in shake flask cultures, and 500-fold in a solid-state fermentation. In all cases described above the heterologous genes were fused to highly expressed fungal genes. Baron et al. (1992) showed that the use of fungal genes is not necessarily required. Fusion of human lysozyme to the 3' end of the bacterial *Streptoalloteichus hindustanus* phleomycin resistance gene, provided with a synthetic secretion signal, resulted in a 10 - 150 fold increase in lysozyme protein levels in *Tolypocladium geodes*, as compared to non-fused lysozyme (Table III).

In general, the N-terminal fungal protein is believed to serve as a carrier, improving the translocation of the protein into the ER, to aid folding and to protect the heterologous protein from degradation. Further along the secretory pathway, in most cases the fusion protein is cleaved, resulting in the secretion of the separate proteins. Cleavage have been shown to occur either by autocatalytic processing of the heterologous protein (Ward et al. 1990), an unknown fungal protease (Roberts et al. 1992, Baron et al. 1992, Nyyssönen et al. 1993, Nyyssönen et al. 1995) or by a KEX2-like protease, for which a recognition site had been

Table III. **Production of heterologous proteins by filamentous fungi using a gene-fusion strategy.**

Host	promoter	pre(pro)seq.	"carrier"-gene	KEX2	protein level	improvement factor	reference
BOVINE PROCHYMOSIN							
A. nidulans	*A. niger glaA*	*glaA*	–		3 mg l^{-1}		Cullen et al. 1987
A. niger	*A. niger glaA*	chymosin	–		6.2 mg l^{-1}		Van Hartingsveldt et al. 1990
	A. niger glaA	*glaA*	–		11.3 mg l^{-1}		"
	A. niger glaA	*glaA* (prepro)	–		4.1 mg l^{-1}		"
	A. niger glaA	*glaA* (prepro)	*A. niger glaA*$_{1-71}$	–	10.2 mg l^{-1}		"
A. awamori	*A. niger glaA*	*glaA*	–		8 mg l^{-1}		Ward et al. 1990
	A. awamori glaA	*glaA* (prepro)	*A. awamori glaA*$_{1-614}$	–	140 mg l^{-1}	10 - 20 x	"
	A. awamori glaA	*glaA* (prepro)	*A. awamori glaA*$_{1-614}$	–	1 g l^{-1} [a,b]	100 x	Dunn-Coleman et al. 1991
A. oryzae	*A. oryzae glaA*	chymosin	–		70 μg l^{-1}		Tsuchiya et al. 1994
	A. oryzae glaA	*glaA* (prepro)	*A. oryzae glaA*$_{1-603}$	–	30 - 80 μg l^{-1}		"
	A. oryzae glaA	*glaA* (prepro)	*A. oryzae glaA*$_{1-511}$	–	0.1 - 0.3 mg l^{-1}	5 x	"
	A. oryzae glaA	*glaA* (prepro)	*A. oryzae glaA*$_{1-511}$	–	150 mg/kg [d]	500 x	"
A. awamori	*A. awamori amyA/B*	*amyA/B*	*A. awamori amyA/B*	–	50 mg l^{-1}		Korman et al. 1990
Fab ANTIBODY FRAGMENTS							
T. reesei	*T. reesei cbh1*	*cbh1*	–		0.3 - 1 mg l^{-1}		Nyyssönen et al 1993; 1995
	T. reesei cbh1	*cbh1*	*T. reesei cbh1*$_{1-466}$	–	5 - 40 mg l^{-1}	5 - 100 x	"
	T. reesei cbh1	*cbh1*	*T. reesei cbh1*$_{1-466}$	–	50 - 150 mg l^{-1} [c]	> 150 x	"

Table III (continued)

Host	promoter	pre(pro)seq.	"carrier"-gene	KEX2	protein level	improvement factor	reference
HEN EGG WHITE LYSOZYME							
A. niger	*A. niger glaA*	HEWL	_		12 mg l^{-1}		Archer et al. 1990b
	A. nidulans gpdA	HEWL	_		1 mg l^{-1}		"
	A. niger glaA	HEWL	_		50 mg l^{-1}		Jeenes et al. 1993
	A. niger glaA	*glaA* (prepro)	*A. niger glaA*$_{1-498}$	+	1 g l^{-1}	> 20 x	"
HUMAN INTERLEUKIN-6							
A. nidulans	*A. niger glaA*	*glaA*	_		25 μg l^{-1}		Carrez et al. 1990
	A. niger glaA	hIL6	_		<1 μg l^{-1}		"
	A. niger glaA	*glaA* (prepro)	_		2 μg l^{-1}		"
	A. niger glaA/ *A. nidulans gpdA*	*glaA*	*A. niger glaA*$_{1-614}$	+	5 mg l^{-1}	200 x	Contreras et al. 1991
A. niger	*A. niger glaA*	*glaA*	_		ND		Carrez et al. 1990
	A. niger glaA	hIL6	_		ND		"
	A. niger glaA	*glaA* (prepro)	_		ND		"
A. niger	*A. nidulans gpdA*	hIL6 (prepro)	_		ND [a]		Broekhuijsen et al. 1993
	A. nidulans gpdA	*glaA* (prepro)	_		ND [a]		"
	A. nidulans gpdA	*glaA* (prepro)	*A. niger glaA*$_{1-514}$	+	15 mg l^{-1} [a]	> 1000 x	"
	A. nidulans gpdA	*glaA* (prepro)	*A. niger glaA*$_{1-514}$	_	> 40 mg l^{-1} [a]	> 1000 x	"
A. nidulans	*A. nidulans alcA*	?	_		< 1 - 5 mg l^{-1}		Hintz et al. 1995
	A. nidulans alcA [e]	?	*A. niger glaA*	+	> 100 mg l^{-1}	20 - 100 x	"

Table III (continued)

Host	promoter	pre(pro)seq.	"carrier"-gene	KEX2	protein level	improvement factor	reference
HUMAN LACTOFERRIN							
A. nidulans	*A. nidulans alcA*	lactoferrin	_		5 mg l^{-1}		Ward et al. 1992a
A. oryzae	*A. oryzae amy*	*amy*	_		25 mg l^{-1}		Ward et al. 1992b
A. awamori	*A. awamori glaA*	*glaA* (prepro)	*A. awamori* $glaA_{1-498}$	+	> 250 mg l^{-1}	> 10 - 50 x	Ward et al. 1995
A. awamori	*A. awamori glaA*	*glaA* (prepro)	*A. awamori* $glaA_{1-498}$	+	> 2 g l^{-1} [b]	> 80 - 400 x	"
HUMAN LYSOZYME (HLz)							
T. geodes	*T. reesei* TR1	HLz			0.5 - 2 mg l^{-1}		Baron et al. 1992
	T. reesei TR1	Synthetic prepro			0.25 - 2 mg l^{-1}		"
	T. reesei	Synthetic prepro	*S. hindustanus ble*	_	10 - 150 mg l^{-1}	10 - 150 x	"
PORCINE PANCREATIC PROPHOSPHOLIPASE A2							
A. niger	*A. niger glaA*	PLA2	_		ND [a]		Roberts et al., 1992
	A. niger glaA	*glaA* (prepro)	*A. niger* $glaA_{1-614}$	_	10 mg l^{-1} [a]		"

ND Not Detectable

[a] Protease mutant strain

[b] High levels obtained by a series of mutagenesis

[c] Submerged fermentations

[d] Solid state fermentation

[e] Similar results were obtained when a glucose-derepressed *alcA* tcr was used, but maximal levels were obtained earlier

introduced specifically in the fusion protein (Contreras et al, 1991; Broekhuijsen et al, 1993, Ward et al. 1995).

Although the strategies described here have been successful in increasing the protein levels of heterologous proteins, in a lot of cases the protein levels are still lower than those obtained with homologous proteins and classical strain improvement programs (random mutagenesis and selection) had to be used to further improve the yields (Dunn-Coleman et al. 1991; Ward et al. 1995). This demonstrates the need to gain more knowledge concerning the parameters that influence heterologous protein production.

V. OUTLINE OF THIS THESIS

As described in this chapter, production levels of heterologous proteins are usually two to three orders of magnitude lower than the levels of homologous proteins. Unfortunately, most data available in the literature are obtained from undefined multi copy strains. The use of different expression cassettes, different strains and the lack of data like mRNA levels makes it difficult to compare the results and draw conclusions with respect to the factors that are responsible for the lower production levels of heterologous proteins. To obtain more insight in these factors a study has been started in which the production levels of a number of different proteins were systematically analyzed in defined isogenic single copy strains of *A. awamori*.

To facilitate efficient selection of transformants containing a single copy of a vector integrated at a specific locus, a transformation system was developed, similar as previously was described for *Aspergillus* species like *A. niger* and *A. nidulans* (Hamer and Timberlake, 1987; Van Gorcom and van den Hondel, 1988). This system (Chapter 2) is based on the transformation of a *pyrG* mutant strain with a vector containing a *pyrG* gene with a mutation that is localized opposite to the one present in the genome. A new method has been developed for the construction of *pyrG* mutants which allows the introduction of specific mutations in the target locus, resulting in defined mutant strains.

For the production of proteins in large quantities efficient gene expression systems are also required. During the years a number of such expression systems has been developed for *Aspergillus* species, mainly based on the *A. niger* or *A. awamori* glucoamylase (*glaA*) tcr. Although the existing tcrs are useful for high level expression of genes, there still is a need for new, efficiently inducible tcrs, especially when they can be induced under conditions different from the expression conditions for already available tcrs. Studies, carried out at the

TNO laboratory in Rijswijk, indicated that the tcr of the 1,4-ß-endoxylanase A gene (*exlA*) could be used for overexpression of its own gene product 1,4-ß-endoxylanase A (Hessing et al. 1994). To further investigate the usefulness of this tcr for the expression of other genes, research was started to study induction of the *exlA* tcr in shake-flask cultures (Chapter 3). To investigate the kinetics of synthesis of mRNAs and proteins driven by the *A. awamori exlA* tcr in more detail, continuous culture fermentations were performed (Chapter 6).

To obtain more insight in the parameters determining the production efficiency of secreted heterologous proteins by filamentous fungi, in Chapter 4 a systematic analysis is described which was carried out to compare the production levels of a homologous protein with those of different heterologous - fungal and non-fungal - proteins. All expression cassettes were integrated in a single copy at the *pyrG* locus (see Chapter 2) and transcription of the genes was controlled by the expression signals of the *A. awamori* 1,4-ß-endoxylanase A (*exlA*) gene (see Chapter 3). Analysis showed that large differences occurred in the steady-state mRNA levels transcribed from the various genes, varying from high levels for genes of fungal origin (*A. awamori* 1,4-ß-endoxylanase A, *A. niger* glucoamylase and *Thermomyces lanuginosus* lipase) to low or undetectable levels for genes of non-fungal origin (human interleukin-6 and *Cyamopsis tetragonoloba* (guar) α-galactosidase). In all cases except for hIL6, the protein levels corresponded to the amount expected on basis of the mRNA levels. For hIL6, very low protein levels were observed whereas relatively high steady-state mRNA levels were obtained.

In Chapter 5 the experiments are described that were carried out to further investigate the problems observed with hIL6 and α-galactosidase production by making *glaA* gene-fusions. The effect of a gene-fusion was investigated at the mRNA and at the protein level, by comparison with the corresponding non-fused genes, both in defined single copy strains.

With one gene, guar α-galactosidase, full length mRNA was only detected when a synthetic α-galactosidase gene was used with a different codon usage than the wild-type gene (Chapter 4). Fusions with the wild-type *aglA* gene resulted in truncated mRNA lacking most of the *aglA* sequences. Experiments were carried out to identify the reason for this truncation (Chapter 5).

Curr. Genet. (1995) **27**:536-540

CHAPTER 2

A novel strategy for the isolation of defined *pyrG* mutants and the development of a site-specific integration system for *Aspergillus awamori*

Robin J. Gouka, Johanna G.M. Hessing, Hein Stam*, Wouter Musters* and Cees A.M.J.J. van den Hondel

Department of Molecular Genetics and Gene Technology, TNO Nutrition and Food Research Institute, PO Box 5815, NL-2280 HV Rijswijk, the Netherlands and *Department of Gene Technology and Fermentation, Unilever Research Laboratorium, Olivier van Noortlaan 120, NL-3133 AT Vlaardingen, the Netherlands

SUMMARY

A homologous gene transfer system for *Aspergillus awamori* for site-specific integration is described, based on two components. First, a defined *A. awamori pyrG* mutant strain constructed by a selection strategy for gene-replacement in fungi. Second, a vector with a homologous *pyrG* selection marker containing a defined mutation at a site different from that of the mutation in the *pyrG* gene of the defined mutant strain. Defined mutations in the *A. awamori pyrG* gene, isolated from a genomic library by heterologous hybridisation with the *A. niger pyrG* gene as a probe, were introduced by specifically altering sequences at restriction sites in the coding region of the gene. After transformation of the *A. awamori* wild-type strain with vectors containing these mutated *pyrG* genes, and selection for 5-fluoro-orotic acid resistance (5-FOA^R), on the average 60 percent of the 5-FOA^R colonies originated from replacement of the wild-type *pyrG* gene by the mutated *pyrG* allele. After transformation of a mutant strain, carrying a mutation near the 5' end of the *pyrG* gene with vectors containing a mutation near the 3' end of the *pyrG* gene, 35% of the resulting transformants contained one copy of the vector at the *pyrG* locus.

INTRODUCTION

Filamentous fungi of the genus *Aspergillus* are widely used for the commercial production of extracellular proteins due to their ability to secrete these in large amounts. For homologous proteins production levels of grams per litre have been reported, e.g. for *Aspergillus niger* glucoamylase and *Aspergillus oryzae* α-amylase (Verdoes et al. 1995). In contrast, production levels of heterologous proteins are usually orders of magnitude lower. At best, the extracellular yields of foreign gene products reach 50 mg l^{-1}, even in optimized situations (Van den Hondel et al. 1991). In our laboratory a study has been initiated to identify factors that cause these differences in homologous and heterologous protein production levels in filamentous fungi. In this study the expression of genes of different origin - fungal and non-fungal - will be compared. The results might lead to the identification of factors that cause these differences.

As in our laboratory good experience was obtained with *Aspergillus awamori* for the overproduction of a 1,4-ß-endoxylanase (Hessing et al. 1994) and also in other laboratories *A. awamori* has been used for (over)production of proteins like bovine chymosin and *A. niger* glucoamylase (reviewed by Verdoes et al. 1995), we decided to use *A. awamori* as a host organism for our study.

For a reliable comparison of production levels of different proteins it is a prerequisite that the strains to be analyzed carry the different expression(-analysis) vectors in a single copy situation and in an identical chromosomal environment, as both copy number and chromosomal environment were shown to influence gene expression (Van Gorcom et al. 1985, Miller et al. 1987). Expression(-analysis) systems have been developed for *Aspergillus* species like *A. niger* and *A. nidulans* which facilitate efficient selection of transformants containing a single copy of an expression(-analysis) vector integrated at a specific locus. These systems are based on an *A. niger* mutant *pyrG* gene (Van Gorcom and Van den Hondel, 1988) or *A. nidulans* mutant *argB* genes (Hamer and Timberlake, 1987, Punt et al. 1990) as selection marker and the corresponding *pyrG* or *argB* mutant strains as recipients. In this paper the development of a similar system for *A. awamori* is described. The system is based on the transformation-mediated isolation of defined *A. awamori pyrG* recipient strains and the efficient selection of *A. awamori* transformants with one or more copies of an expression vector integrated at the *pyrG* locus.

MATERIALS AND METHODS

Strains, media and plasmids. The *A. awamori* strain ATCC 11358 (CBS 115.52) has been used as a recipient strain for transformation. *Escherichia coli* strain JM109 (Sambrook et al. 1989) was used for propagation of plasmids. *A. awamori* was cultivated in Aspergillus minimal medium (MM; Bennett and Lasure 1991) or enriched MM (MM supplemented with 5 g l^{-1} yeast extract (Difco) and 0.1% vitamin-free casamino acids (Difco)). Plasmid pUC19 (New England Biolabs) was used for subcloning of genomic fragments. Plasmid pAB4-1 comprises the *A. niger pyrG* gene (Van Hartingsveldt et al. 1987).

Transformation and selection of *A. awamori* strains. Protoplasting and transformation of *A. awamori* was performed as described by Punt and Van den Hondel (1993). *A. awamori Pyr*$^-$ transformants were selected on osmotically stabilized (1.2 M sorbitol) MM with 10 mM proline as N-source and supplemented with 0.1% casamino acids, 10 mM uridine, 0.75 g l^{-1} 5-fluoro-orotic acid (Sigma), and 1.5% bacteriological agar (Oxoid). *A. awamori* Pyr$^+$ transformants were selected on osmotically stabilized (1.2 M sorbitol) agar plates containing MM.

DNA manipulations. Standard recombinant DNA techniques were used for cloning and analysis of the *A. awamori pyrG* gene (Sambrook et al. 1989). *Aspergillus* chromosomal DNA was isolated as described by Kolar et al. (1988). Colony blot hybridization experiments for *A. awamori* were carried out according to Kinsey (1989).

RESULTS AND DISCUSSION

Isolation and characterization of the *A. awamori pyrG* gene

For the development of a site-specific integration system for *A. awamori* we chose to use the homologous *pyrG* gene as selection marker. Therefore an *A. awamori* genomic library in λEMBL3 (Hessing et al. 1994) was screened with the 3.8 kb *Xba*I fragment from pAB4-1, containing the *A. niger pyrG* gene (Van Hartingsveldt et al. 1987), as a probe. This screening resulted in the isolation of ten hybridizing phage clones that contained the putative *pyrG* gene. Southern blot analysis of DNA of five λ phages, using the *A. niger pyrG* fragment as a probe, revealed that all of them contained DNA fragments that were also present in corresponding digests of *A. awamori* chromosomal DNA. A restriction map of the genomic

region is shown in Fig.1A. A 2.4 kb *Bam*HI/*Hin*dIII fragment, a 3.0 kb *Bam*HI fragment and two *Sal*I fragments of 3.3 and 3.8 kb respectively, all hybridizing with the 3.8 kb *A. niger pyrG* fragment, were isolated and subcloned in pUC19, resulting in the vectors pAW4-1 to pAW4-4 (Fig.1A). After transformation of the *A. niger pyrG* mutant strain AB4.1 (Van Hartingsveldt et al. 1987) with the vectors pAW4-1 and pAW4-2 Pyr^+ colonies were obtained with a frequency of 2-4 colonies per μg DNA, whereas with vectors pAW4-3 and pAW4-4 no Pyr^+ colonies were observed.

A large part of the DNA sequence of the putative *pyrG* gene was determined. Comparison of the *A. awamori* nucleotide sequence with the *A. niger pyrG* nucleotide sequence showed a similarity of 97%. The deduced aminoacid sequence showed a similarity of 99%. From this comparison the position of the gene and the direction of transcription could be deduced (Fig.1A).

Construction of defined *A. awamori pyrG* mutant strains

Construction of defined *A. awamori pyrG* mutant strains was performed in two steps. First, defined mutations in the cloned *A. awamori pyrG* gene were introduced *in vitro*. To obtain a vector with a defined mutated *A. awamori pyrG* gene, pAW4-1, was digested with *Sal*I that cuts near the 3' end in the coding region approximately 0.6 kb downstream of the ATG codon. Treatment of the *Sal*I digest with T4 DNA polymerase to fill in the sticky ends and subsequent ligation resulted in the isolation of plasmid pAW4-10. In this plasmid the *Sal*I site (gtcgac) was converted to the sequence gtcgatcgac, introducing a *Pvu*I site, resulting in a frame-shift mutation in the coding sequence. With the same method, another mutation was introduced near the 5' end of the *pyrG* gene. For this purpose pAW4-2 was digested with *Bgl*II that cuts in the coding region at 144 nucleotides downstream of the ATG codon as was shown by sequence analysis. One of the clones obtained after transformation of *E. coli*, named pAW4-20, appeared to contain a deletion of 49 nucleotides (from nucleotide 109 up to 157 downstream of the ATG codon; Fig. 1B). As was shown by nucleotide sequence analysis, this deletion resulted in a frame-shift mutation in the coding sequence. Since the mutation caused by the deletion was likely to be more stable than the one due to the filled-in *Bgl*II site, it was decided to use the vector with the deletion. In the next step of the construction of a *pyrG* mutant strain, the mutated *pyrG* genes were used to replace the chromosomal wild-type *pyrG* gene by transformation and subsequent selection of 5-FOA^R transformants. Since $PyrG^-$ strains are resistant to 5-FOA (Boeke et al. 1984) *pyrG*$^-$ transformants might be selected directly from wild-type strains. Transformation of the wild-type

A. awamori strain was performed with *Bam*HI/*Hin*dIII digested pAW4-10 or *Bam*HI digested pAW4-20. Per transformation 2.10^6 protoplasts were transformed with 10 μg of DNA. With both vectors a number of 5-FOA^R colonies were obtained on plates supplemented with 5 mM uridine and 0.75 g l^{-1} 5-FOA. In three transformations in total eleven 5-FOA^R colonies were obtained with 30 μg of pAW4-10 DNA and 16 with 30 μg of pAW4-20 DNA. Although the number of 5-FOA^R colonies obtained was not significantly higher than that obtained on the control plates without DNA, all of the 27 5-FOA^R colonies were analyzed further. Of the total number of 27 5-FOA^R colonies twenty were not able to grow on MM plates without uridine, corresponding with a *pyr*$^-$ phenotype. Southernblot analysis was carried out with genomic DNA of all 20 5-FOA^R, uridine requiring colonies (eight obtained with pAW4-10 and twelve with pAW4-20). DNA was digested with *Sal*I or *Bgl*II and probed with a 3.0 kb *Bam*HI fragment from pAW4-1 containing the *A. awamori pyrG* gene (Fig.1C). In the DNA of three (#2, 3, and 8) of the eight colonies obtained with pAW4-10, both *Sal*I fragments of 3.3 and 3.8 kb, present in the wild-type strain, were replaced by one hybridizing fragment of approximately 7 kb. This result indicates that these mutants originate from a replacement of the wild-type *pyrG* gene by the mutated *pyrG* gene from pAW4-10. One of these *pyrG* mutant strains was designated AW4.10. Four transformants had a similar hybridization pattern as observed for the wild-type strain and thus are likely to be the result of a spontaneous mutation. In the remaining transformant (#6) both *Sal*I fragments were replaced by a fragment of 4.5 kb which might be a deletion. In DNA of nine of the twelve colonies obtained with pAW4-20 both *Bgl*II fragments of approximately 9 and 2.7 kb, present in wild-type DNA, were replaced by a single hybridizing fragment of approximately 12 kb, confirming that in those colonies the wild-type *pyrG* gene was replaced by the mutated *pyrG* gene present on pAW4-20. One of those strains was designated AW4.20. In one mutant (#17) a slightly smaller *Bgl*II fragment was visible indicating a deletion. The remaining two mutants (#25 and 27) are likely to be the result from spontaneous mutations as they have a hybridization pattern similar to the wild-type.

Taken together, these results indicate that defined *pyrG* mutant strains can be obtained very efficiently by replacing the wild-type gene for a mutated gene using transformation and subsequent selection on medium containing 5-FOA. Although the number of colonies obtained on transformation plates was not significantly elevated compared to the number of colonies obtained on control plates, on the average 60 percent of the transformants resulted from gene-replacement.

These results are similar to those obtained by Alic et al (1993), however, in our approach no second selection marker is needed, which allows us to use the approach in wild-type

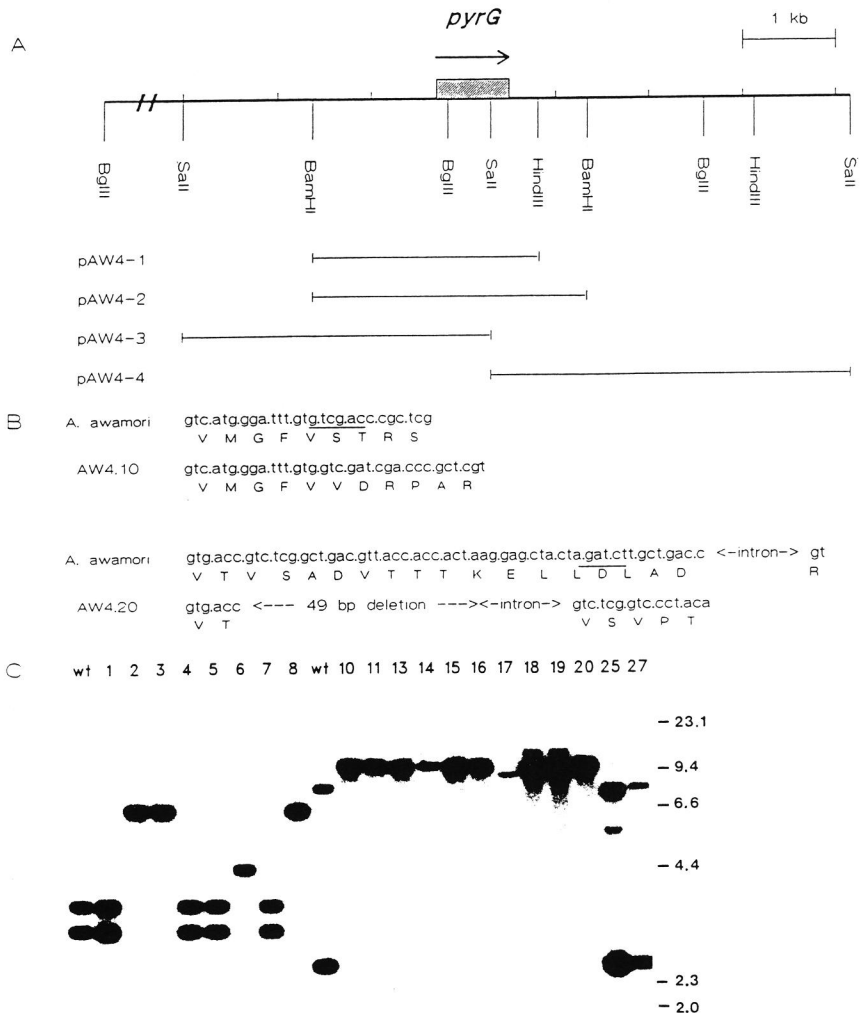


Fig.1A. Restriction map of the *A. awamori* wild-type genomic region containing the *pyrG* gene encoding orotidine-5'-monophosphate decarboxylase. The position of the gene is indicated by a dashed box. The direction of transcription is indicated by an arrow. The length of the most 5' located *Bgl*II-*Sal*I fragment, indicated by -//-, is approximately 7 kb. The fragments which were subcloned in pUC19 resulting in the vectors pAW4-1 to pAW4-4 are shown below.

strains.

The selection strategy to obtain defined mutants is not restricted to *pyrG* mutants, but can also be extended to other genes of which a selection method for mutants is available. For example selection on chlorate for nitrate reductase (*niaD*) mutants (Gouka et al. 1991) and selection on fluoro-acetate for acetyl-coenzyme A synthetase (*facA*) mutants (Gouka et al. 1993).

Site-specific integration of vectors at the *A. awamori pyrG* locus

Site-specific integration of vectors at the *pyrG* locus was analyzed by transformation of *A. awamori* strain AW4.20 (containing a defined mutation at the *Bgl*II site) with vectors containing the mutated *pyrG* gene from pAW4-10 (with a defined mutation at the *Sal*I site) as selection marker. This mutated *pyrG* gene was preferred as selection marker over a wild-type *pyrG* gene since in related species it has been demonstrated that integration of one copy of a vector at a specific locus is enhanced when using a mutant selection marker (Van Gorcom and Van den Hondel 1988; Van den Hondel and Punt 1991).

In total 151 Pyr^+ transformants of *A. awamori* AW4.20 were obtained, at a frequency of 0.1-0.2 transformants per μg. This number was 10 times less compared to the number of transformants obtained with vectors containing the wild-type *pyrG* gene. A similar difference in transformation frequency between a vector with a mutated or a wild-type gene as selection marker is also observed in other systems (Van Gorcom and Van den Hondel 1988; Punt et al. 1990).

Fig.1B. Part of the nucleotide and deduced amino acid sequence of the *pyrG* gene of the *A. awamori* wild-type strain and mutant strains AW4.10 (upper part) and AW4.20 (lower part). In AW4.10 the *Sal*I site (gtcgac, underlined) in the wild-type strain, present approximately 0.65 kb downstream of the ATG codon, is converted into the sequence 5'-gtcgatcgac resulting in a frame shift. The sequence around the *Bgl*II site (underlined), corresponding to the nucleotides 103 to 160 of the coding region of the *A. awamori pyrG* gene is given in the lower part. The 49 basepairs deletion of AW4.20 is shown below. The position of the intron is indicated between arrows.

Fig.1C. Southernblot analysis of *A. awamori* wild-type and twenty *A. awamori* $PyrG^-$ transformants obtained with pAW4-10 or pAW4-20. The numbers in the lanes correspond with transformant numbers in the text. DNA was digested with *Sal*I (wt and no. 1-8) or *Bgl*II (wt and no. 10-27), separated on an 0.7% agarose gel, transferred to nitrocellulose and probed with ^{32}P-labelled pAW4-1 DNA. Markers for molecular size in kb are indicated on the right.

Table I. Results of Colony blot and Southern analysis of *A. awamori* transformants obtained with vectors containing a mutated *pyrG* gene as selection marker

Exp[1].	No. of transformants analyzed	Type of integration[2] A	B	C	D
1	17	6	9	2	0
2	23	12	6	4	1
3	8	4	2	2	0
4	13	9	3	1	0
5	16	8	4	2	2
6	10	4	6	0	0
7	32	21	8	2	1
8	32	13	15	4	0
Total	151	77 (51%)	53 (35%)	17 (11%)	4 (3%)

[1] In all experiments the mutant *pyrG* gene containing the filled in *Sal*I site was present as selection marker on the vector.

[2] Transformants were classified into four categories; A, gene-replacement or gene-conversion; B, single copy integration of the vector at the *pyrG* locus; C, multicopy integration of the vector at the *pyrG* locus; D, transformants with a hybridization pattern that could not be explained, although integration at the *pyrG* locus had occurred as was shown by the disappearance of the hybridizing band in the host strain.

Between ( ) the percentage of transformants is given for accumulated data.

To have a first discrimination between transformants obtained by gene-conversion/gene-replacement and transformants obtained by integration of the vector at the *pyrG* locus, a colony blot hybridization was carried out using pUC19 DNA, present in all vectors, as a probe. As shown in Table I 77 transformants, being 51%, contained no vector DNA indicating a gene-replacement/gene-conversion. Genomic DNA of the remaining 74 transformants was isolated to analyze the integration pattern. Hybridization was carried out using vector pAW4-1 as a probe. From Table I, which shows the data obtained from this experiment, it can be concluded that 70% of the integration-type transformants contained one copy of the vector at the *pyrG* locus, whereas 23% contained more than one copy. The remaining 7% displayed

a hybridization pattern that was not easily explained, although it was clear that integration had occurred at the *pyrG* locus.

From these data it is clear that with the use of a defined *pyrG* mutant strain and a mutant selection marker with a relatively high frequency site specific single copy transformants can be isolated.

Currently we are using the *pyrG* based integration strategy to study heterologous gene expression of strains containing one copy of different expressioncassettes at the *pyrG* locus.

Acknowledgements

The experiments described in this paper were carried out as contract research for Unilever Research Laboratories. Prof. Dr. P.H. Pouwels and Dr. P.J. Punt are greatly acknowledged for critically reading the manuscript.

Appl Microbiol Biotechnol, in press

CHAPTER 3

An expression system based on the promoter region of the *Aspergillus awamori* 1,4-ß-endoxylanase A gene

Robin J. Gouka[1], Johanna (H.) G.M. Hessing[1], Peter J. Punt[1], Hein Stam[*2], Wouter Musters[2] and Cees A.M.J.J. van den Hondel[1]

[1] TNO Nutrition and Food Research Institute, Department of Molecular Genetics and Gene Technology, P.O. Box 5815, NL-2280 HV Rijswijk, the Netherlands

[2] Unilever Research Laboratorium, Olivier van Noortlaan 120, NL-3133 AT Vlaardingen, the Netherlands

[*] Present address: Quest International, P.O. Box 2, NL-1400 CA Bussum, the Netherlands

SUMMARY

A new, highly inducible fungal promoter derived from the *Aspergillus awamori* 1,4-ß-endoxylanase A (*exlA*) gene is described. Induction analysis, carried out with the wild type strain in shake flasks showed that *exlA* expression is regulated at the transcription level. Using a ß-glucuronidase (*uidA*) reporter strategy, D-xylose was shown to be an efficient inducer of the *exlA* promoter, whereas sucrose or maltodextrin were not. Upon D-xylose induction the *exlA* promoter was threefold more efficient than the frequently used *Aspergillus niger* glucoamylase (*glaA*) promoter under maltodextrin induction. Detailed induction analyses demonstrated that induction was dependent on the presence of D-xylose in the medium. C-limited chemostat cultures with the *uidA* reporterstrain showed that D-xylose was also a very good inducer in a fermenter, even in the presence of sucrose.

INTRODUCTION

Filamentous fungi, such as *Aspergillus* and *Trichoderma* species, have a number of favourable characteristics which makes them attractive hosts for the large scale production of proteins. These characteristics include (a) their capacity to secrete substantial amounts of proteins into the medium (reviewed by a.o. Fowler and Berka 1991; Van den Hondel et al. 1991; Verdoes et al. 1995), (b) well-established and safe use of the enzymes they produce in food and food processing industry and (c) their long history in large scale fermentation processes which resulted in a considerable knowledge in this field.

For the production of proteins in large quantities efficient gene expression systems are required. During the years a number of such expression systems has been developed for *Aspergillus* and *Trichoderma* species. One of the most frequently used expression systems in the black *Aspergilli* is based on the *A. niger* or *A. awamori* glucoamylase (*glaA*) promoter. This promoter was successfully utilized to produce heterologous proteins like bovine prochymosin (Cullen et al. 1987) and human α-interferon (Gwynne et al. 1987) in *A. nidulans* and *A. niger*. Besides the *glaA* promoter a number of other inducible promoters have been used successfully for production of homologous and heterologous proteins in *Aspergilli*. These promoters are a.o. derived from the alcohol dehydrogenase (*alcA*) gene of *A. nidulans* (Gwynne et al. 1987), the *A. oryzae* α-amylase (*amy*) gene (Christensen et al. 1988), the *A. nidulans* aldehyde dehydrogenase (*adhA*) gene (Upshall et al. 1987) and the constitutively expressed *A. nidulans* glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) gene (Punt et al. 1991). Although the expression systems described have been used successfully to produce both homologous and heterologous proteins, the secreted yields of most heterologous proteins are very low as compared to homologous proteins and reach levels that initially do often not exceed a few milligrams per litre. Therefore, in our laboratory, a comparative study was started to identify factors that hamper efficient expression of heterologous genes. For this study, the filamentous fungus *A. awamori*, of which has been shown that it can produce homologous proteins very efficiently (Finkelstein et al. 1989; Berka et al. 1991), was chosen as host strain. For efficient expression of the heterologous genes it was decided to develop an expression system for *A. awamori*. Although several of the promoters indicated have shown to be useful for high level expression of genes, there still is a need for new, strongly inducible promoters, especially when they can be induced under conditions different from expression conditions for already available promoters. Recent studies, carried out in our laboratory, indicated that the promoter of the 1,4-ß-endoxylanase A gene (*exlA*) can be used

efficiently for the overexpression of its gene product 1,4-ß-endoxylanase A (Hessing et al. 1994). To further investigate the usefulness of this promoter in an alternative expression system, we started research to study induction of the *exlA* promoter in detail.

MATERIALS AND METHODS

Strains and media. The *A. awamori pyrG* mutant strain AW4.20 (Gouka et al. 1995), used as a recipient strain for transformation, is a derivative of *A. awamori* ATCC 11358 (CBS 115.52). *Escherichia coli* strain JM109 (Sambrook et al. 1989) was used for propagation of plasmids. For shake flask cultivations preculture medium consisted of Aspergillus minimal medium (MM; Bennett and Lasure 1991) with 1% sucrose as carbon source and 0.1 % yeast extract (Difco). For induction medium Aspergillus MM with 5% D-xylose as carbon source and 0.1 % yeast extract was used. Medium for the batch phase of the fermentation consisted of (g l^{-1}): sucrose 5.0, NH_4Cl 9.0, KH_2PO_4 1.0, $MgSO_4.7H_2O$ 1.0, $CaCl_2.2H_2O$ 0.3, yeast extract 1.0. For induction 5 g l^{-1} of D-xylose was added.

Induction of the *exlA* promoter in shake flask cultures. 1.5 x 10^6 spores/ml were inoculated in a 0.5 l shake flask with 200 ml Aspergillus preculture medium. After 24 hours of cultivation in a rotary shaker (300 rpm, 30°C) the mycelium was harvested by filtration using a Miracloth (Calbiochem) filter, washed with MM without carbon source and transferred to a 0.5 l shake flask containing 100 ml Aspergillus induction medium.
In the experiments carried out to compare the *exlA* and the *glaA* promoter, precultivation was omitted to avoid that the *glaA* promoter would be active in preculture medium (see Table I).

Bioreactor cultivations. Fermentation experiments were carried out in a 3 litre Chemoferm glass fermenter with baffles, provided with a magnetically driven stirrer (8 blades Rushton impeller) on the bottom plate. The dissolved oxygen was measured by an Ingold O_2 probe, pH by an Ingold pH probe and temperature with an PT 100 sensor. Gelman bacterial filters were used to filter incoming air and outlet gasses before analysis. pH was regulated by 12.5% NH_4OH and 3M H_3PO_4 using a Watson Marlow peristaltic pump. Polypropylene glycol 2000 was used as antifoam. Working volume was maintained by using a teflon overflow tube. The outlet gasses were cooled by a condensor and led to the mass-spectrometer (MM8-80, VG) for measurements and calculations. The feed rate was checked with a burette during addition by a Meredos SPGL70 pump.

The set points for fermentation were: pH 4.5, stirrer speed 225 rpm, 25°C, airflow in 1.5 $l.min^{-1}$. After inoculation regulation of the pH was started at the moment it reached pH 4.5. The fermentation was switched from batch to continuous ($D=0.1\ h^{-1}$) when the culture reached the end of the exponential phase. Simultaneously, the stirrer speed was increased to 300 rpm to maintain the DO_2 above 35%.

Induction with D-xylose was started when a steady-state phase was achieved (defined as constant outlet gasses during at least 20 hours). D-xylose was added to the fermenter and to the feed vessel at the same time.

Proteolytic activity. Proteolytic activity in the induction medium was tested by incubation of low molecular weight marker proteins (LMW; Pharmacia) in medium samples taken 22, 37 and 48 hours after transfer to induction medium. Analysis of these samples after 24 hours of incubation at 37°C using SDS-PAGE and subsequent Coomassie Brilliant Blue staining showed that none of the LMW marker proteins was degraded. Acidification of the medium to pH 3 with acetic acid resulted in degradation of almost all of the LMW marker proteins within 2 hours of incubation, whereas no spontaneous degradation of the marker proteins occurred at this pH.

Transformation of *A. awamori*. For transformation of *A. awamori* AW4.20 $2x10^6$ spores were inoculated in Aspergillus MM with 10 mM uridine. After overnight growth, the mycelium was diluted 1 to 5 and cultivated for another 16-18 hours. Transformation of *A. awamori* was carried out as described by Punt and Van den Hondel (1993). Selection of *A. awamori* Pyr^+ transformants was performed on Aspergillus MM with 1.2 M sorbitol. For hygromycin B selection 10 mM uridine and 150 mg l^{-1} hygromycin B were added.

Construction of GUS reporter strains. Two reporter strains (AW16 and AWGUS64-2A) were constructed which contained the following constructs. A vector containing a P*exlA-uidA* reporter construct was obtained as follows: the *uidA* gene was isolated as 1.9 kb *Nco*I - *Afl*II fragment from a derivative of pNOM102 (Roberts et al. 1989) in which the *Nco*I site present downstream of the *uidA* stopcodon was converted to an *Afl*II site with a synthetic oligonucleotide (MBL 888: 5' catgccttaagg 3'). This fragment was ligated to a partial 7.2 kb *Bsp*HI - *Afl*II fragment from plasmid pAW14B (Hessing et al. 1994). The *Bsp*HI site which was used (*Bsp*HI has five sites in the vector) was the one that comprised the translation initiation codon (TC**ATG**A) and the *Afl*II site (*Afl*II has two sites in the vector) the one comprising the

stopcodon (CT**TAA**G). The resulting vector was called pAW15-1. The fusions were verified by sequence analysis (data not shown). For insertion of a selection marker, in pAW15-1 a *Not*I site was created by converting the *Eco*RI site present 1.2 kb upstream of the ATG codon into a *Not*I site with a synthetic oligonucleotide (MBL 52-7A: 5' aattgcggccgc 3'), resulting in pAW15-1Not. In pAW15-1Not the *A. awamori pyrG* gene with a frame-shift mutation in the coding region at the *Sal*I site (*pyrG**; Gouka et al. 1995) was cloned. This *pyrG** gene was isolated as a 2.4 kb *Not*I fragment from a derivative of pAW4-10 (Gouka et al. 1995) in which respectively an upstream *Xba*I and a downstream *Hin*dIII site were converted into a *Not*I site using synthetic adaptors. The resulting vector was called pAW16 (Fig.1A). The P*glaA-uidA* reportervector was derived from pGUS64 (Verdoes et al. 1994a). In this vector, designated pAWGUS64-2A, the *A. niger pyrG* gene was replaced by the 2.4 kb *Not*I fragment containing the *A. awamori pyrG** gene as in pAW16. Both vectors were used to transform *A. awamori* AW4.20. To allow comparison, transformants containing a single copy of the vector at the *A. awamori pyrG* locus were first identified by colony blot hybridization and further by Southern blot analysis. Two single copy transformants, AW16 and AWGUS64-2A respectively, were used for further research.

Construction of *exlA*$^-$*::uidA*$^+$ strain AW15.7. To obtain a strain in which the *exlA* coding region was replaced by the *uidA* coding region replacement vector pAW15-7 (Fig.1B) was

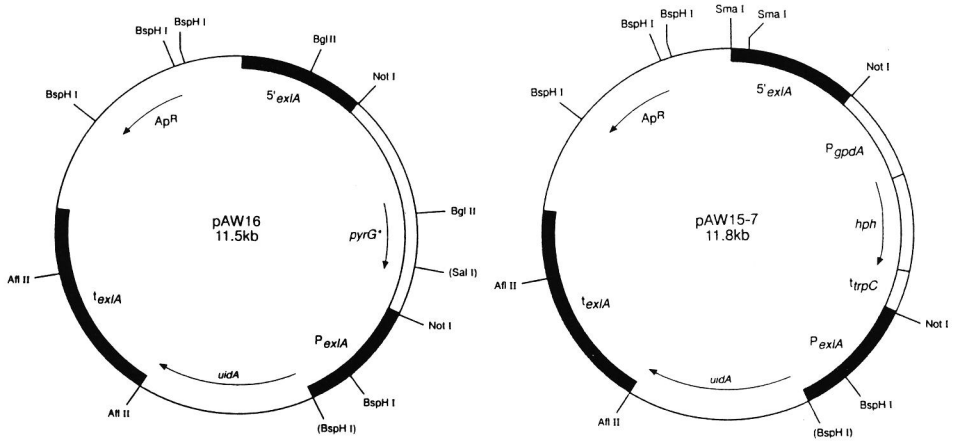


Fig. 1. Restriction map of the vectors pAW16 (A) and pAW15-7 (B). The construction of these vectors is described in M&M. Filled thick lines represent *A. awamori exlA* sequences, open thick lines *A. awamori pyrG* sequences (A) or the hygromycin resistance cassette (B) and thin lines *E. coli* DNA.

constructed as follows: in the *Not*I site of pAW15-1Not a fragment containing the hygromycin B resistance cassette from pAN7-1 (Punt and van den Hondel 1993) was cloned. This cassette, which was originally present on a 2.6 kb *Sac*I-*Hin*dIII fragment, was provided with *Not*I sites using synthetic oligonucleotides.

An *A. awamori exlA*$^-$::*uidA*$^+$, *pyrG*$^-$ strain was obtained by transforming *A. awamori* AW4.20 with replacement vector pAW15-7, which was linearized with *Sma*I. A frequency of 5 hygromycin resistant colonies per μg of DNA was obtained. Two transformants with a proper replacement, as identified by colony blot hybridization, GUS plate tests using X-gluc (5-brom-4-chlor-3-indolyl-ß-D-glucuronide) and Southern blot analysis (data not shown), were obtained out of 50 tested. One of these strains, designated AW15.7, was used for further research.

Initial analysis of AW15.7 revealed that this strain grew much poorer in the presence of uridine than *pyrG*$^+$ strains. This is in contrast to *pyrG* mutants of other *Aspergillus* species. Therefore, to acquire a replacement strain with identical growth characteristics as the wild-type strain, a *pyrG*$^+$ transformant from AW15.7 was obtained by transformation of AW15.7 with vector pAW4-10. Transformants carrying one copy of pAW4-10 at the *pyrG* locus were identified by Southern blot analysis and one of them was designated AW15.7-1.

Recombinant DNA techniques. Standard recombinant DNA techniques were used for cloning procedures (Sambrook et al. 1989). *Aspergillus* chromosomal DNA and RNA was isolated as described by Kolar et al. (1988). Colony blot hybridization of *A. awamori* was carried out according to Kinsey (1989).

ß-Glucuronidase (GUS) assays. For a qualitative GUS assay, spores were inoculated on solidified induction medium, overlaid with a Hybond filter and incubated for two days at 30°C. Then, the filter was transferred to liquid nitrogen for 30 seconds to break open the cells and developed in Z-buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 0.27% (v/v) ß-mercapto-ethanol) containing 0.04% X-gluc. Quantitative GUS assays were performed according to Roberts et al. (1989).

Xylanase activity and D-xylose concentration measurements. Total xylanase activity was determined by measuring the production of newly liberated reducing groups using the DNS (2-hydroxy-3,5-dinitrobenzoicacid) method as described by Hessing et al. (1994). D-xylose concentrations in the medium were determined according to the same method with the exception that oat-spelt xylan and the incubation step at 40°C were omitted.

RESULTS AND DISCUSSION

Induction analysis of the *A. awamori exlA* promoter

To study regulation of *exlA* gene expression, induction analysis was started with the *A. awamori* wild-type strain. The promoter was induced with D-xylose, an efficient inducer for xylanases in related species like *A. terreus* (Hrmova et al. 1991), *A. sydowii* (Ghosh and Nanda 1994) and *A. tubigensis* (De Graaff et al. 1994). In several studies focused on induction of xylanases, besides D-xylose and xylan, two well-known inducers of xylanases (reviewed by a.o. Thomson 1992; Coughlan and Hazlewood 1993), other inducing compounds have been described. These compounds are usually xylo-oligosaccharides, non-metabolized analogues of xylobiose such as ß-methyl-xyloside and/or positional isomers of xylobiose (Bailey and Poutanen 1989; Hrmova et al. 1991; Ghosh and Nanda 1994; De Graaff et al. 1994). However, all these compounds are too expensive to use in large scale cultivations. As D-xylose is a more defined chemical substance than xylan, it was chosen as inducer. Furthermore, it was observed that with xylan usually long cultivation periods are required to obtain high levels of protein produced (JGM Hessing, unpublished results).

The wild-type strain was grown according to the induction protocol described in M&M and at different time points medium and mycelium samples were taken. The medium samples were analyzed for extracellular xylanase activity (Fig.2A). From these data it is clear that after induction xylanase is secreted into the medium reaching approximately 15,000 U ml^{-1} 48 hours after transfer of the mycelium to induction medium. From data obtained by Hessing et al. (1994) it is clear that *A. awamori* produces more than one xylanase. Therefore the xylanase activity measurements (Fig.2A) did not reflect *exlA* expression alone. To determine the contribution of 1,4-ß-endoxylanase A activity to the total xylanase activity, strain AW15.7-1 was constructed, in which the *exlA* gene was replaced by the *E. coli* ß-glucuronidase (*uidA)* gene. Xylanase induction of AW15.7-1 was performed as described for the wild-type strain. As the total xylanase activity was not clearly different from the *exlA*$^+$ strain (Fig.2A), we concluded that the *exlA* gene has only a minor contribution to the total xylanase activity for the substrate used (oat spelt xylan) in the xylanase assay. These results indicated that *exlA* expression has to be investigated by other methods. Therefore, Northern blot analysis and GUS activity assays were performed.

Northern blot analysis was carried out with RNA isolated from mycelium samples taken before induction (t=0 hr) and 22, 37 and 48 hours after induction (Fig.2B). Before induction, no *exlA* mRNA could be detected. After 22 and 37 hours *exlA* mRNA was present, as was

shown by a hybridizing band of approximately 1 kb in size. However, after 48 hours this hybridizing band had disappeared again. Northern blot analysis of samples taken from AW15.7-1 showed an identical pattern for *uidA* mRNA.

To investigate whether the disappearance of *exlA* and *uidA* mRNA is related to the absence of inducer in the medium, the levels of D-xylose were determined. From Fig. 2A it is clear that within the period of 48 hours the 5% of D-xylose was almost completely consumed, leaving less than 0.05% after 48 hours. These results suggest that as soon as the fungus is deprived from D-xylose, mRNA synthesis stops and apparently the previously synthesized *exlA* or *uidA* mRNA is degraded.

From Fig.2A, it is clear that in AW15.7-1 the GUS activity increases from hardly detectable before induction to 1760 U/mg protein 48 hours after induction.

Altogether it can be concluded that regulation of the *exlA* gene is at the transcription level. From the results on GUS activity in AW15.7-1 it is clear that *uidA* expression is a good alternative to study induction of the *exlA* promoter in a more detailed manner.

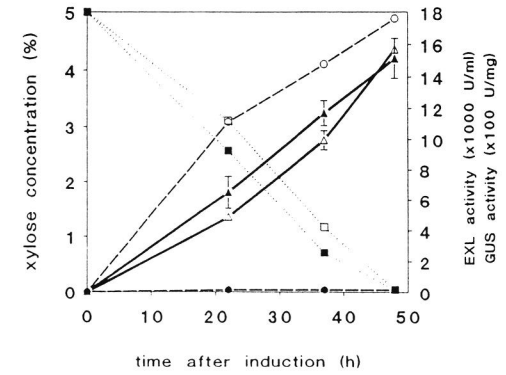


Fig.2A. Xylanase- and GUS activity, and remaining extracellular D-xylose concentration of *A. awamori* wt and AW15.7-1 after transfer of mycelium to medium with 5% D-xylose: EXL activity wt (▲) and AW15.7-1 (△), GUS activity wt (●) and AW15.7-1 (○), D-xylose conc. wt (■) and AW15.7-1 (□).

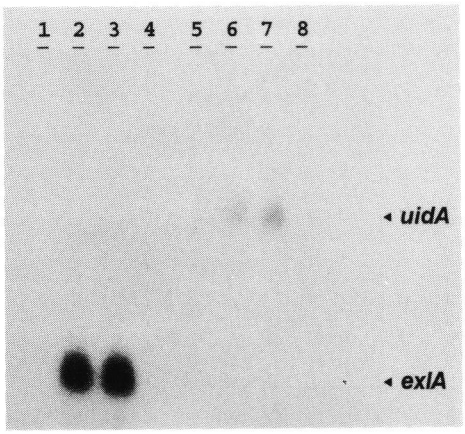


Fig.2B. Northern blot of *A. awamori* wild-type and AW15.7-1. Equal amounts of total RNA, isolated from mycelium samples before and 22, 37 and 48 hours after transfer of mycelium to medium with 5% D-xylose, were probed with a ^{32}P-labelled 0.4 kb *exlA* 5' non-coding region DNA fragment. *Lanes 1, 2, 3 and 4*: RNA from wild-type at 0, 22, 37 and 48 hours, *lanes 5, 6, 7 and 8*: RNA from AW15.7-1 at 0, 22, 37 and 48 hours, respectively.

Comparison of promoter efficiency of the *A. awamori exlA* gene and the *A. niger glaA* gene.

To compare the efficiency of the *exlA* promoter with the *A. niger glaA* promoter a reporter strategy was used. For this purpose strains AW16 and AWGUS64-2A were constructed containing respectively the *exlA* and the *glaA* promoter fused to the *uidA* gene (see Materials and Methods). Both reporter strains were cultivated in shake flasks using Aspergillus MM supplemented with 0.5% yeast extract, 0.1% casaminoacids and 5% of either sucrose, D-xylose or maltodextrin. Maltodextrin has been shown to be an efficient inducer for the *glaA* promoter (Verdoes et al. 1994a). After 48 hours mycelial extracts were prepared and GUS specific activities were determined (Table I). With the *exlA* promoter (AW16) a very low GUS activity (~ 2 U/mg) was measured with 5% sucrose or 5% maltodextrin as C-source, indicating that the *exlA* promoter is practically inactive in those media. AW16 grown in medium containing 5% D-xylose showed a GUS activity with an average of 2559 U/mg at 48 hours. With the *glaA* promoter (AWGUS64) a GUS activity of 800-900 U/mg was measured at 48 hours when grown in maltodextrin containing medium. In sucrose medium an activity of 147 U/mg was observed and in D-xylose medium a very low GUS activity. The GUS activities found for the *glaA* promoter with D-xylose and maltodextrin are in agreement with those found in *A. niger* (Verdoes et al. 1994a). The use of sucrose as inducer, which was not tested in that study, results in lower GUS activities than obtained with glucose (150 vs 1000 U/mg; Verdoes et al. 1994a).

Table I. ß-glucuronidase activities in AW16 and AWGUS64-2A strains grown in the presence of different carbon sources

Carbon source	Strains	
	AW16 (P*exlA*)	AWGUS64-2A (P*glaA*)
5% sucrose	2.3 ± 0.4	147 ± 13
5% D-xylose	2559 ± 35	6.5 ± 0.3
5% maltodextrin	1.9 ± 0.1	874 ± 50

Activities (± standard error) are given in Units per mg of total intracellular protein and are an average of 4 independent cultures.

These results indicate that the *exlA* promoter is, as the *glaA* promoter, an efficient promoter. Under inducing conditions, the *exlA* promoter results in threefold more GUS activity than the *glaA* promoter at its induced level. Moreover, the *exlA* promoter is induced in the presence of a C-source in which the *glaA* promoter in not active and vice versa.

From these data we conclude that the regulation of the *exlA* promoter is less complex than that of the commonly used *glaA* promoter since the *exlA* promoter is only active at considerable levels with a very limited set of C-sources (De Graaff et al. 1994; Ghosh and Nanda 1994), whereas the *glaA* promoter is expressed with most C-sources tested (Fowler et al. 1990; Verdoes et al. 1994a) except D-xylose.

Although a direct comparison of promoter strength is difficult, since i) both mRNAs have different 5' and 3' untranslated regions which may result in different mRNA stability and ii) the culture conditions in different growth media may affect overall cell physiology and protein synthesis, we conclude that the *exlA* promoter is a highly inducible promoter and may be a good alternative for the *glaA* promoter for the production of proteins.

Development and analysis of an induction procedure for P*exlA* driven expression

To analyze the induction of the *exlA* promoter in further detail, reporterstrain AW15.7-1, in which the *exlA* coding region is replaced for the *uidA* reportergene, was cultivated under various shake flask culture conditions. An induction method was used in which a standard amount of mycelium was transferred to induction medium (see Materials and Methods). This procedure avoids large differences between culture biomass and shortens the production phase, which is particularly important when products are toxic for the fungus or susceptible to proteolytic degradation. The experiments were carried out with 1%, 5% and 10% D-xylose as inducing C-source. Induction of the P*exlA* driven expression was analyzed by measuring GUS activity. Time-course experiments (Fig.3) showed an increase in GUS activity after transfer to D-xylose containing medium. The period and maximal level of GUS production depended on the concentration of D-xylose in the medium. In 1% D-xylose the GUS activity was maximal after 15 hours (700 U/mg; Fig.3) and then decreased. The maximal level observed in 5% D-xylose cultures was 1500 U/mg protein after 48 hours of induction, whereas after 67 hours the activity had decreased to 800 U/mg and after 137 hours to 230 U/mg. In cultures with 10% D-xylose, the prolonged presence of D-xylose resulted in extended induction as was shown by an increasing GUS activity up to 1700 U/mg, which did not decrease even after 67 hours.

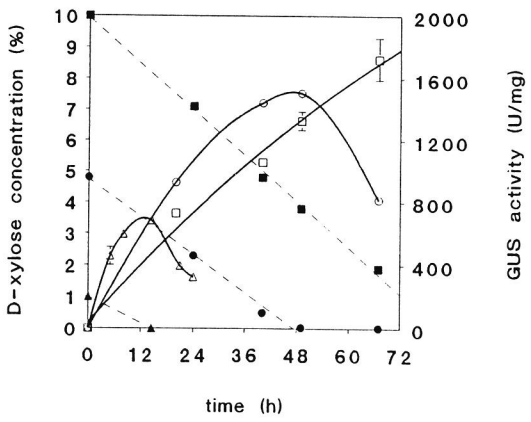


Fig. 3. ß-glucuronidase activity of *A. awamori* AW15.7-1 upon transfer into 1% (△), 5% (○) and 10% (□) D-xylose containing medium and remaining D-xylose levels in cultures with starting concentrations of 1% (▲), 5% (●) and 10% (■) D-xylose.

Determination of D-xylose concentrations in the medium showed that in cultures containing 1% D-xylose, D-xylose was depleted after about 15 hours, and in cultures with 5% D-xylose after 48 hours. In the cultures with 10% D-xylose, D-xylose was still present at 67 hours.

Northern analysis showed that *uidA* mRNA was not detectable anymore when D-xylose was depleted from the medium (data not shown, see also Fig. 2B).

Taken together, these results indicate that P*exlA* driven induction depends on the presence of D-xylose in the medium. The period and maximal level of GUS production correlated with the concentration of D-xylose in the medium. Depletion of D-xylose from the medium resulted in the disappearance of *uidA* mRNA and a decrease in GUS activity. Apparently, the existing *uidA* mRNA is rapidly degraded and no *de novo* synthesis of *uidA* mRNA occurs. The decrease in GUS activity suggests that the existing GUS protein present was degraded by intracellular proteases.

Induction analysis of the *exlA* promoter in chemostat cultures

To investigate whether the *exlA* promoter could also be induced efficiently in a fermenter, continuous culture experiments were carried out with *A. awamori* AW15.7-1. In addition, these fermenter experiments enabled us to investigate whether induction levels are maintained by a successive addition of D-xylose to the culture medium, as this was suggested by the results obtained with shake flask cultures. The experiments were performed in a 2.5 l fermenter under carbon limited conditions using D-xylose as inducer. To realize high D-xylose concentrations instantly, at t=0 hours 12.5 g D-xylose, dissolved in water, was added to the fermenter vessel. To maintain a certain level of induction also 2.5 g l^{-1} D-xylose was added to the feed. After induction a number of mycelium and medium samples were taken to analyze mRNA kinetics

and GUS protein accumulation.

Northern blot analysis with total RNA (Fig.4A), extracted from mycelium taken at several time points after induction with D-xylose, showed a quick response on *uidA* mRNA synthesis, visualized by the appearance of *uidA* mRNA already after 10 minutes. After 45 minutes the *uidA* mRNA level reached a steady-state level which was maintained during the total period of cultivation. Between 10 and 26 hours upon induction D-xylose levels were too low to be measurable (Fig.4B), indicating that the effect of the D-xylose shot had disappeared and that the culture grew under carbon limited conditions again. The presence of *uidA* mRNA at 26 hours indicated that the amount of D-xylose present in the feed, was still capable of maintaining induction. These experiments also showed that the concentration of sucrose in the feed still allowed induction.

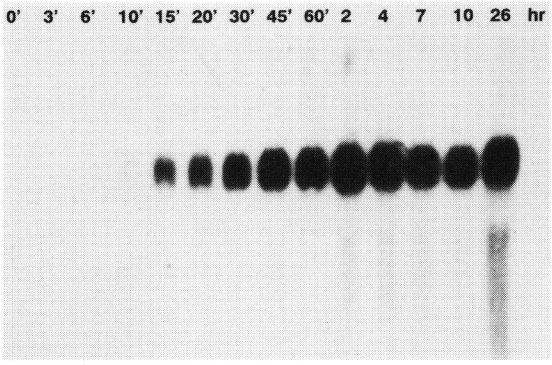


Fig. 4A. Northern blot analysis of *A. awamori* AW15.7-1 grown in a chemostat under D-xylose inducing conditions. The blot contains equal amounts of total RNA and was probed with a ^{32}P-labelled 1.9 kb *Nco*I DNA fragment from pNOM102 containing the *uidA* coding region.

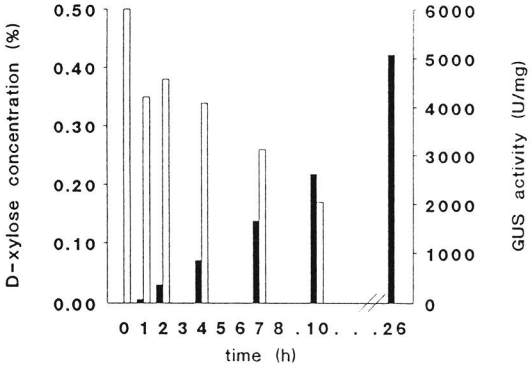


Fig. 4B. ß-glucuronidase activity of *A. awamori* AW15.7-1 (■) grown in a chemostat under D-xylose inducing conditions. D-xylose concentrations (□) are indicated on the left.

To see whether steady-state levels of GUS were obtained, at all time points the activity was determined. GUS activity (Fig.4B) became detectable 1 hour after induction increasing up to a level of 5000 U/mg protein, which was present 26 hours after induction. From an extended continuous culture experiment (data not shown) it was demonstrated that after 26 hours the level of specific GUS activity did not increase, suggesting that a steady-state in protein level was reached.

From these results it can be concluded that induction with D-xylose is also useful in a fermenter, in particular since relatively low levels of D-xylose can even induce in the presence of a cheap C-source like sucrose, which makes *exlA* induction cost effective and attractive for large scale fermentations.

Currently, we are using the *exlA* expression signals in combination with the induction protocol to study heterologous gene expression, both in shake flasks and in continuous fermenter cultures.

Acknowledgements.

This work was initially sponsored by Unilever Research Laboratorium and subsequently by a collective of Unilever, Gist-brocades, DSM, Solvay-Duphar and Ministry of Economic Affairs. Rob Muijsenberg is acknowledged for carrying out fermenter experiments.

Appl Environ Microbiol, in press

CHAPTER 4

Analysis of heterologous protein production in defined recombinant *Aspergillus awamori* strains: Comparison of expression levels of genes from fungal and non-fungal origin

Robin J. Gouka, Peter J. Punt, Johanna (H.) G.M. Hessing and Cees A.M.J.J. van den Hondel

TNO Nutrition and Food Research Institute, Department of Molecular Genetics and Gene Technology, P.O. Box 5815, NL-2280 HV Rijswijk, the Netherlands

SUMMARY

A study was carried out to obtain more insight in the parameters determining the secretion of heterologous proteins from filamentous fungi. A strategy was chosen in which mRNA and protein levels of a number of heterologous genes from different origin were compared. All genes were under control of the *Aspergillus awamori* 1,4-ß-endoxylanase A (*exlA*) expression signals and were integrated in a single copy at the *A. awamori pyrG* locus. Northern analysis showed that large differences occurred in the steady-state mRNA levels obtained from the various genes, varying from high levels for genes of fungal origin (*A. awamori* 1,4-ß-endoxylanase A, *A. niger* glucoamylase and *Thermomyces lanuginosus* lipase) to low levels for genes of non-fungal origin (human interleukin-6 and *Cyamopsis tetragonoloba* (guar) α-galactosidase). With the *C. tetragonoloba* α-galactosidase wild-type gene full length mRNA was even undetectable. Surprisingly, small amounts of full length mRNA could be detected when a *C. tetragonoloba* α-galactosidase gene with an optimized *S. cerevisiae* codon preference was expressed. In all cases except for hIL6, the protein levels corresponded to the amount expected on basis of the mRNA levels. For hIL6, very low protein levels were observed whereas relatively high steady-state mRNA levels were obtained. Our data suggest that intracellular protein degradation is the most likely explanation for the low levels of secreted human interleukin-6.

INTRODUCTION

In nature, filamentous fungi, including members of the genus *Aspergillus*, are able to use a great variety of carbon and nitrogen sources by secreting a wide range of different enzymes into their environment. This aspect, together with their capacity to secrete these enzymes in large amounts, has made filamentous fungi attractive hosts for the large scale production of proteins (reviewed by a.o. Fowler and Berka 1991; Van den Hondel et al. 1991, Verdoes et al. 1995).

In literature a large number of data concerning production of heterologous proteins in filamentous fungi has been reported (reviewed by Fowler and Berka 1991; Van den Hondel et al. 1991). From these data it is clear that, in general, the secreted yields of most heterologous proteins are low as compared to homologous proteins and reach levels that in most cases do not exceed a few milligrams per litre of culture medium. Several strategies have been developed to improve these yields, including (i) the introduction of a large number of gene copies (Verdoes et al. 1995), (ii) the use of strong fungal promoters and efficient secretion signals (Verdoes et al. 1995), (iii) gene fusions with a gene encoding a well-secreted protein (Ward et al. 1990; Contreras et al. 1991; Roberts et al. 1992; Broekhuijsen et al. 1993; Ward et al. 1995), (iv) the use of protease deficient host strains (Roberts et al. 1992; Broekhuijsen et al. 1993), (v) medium development (Archer et al. 1990a), and (vi) modification of the protein and random mutagenesis and subsequent screening for higher production levels (Dunn-Coleman et al. 1991). However, a more systematic study is necessary to identify and eliminate the factors that cause the low expression levels of heterologous proteins.

The aim of this study is to obtain more insight in the parameters that determine the production and secretion of heterologous proteins. For this study, the filamentous fungus *Aspergillus awamori*, which has been shown to be an efficient producer of proteins (Finkelstein et al. 1989; Berka et al. 1991; Hessing et al. 1994), was chosen as host strain. A strategy was developed to compare the production levels of a homologous protein with different heterologous - fungal and non-fungal - proteins. To exclude that differences in production levels are caused by differences in gene copy number or integration site, all expression cassettes were integrated in a single copy at a defined locus (Gouka et al. 1995). To avoid differences in transcriptional regulation, all genes were controlled by the expression signals of the *A. awamori* 1,4-ß-endoxylanase A (*exlA*) gene (Hessing et al. 1994).

MATERIALS AND METHODS

Strains, media and transformation. The *A. awamori* strain AW15.7 (*exlA*$^-$::*uidA*$^+$, *pyrG*$^-$, Hm^R; Gouka et al. 1996a), used as a recipient strain for transformation, is a derivative of *A. awamori* ATCC 11358 (CBS 115.52). *Escherichia coli* strain JM109 (Sambrook et al. 1989) was used for propagation of plasmids. Induction of the *A. awamori exlA* promoter was carried out as described by Gouka et al. (1996a), using a preculture medium containing 1% sucrose as C-source and an induction medium with 5% D-xylose as C-source. For transformation of *A. awamori* AW15.7 2x10^6 spores were inoculated in Aspergillus MM (Bennett and Lasure 1991) with 10 mM uridine. After 20 - 24 h of agitated growth at 30°C the mycelium was diluted 1 to 5 and cultivated for 16 - 20 h. Transformation of *A. awamori* was carried out as described by Punt and Van den Hondel (1993). Selection of *A. awamori* Pyr$^+$ transformants was performed on Aspergillus MM which was osmotically stabilized with 1.2 M sorbitol.

Construction of Aspergillus expression vectors. Expression vectors (Table I) were constructed carrying the genes coding for *A. awamori* 1,4-ß-endoxylanase A (Hessing et al. 1994), truncated *A. niger* glucoamylase G-2 form lacking the starch-binding domain (GLA_{G2}; Broekhuijsen et al. 1993), *Thermomyces lanuginosus* lipase (Boel and Huge-Jensen 1989), human interleukin-6 (Broekhuijsen et al. 1993) and *Cyamopsis tetragonoloba* α-galactosidase. For the latter enzyme two gene versions were used: the wild-type gene (Overbeeke et al. 1989) and a synthetic gene with an optimized yeast codon bias ($agl\!A_{syn}$; Verbakel 1991) differing 22% in nucleotide sequence from the cDNA clone. All genes were preceded by the *exlA* promoter and followed by the *exlA* transcription terminator. Furthermore, the genes were either expressed with the *exlA* preprosequence and/or their own preprosequence (Table I).

All vectors were based on plasmid pAW14ANot, a vector derived from pAW14B (Hessing et al. 1994). In pAW14B, the *Eco*RI site, present in the polylinker, was changed into a *Not*I site using a synthetic oligonucleotide. Two types of fusions were made for which the following restriction sites were used: (a) for translation start fusions, the *Bsp*HI site (TC**ATG**A), comprising the translation start codon ATG in pAW14ANot, was used (Hessing et al. 1994), (b) for fusions of the *exlA* preprosequence with the part of a heterologous gene encoding the mature protein, the *Nru*I site (TCGCGA), present 3 nucleotides upstream of the cleavage site of the prosequence with the mature sequence (Hessing et al. 1994), was used. In all cases fusion with the *exlA* transcription terminator was accomplished at an *Afl*II site (CT**TAA**G) that comprised the stopcodon TAA. To create correct fusions either a *Bsp*HI - *Afl*II or a *Nru*I - *Afl*II

fragment was isolated from pAW14ANot, containing the 5' and 3' *exlA* regulatory sequences together with pUC19. The 5' and 3' end sequences of the heterologous genes were subsequently adapted to these sites to regenerate a correct fusion using an appropriate restriction site in the coding sequence together with annealed synthetic oligonucleotides or fragments generated by PCR. The correct DNA sequence of all fusions was verified by sequence analysis. As selection marker, a mutant *A. awamori pyrG* gene (Gouka et al. 1995), present on a 2.4 kb *Not*I fragment, was inserted in all expression vectors. All vectors were integrated at the *pyrG* locus in *A. awamori* AW15.7 (*exlA*$^-$::*uidA*$^+$, *pyrG*$^-$, Hm^R) using the recently developed integration system by Gouka et al. (1995).

Recombinant DNA techniques. Standard recombinant DNA techniques were used for cloning procedures (Sambrook et al. 1989). *Aspergillus* chromosomal DNA and RNA was isolated as described by Kolar et al. (1988). For Northern blot analysis, as a probe a 400 bp DNA fragment was used which allowed a direct comparison of the amount of specific mRNA in the different expression strains. This 400 bp DNA fragment was isolated by PCR from a vector derived from pAW14ANot, in which the *exlA* coding region was deleted and which only contained the 5' and 3' *exlA* non-coding sequences. With this probe in each transformant the *uidA* mRNA provided with *exlA* expression signals could also be detected and used as a reference for the induction level.

Quantification of protein production levels. ß-glucuronidase assays were performed according to Roberts et al. (1989). For quantification of lipase activity the formation of free fatty acids was measured (Frenken et al. 1992). Determination of glucoamylase activity was as performed by Metwally et al (1991). α-Galactosidase activity assays were carried out essentially according to Overbeeke et al. (1989). Human interleukin-6 activity was measured with a biological assay using hIL6-dependent mouse-hybridoma cells as described by Van Snick et al. (1986).

Quantification by Western blotting was carried out according to Sambrook et al. (1989). For determination of protein concentrations purified protein was used as a standard. Purified 1,4-ß-endoxylanase A, lipase and α-galactosidase protein and the respective polyclonal antibodies were obtained from Unilever Research Laboratorium. Purified glucoamylase was obtained from Boehringer-Mannheim (lot 12005320-30). Monoclonal antibodies raised against glucoamylase were produced in our laboratory. Purified human interleukin 6, produced by Peprotech (lot nr 4602), was supplied by Sanvertech. Polyclonal antibodies against hIL6,

produced by Endogen (lot nr 403015), were obtained from Sanvertech. For detection either phosphatase labelled antibodies (Promega) or the ECL Western blot detection kit (Amersham) based on peroxidase labelled antibodies were used. Total protein amounts were determined using the BioRad protein assay kit.

Dry weight measurements. Dry weight was measured by filtration of 25 ml of homogenous culture samples over a Miracloth filter. The mycelium was squeezed and dried overnight in a vacuum oven at 80°C.

Determination of D-xylose concentrations. D-xylose concentrations were determined by measuring the reducing sugar content with the DNS (2-hydroxy-3,5-dinitrobenzoicacid) method as also used for xylanase activity assays as described by Hessing et al. (1994). Appropriate amounts of medium sample were supplemented with AD to a total volume of 0.5 ml. To this solution 0.5 ml DNS was added, mixed and incubated for 10 minutes at 100°C. The reducing sugar content was determined at OD 543 nm, using a D-xylose standard as a reference.

RESULTS

Construction of recombinant fungal strains and analysis of their growth and induction characteristics

Strain *A. awamori* AW15.7 (*exlA*$^-$::*uidA*$^+$, *pyrG*$^-$, Hm^R) was transformed with the expression vectors described in M&M using the recently developed integration system by Gouka et al. (1995). Transformants with a single copy of the vector at the *pyrG* locus were identified by Southern blot analysis and for each of the expression vectors two independent single copy transformants were used for further analysis (see Table I). In some cases also multi copy transformants were analyzed. As a control AW15.7-1, a *pyrG*$^+$ derivative of AW15.7, was used (Gouka et al. 1996a). For analysis of mRNA and protein levels, each transformant was cultivated in duplicate according to the induction procedure described in M&M. Since mRNA and protein levels can only accurately be compared if no differences occur in growth and induction, a number of control parameters were analyzed for each strain.

Growth of the cultures was checked by analysis of three parameters, namely culture morphology, dry weight and D-xylose consumption. All cultures showed a similar morphology of small pellets. Dry weight measurements were carried out 38 hours after induction.

Table I Expression strains

Strain	transf number	copy number	prepro-sequence	gene	encoded protein	origin	origin of plasmid	
AW15.7-1	–	–	–	–	–	–		
AW14A	2	1	*exlA*	*exlA*	1,4-ß-endoxylanase A	*A. awamori*	pAW14B	(Hessing et al. 1994)
AWGLA	3, 4	1	*glaA*	$glaA_{G2}$	glucoamylase G2	*A. niger*	pAN56-1[1)]	(Broekhuijsen et al. 1993)
AWLPL1	20	1	*lplA*	*lplA*	lipase	*Thermomyces lanuginosus*	pTL-1[2)]	(Boel et al. 1989)
AWLPL2	42	1	*exlA*	*lplA*	lipase	*Thermomyces lanuginosus*	pTL-1[2)]	(Boel et al. 1989)
AWHIL6	7, 12	1	*exlA*	hIL6	interleukin-6	Human	pAN56-4	(Broekhuijsen et al. 1993)
AWAGL1			*aglA*	*aglA*	α-galactosidase	*Cyamopsis tetragonoloba*	pUR2303	(Verbakel 1991)
AWAGL2	69	> 5	*exlA*	*aglA*	α-galactosidase	*Cyamopsis tetragonoloba*	pUR2303	(Verbakel 1991)
AWAGLS	9, 11	1	*exlA*	$aglA_{syn}$	α-galactosidase	*Cyamopsis tetragonoloba*	pUR2746	(Verbakel 1991)

1) pAN56-1 is identical to pAN56-3 (Broekhuijsen et al. 1993) with the exception that pAN56-1 contains a shorter promoter fragment

2) pTL-1 contains a synthetically constructed lplA gene with a slightly modified DNA sequence to introduce restriction sites

An average value of 1.3 g (SE 0.1) mycelium per 100 ml medium was obtained. As a measure for growth rate also D-xylose consumption in the culture medium was determined for all single copy transformants. This analysis showed that after 22 hours the D-xylose concentration decreased from 5% to 3%, whereas after 38 hours it decreased further to about 1%. These results are in agreement with the D-xylose consumption of a wild-type strain (Gouka et al. 1996a).

For efficient production of heterologous proteins also the pH of the medium is an important parameter, since an acidic pH might either irreversibly inactivate heterologous proteins and/or activate acidic proteases which have been shown to degrade heterologous proteins such as hIL6 (Broekhuijsen et al. 1993). Measurements of the pH, showed that the pH remained almost neutral (6.4 - 6.7) for all strains.

To analyze induction levels, in host strain AW15.7 used for transformation, the endogenous *exlA* encoding sequence had been replaced by a DNA fragment containing an expression cassette with the gene encoding mature *E. coli* ß-glucuronidase under control of the *exlA* expression signals (Gouka et al. 1996a). The presence of this reporter construct in all different recombinant strains enabled the use of GUS activity as a control for the induction of the *exlA* promoter. Both after 22 hours (average GUS activity 660 U/mg, SE 88) and after 38 hours (average GUS activity 884 U/mg, SE 107) of induction no significant differences in GUS activity were observed between the various single copy strains, indicating that the induction level was similar in all strains.

From the results of the growth and induction analysis it can be concluded that growth and induction of all single copy strains was similar, precluding any effects of them on the specific mRNA and protein levels of the cultures.

Analysis of protein production

For the different expression strains the extracellular and intracellular protein levels were quantified by Western blot analysis and enzyme activity assays.

Using Western blot analysis, EXL, GLA_{G2}, LPL (2x) and AGL_{syn} could be detected in medium samples taken 22 and 38 hours after induction. Samples of these strains are shown in Figure 1. Human IL6 could only be detected in very low amounts after 20 times concentration of the medium samples. In medium samples of the strains containing the wild-type guar *aglA* gene (AWAGL1 or AWAGL2; data not shown), no AGL could be detected. In Figure 1 it can be seen that the amount of protein present at 22 hour was always about half the amount present at 38 hours. To determine whether the proteins determined with Western blot analysis were also

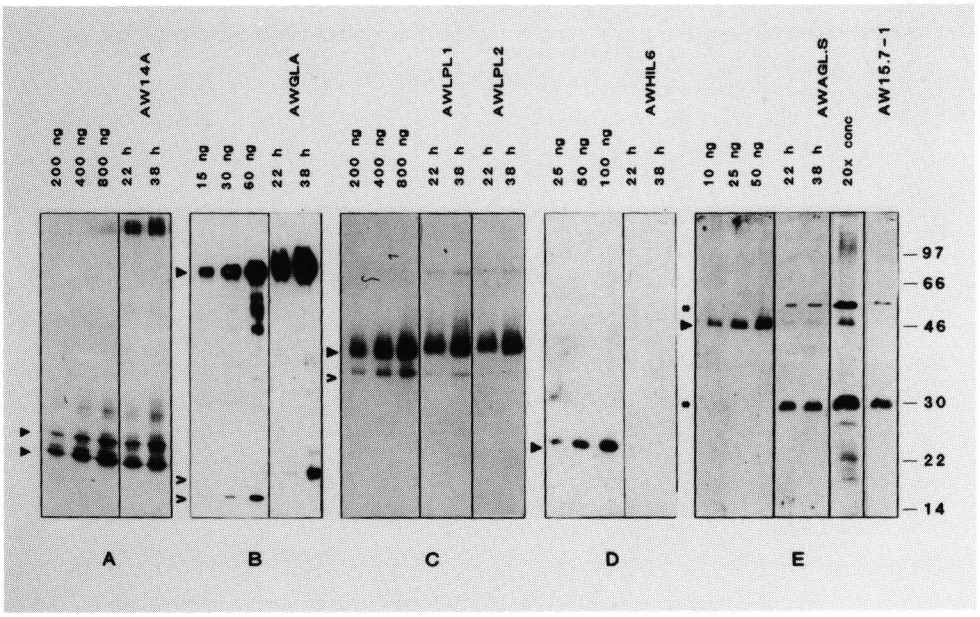


Fig.1 Western blot of medium samples taken after induction with D-xylose. Blots were incubated with antisera against 1,4-ß-endoxylanase (A), glucoamylase (B), lipase (C), human interleukin-6 (D), and α-galactosidase (E). Each panel contains a concentration range of purified protein, followed by a 15 μl medium sample of 22 and 38 hours cultures except the lanes in panel B which contain a 1 μl medium sample. The purified GLA protein is a deglycosylated sample. Panel E also contains 10 μl of a 20 times concentrated sample of 38 hours medium. The position of the specific proteins is indicated by a closed triangle (▸). Specific degradation products are indicated by a >, background signals as identified with control strain AW15.7-1 by a *. A marker for molecular weight (kD) is indicated on the right.

enzymatically active, enzyme activity tests were performed. Assuming that the specific activity of the proteins was similar to that in the original organism, it can be concluded that the amounts were indeed comparable with the amounts determined on the basis of Western blot analysis (Table II), indicating that the proteins were also enzymatically active. Specific EXL activity assays could not be carried out since EXL has only a minor contribution to the total xylanase activity (Gouka et al. 1996a).

In order to compare the production levels of the different proteins, all protein levels were calculated on molar basis and expressed relatively to the EXL protein level at 38 hours. As

shown in Table II production of heterologous fungal proteins is very efficient. The amount of GLA produced was even higher than the amount of EXL, whereas the levels observed for the LPL producing strains were comparable with the EXL protein level. Expression of the *lplA* gene with either the *lplA* preprosequence or with the *exlA* preprosequence resulted in similar amounts of lipase, suggesting that the different preprosequences did not influence protein secretion.

Table II Relative mRNA and extracellular protein levels at 38 hours

Strain	Gene	Prepro seq	RNA levels[1]	Protein levels[2] A	Protein levels[3] B	MW protein in kD	Relative protein amount
AW14A	*exlA*	*exlA*	++++	20-26		23	100
AWGLA	*glaA*	*glaA*	+++	> 100	100-150	80	200
AWLPL1	*lplA*	*lplA*	+++	26-35	30	38	90
AWLPL2	*lplA*	*exlA*	+++	26-35	30	38	90
AWHIL6	hIL6	*exlA*	+	< 0.1	< 0.1	23	< 1
AWAGL1	*aglA*	*aglA*	-	ND	ND	46	ND
AWAGL2	*aglA*	*exlA*	-	ND	ND	46	ND
AWAGLS	$aglA_{syn}$	*exlA*	+/-	0.2 - 0.4	0.4	46	1

ND Not Detectable.

[1] mRNA levels are relative to *exlA* mRNA level: - = no mRNA detectable, +/- = 1-5 %, + = 6-25 %, ++ = 26-50 %, +++ = 51-75 %, ++++ = 76-100 %

[2] Protein levels (mg l^{-1}) as determined by Western blot analysis

[3] Protein levels (mg l^{-1}) as determined by enzyme activity assays assuming an identical specific activity as the protein in the original organism. GLA_{G2} levels have been determined using pure GLA (G1-form) as standard.

When necessary, the protein amounts have been corrected for background values using strain AW15.7-1

For all cases tested, the protein levels obtained with the non-fungal genes were always clearly lower. For human interleukin 6 protein amounts of a few tens of μg l^{-1} could be detected in a biological activity assay and on a Western blot after 20 times concentration of the medium samples (data not shown). Guar α-galactosidase could not be detected in the culture medium, even after concentration of medium samples. The same was the case when the medium of a multi copy strain was analyzed. Surprisingly, α-galactosidase, encoded by a synthetic gene differing only in codon usage from the wild-type clone, was produced at levels of 0.2-0.4 mg l^{-1}, being approximately 1% of the EXL protein level.For all strains also the intracellular protein levels were determined. The proteins EXL and LPL could be detected in the soluble fraction of the extracts at a level of about 5% of the extracellular level and at a lower level in the insoluble fraction. Intracellular GLA_{G2} levels were higher than found for EXL and LPL but still markedly less than the GLA_{G2} levels observed in the medium, suggesting that the high production level of GLA_{G2} also results in higher intracellular GLA_{G2} levels. AGL_{syn}, AGL and hIL6, which could not or hardly be detected in the culture fluid, could also not be detected in the mycelial extracts, neither in the soluble nor in the insoluble fraction.

One of the possible explanations for the undetectable levels of AGL and the low levels of hIL6 in the medium could be proteolytic degradation. Although no proteolytic activity towards a mixture of marker proteins could be detected in the culture medium of our *A. awamori* strain (Gouka et al. 1996a), it could be possible that both AGL and hIL6 were degraded specifically. To check this, purified AGL (10 μg) and hIL6 (50 ng) were added to 22 and 38 hour medium of AW15.7-1 and incubated at 30°C. At time points up to 24 hours of incubation samples were taken, which were analyzed by activity assays (AGL) or Western blot analysis (hIL6). Both proteins appeared to be stable for at least 24 hours, indicating that degradation by extracellular proteases was not the reason for the undetectable respectively low protein levels.

Analysis of mRNA synthesis

To investigate whether the results obtained from the protein analysis could be explained by differences at the transcriptional level, the specific mRNA levels at 22 and 38 hours after induction were determined by Northern blot analysis. A DNA fragment comprising the 5' and 3' untranslated *exlA* sequences (see M&M) was used as a probe. As a control for the amount of RNA on the blot, hybridization was carried out with the *A. niger gpdA* gene (Verdoes et al. 1994) as a probe. Figure 2 shows the results of the 38 hours samples. Similar results were obtained with the 22 hours samples (results not shown). It is clear from this Figure that considerable differences occur at the steady-state mRNA level for the different genes.

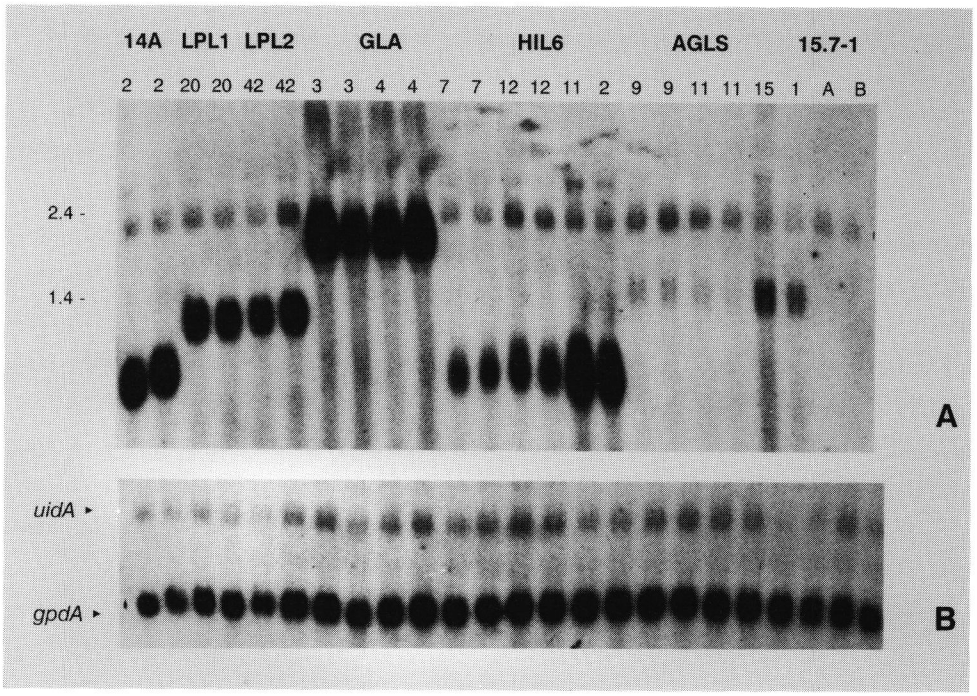


Fig.2 Northern blot analysis of total RNA isolated from all expression strains after 38 hours of induction. As probes a DNA fragment was used that contained both the 5' and 3' *exlA* non-coding regions (Panel A) and a mixture of a 1.4 kb *Hin*dIII DNA fragment from pAN5-2, containing the *A. niger gpdA* gene and a 1.9 kb *Nco*I fragment from pNOM102, containing the *uidA* gene (panel B). The numbers above the lanes correspond to the transformant numbers as shown in Table I. Identical numbers indicate RNA from duplicate cultures. A marker for molecular size is indicated on the left.

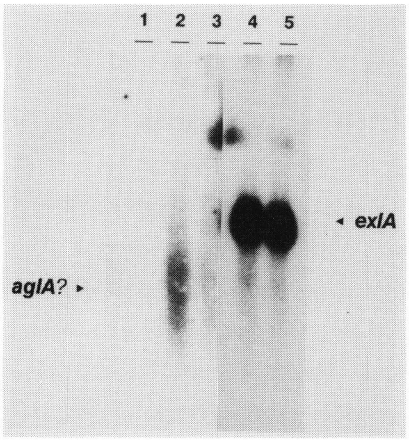


Fig.3 Northern blot analysis of total RNA isolated from AWAGL2 #69 after 0 (*lane 1*), 4 (*lane 2*) and 6 (*lane 3*) hours and AW14A after 4 (*lane 4*) and 6 (*lane 5*) hours of induction. As probe a DNA fragment was used containing the 0.7 kb *exlA* coding region (for AW14A) or the 1.2 kb *aglA* coding region (for AWAGL2).

The highest mRNA levels were obtained with the fungal genes encoding *exlA*, *glaA* and *lplA*. No differences in mRNA levels were observed with the *lplA* gene expressed with either the *lplA* preprosequence (AWLPL1) or the *exlA* preprosequence (AWLPL2), indicating that the difference in preprosequence does not influence transcription efficiency and/or mRNA stability.

The steady-state mRNA levels observed for the non-fungal genes were remarkably lower. The hIL6 steady-state mRNA levels were approximately 4-8 fold lower than those of the fungal genes. No full length guar *aglA* mRNA was detected, neither when the *aglA* nor the *exlA* preprosequence was used (data not shown). When the total *aglA* coding region was used as a probe a faint hybridization signal was visible which might represent a truncated mRNA (Fig. 3). This truncated mRNA could probably not hybridize efficiently with the 5' and 3' *exlA* probe. In a multi copy strain this signal was stronger. Interestingly, low but significant levels of full length mRNA were detected when the $aglA_{syn}$ gene was expressed. As expected, in the hIL6 and $aglA_{syn}$ multi copy strains the mRNA levels were higher than those in a single copy strain.

Considering the results obtained with the guar *aglA* expression vectors, it was important to exclude that unwanted sequence alterations in the *aglA* coding sequence were introduced during the construction of the vectors. Therefore, the *aglA* coding region from both *aglA* containing vectors (pAWAGL1 and pAWAGL2, corresponding to strains AWAGL1 and AWAGL2, Table I) was placed in a yeast expression vector and expressed in *Saccharomyces cerevisiae*. With both vectors transformants were obtained which produced high levels of α-galactosidase mRNA (full length) and enzymatically active α-galactosidase, indicating that no mutations had occurred in the sequence.

DISCUSSION

Heterologous protein production

The objective of this research was to obtain more insight in the parameters determining the differences in heterologous and homologous protein production in filamentous fungi. Since it is clear from literature that, in general, protein production levels can be roughly divided in two groups, high levels for fungal and low levels for non-fungal proteins (Van den Hondel et al. 1991; Verdoes et al. 1995), we chose to express three fungal genes, 1,4-ß-endoxylanase (*exlA*), glucoamylase (*glaA*) and lipase (*lplA*) and two non-fungal genes, human interleukin-6 (*hil6*) and guar α-galactosidase (*aglA*). The latter gene was available in two versions, a plant cDNA clone and a synthetic gene with optimized yeast codon bias. With the fungal genes

high extracellular protein levels were obtained, which showed that the fungal genes were expressed and their protein products secreted almost equally efficient. These results are in agreement with the results obtained with other homologous and heterologous fungal proteins showing that production of fungal proteins is usually efficient (Verdoes et al. 1995).

High production levels of non-fungal proteins are generally very difficult to obtain without optimization procedures (e.g. gene-fusions or random mutagenesis and selection of higher producing strains) and the initial levels that are produced do often not exceed a few milligrams per litre (reviewed by a.o. Van den Hondel et al. 1991). The results which were obtained in this study with human interleukin-6 and plant α-galactosidase are in agreement with these observations. Production of human interleukin-6 in other *Aspergillus* species was also very inefficient. In *A. nidulans* initial levels of a few tens of $\mu g\ l^{-1}$ were found (Carrez et al. 1990), whereas in *A. niger* hIL6 could not even be detected (Broekhuijsen et al. 1993). As possible explanations for the initial low hIL6 levels transport efficiency or extracellular proteolytic degradation has been suggested. Our results show that in our strain other factors are involved (see below).

The total absence of guar α-galactosidase production was even more remarkable since production of the protein, although with variable efficiencies, was possible in a range of organisms as *Bacillus subtilis* (Overbeeke et al. 1990), *Hansenula polymorpha* (Fellinger et al. 1991), *S. cerevisiae* (Verbakel 1991) and *Kluyveromyces lactis* (Bergkamp et al. 1992). Data obtained from Northern analysis showed that the absence of AGL protein was due to the absence of full length *aglA* mRNA (see below).

Heterologous steady-state mRNA levels

The results of the mRNA analyses showed that mRNA stability is an important reason for the low or undetectable protein levels (especially AGL) of heterologous genes, since considerable differences in the steady-state mRNA levels were observed. The mRNA levels in the strains expressing the three fungal genes were clearly higher than those obtained with the strains expressing the non-fungal genes. In particular in transformants containing the plant α-galactosidase expression cassette no detectable levels of full length α-galactosidase mRNA could be found. A similar result had also been observed previously when the gene was expressed in *A. nidulans* and *A. niger* under control of the *A. nidulans gpdA* promoter (PJ Punt, unpublished results). Surprisingly, changing the codon usage of the gene resulted in low, but detectable levels of full length *aglA* mRNA.

In the synthetic *aglA* gene 173 minor codons out of 364 codons, according to the yeast

codon bias, had been modified in major codons in such a way that the occurrence of strong secondary structure was avoided (in total 192 codons have been changed). These modifications resulted in an overall change of 22% in nucleotide sequence without affecting the amino acid sequence (Verbakel 1991).

Preliminary results obtained with nuclear run-on transcription assays performed with the strain AWAGL2 indicated that transcription initiation is occurring. These results support our data that the mRNA that was observed by Northern blot analysis might indeed be a truncated version of the *aglA* mRNA. Experiments are now in progress to determine why no detectable levels of full length mRNA were obtained.

Since transcription initiation of the guar *aglA* gene occurs, the absence of full length mRNA must have been caused by other factors than transcription initiation per se. These include problems in transcription elongation, post transcriptional maturation events like the addition of a 3' poly(A) tail or a 5' cap, transport to the cytoplasm and/or cytoplasmic degradation. Considering the results obtained it is still unknown at which of these levels the problems occur. The fact that changing of the codon sequence of the gene results in full length mRNA might suggest that the coding region contains sequences affecting mRNA synthesis and/or mRNA instability. In higher eukaryotes several factors have been described which determine the stability of an mRNA molecule (reviewed by a.o. Atwater et al. 1990; Hentze 1991). Experiments are currently in progress to determine which of these factors are the case for *aglA*.

Proteolytic degradation

A low level of protein in the medium could also be due to extracellular proteolytic degradation as was suggested in some studies (Archer et al. 1992; Broekhuijsen et al. 1993). Also in our study no hIL6 protein was observed whereas relatively high levels of hIL6 mRNA were found. Since extracellular proteolytic degradation of hIL6 did not occur, it is very likely that hIL6 was degraded before entering the culture medium. One reason for this degradation could be that the protein is improperly folded. A protein which is shown to be involved in folding is the immunoglobulin heavy chain binding protein BiP (Haas and Wabl 1983). BiP has been demonstrated also to interact with misfolded proteins and it is believed that it plays a role in degradation of these misfolded proteins (Knittler et al. 1995 and references therein). Recently, in our laboratory the *A. niger* BiP-encoding gene (*bipA*) was isolated (Van Gemeren and Punt, in preparation). Analysis of *A. niger* strains which overproduce certain proteins showed that BiP mRNA is induced (Punt et al, in preparation). This induction is probably

triggered by a so-called unfolded protein response (Shamu et al. 1994), as was also observed by protein overproduction in other systems. To investigate whether in our case the expression of heterologous genes also resulted in increased *bipA* mRNA levels, Northern blot analysis of all expression strains was carried out using the *A. niger bipA* gene as a probe. This analysis showed that in all strains BiP mRNA levels were similar to the level observed in the control strain AW15.7-1 which contained no expression cassette (data not shown). These observations indicated that expression of the homologous and heterologous genes did not lead to *bipA* induction. Also hIL6 and *aglA* multi copy strains did not show increased *bipA* mRNA levels. From these data we conclude that protein degradation through an unfolded protein response is not occurring in our expression strains.

Acknowledgements.

The research described in this paper was partially paid by Unilever Research Laboratorium, Gist-brocades, Solvay Duphar, DSM and the Ministry of Economic Affairs. We wish to thank Dr. R. Contreras for carrying out the hIL6 biological activity assays, Prof. Dr. Ir. C.T. Verrips for critical reading of the manuscript and Dr. W. Musters, Dr G.C.M. Selten, Dr. C. Groeneveld and Dr. D. Carrez for fruitful discussions.

Submitted

CHAPTER 5

Glucoamylase gene fusions alleviate limitations for protein production in *Aspergillus awamori* at the transcriptional and (post-) translational level

Robin J. Gouka, Peter J. Punt, and Cees A.M.J.J. van den Hondel

TNO Nutrition and Food Research Institute, Department of Molecular Genetics and Gene Technology, P.O. Box 5815, NL-2280 HV Rijswijk, the Netherlands

SUMMARY

In this study the effects of a glucoamylase gene fusion on the mRNA levels and protein levels of human interleukin-6 (hIL6) and guar α-galactosidase (*aglA*) were analyzed. Expression cassettes, containing either hIL6 or *aglA* fused to the *Aspergillus niger* glucoamylase (*glaA*) gene were integrated in a single copy at the *pyrG* locus of *A. awamori*. A *glaA* fusion to the 5' end of hIL6 resulted in a large increase in hIL6 yield, whereas with a *glaA* fusion to the 3' end of hIL6 almost no protein was produced, although both fusion mRNA levels were very similar. Fusions of *glaA* to the 5' end of *aglA* resulted in truncated mRNAs lacking more than 80% of the *aglA* sequence. With a synthetic *aglA* gene with yeast-optimized codon usage, low levels of full length mRNA were obtained. A *glaA* fusion with this synthetic *aglA* gene resulted in a 25-fold increase in mRNA level and an almost similar increase in AGL protein level. Analysis of the truncated wild type *aglA* transcripts by nuclear run-on transcription assays and reverse-transcriptase PCR indicated that truncation most likely occurs by incorrect processing of the *aglA* pre-mRNA in the nucleus.

INTRODUCTION

Filamentous fungi, especially members of the genus *Aspergillus*, have the ability to secrete large amounts of a wide range of different enzymes into their environment. This characteristic has favoured the use of these organisms for large scale production of commercially important enzymes (reviewed by a.o. Fowler and Berka, 1991; Van den Hondel et al. 1991; Verdoes et al. 1995). Although production of fungal proteins is usually quite efficient, non-fungal proteins are poorly produced and reach levels that do often not exceed a few milligrams per litre (Fowler and Berka, 1991; Van den Hondel et al, 1991; Van Gorcom et al. 1994, Gouka et al. 1996b). Therefore, several strategies have been developed to improve these yields. A successful strategy is based on the use of a well-secreted "carrier" protein, usually *A. niger* glucoamylase (Ward et al. 1990; Contreras et al, 1991; Roberts et al. 1992; Broekhuijsen et al, 1993; Ward et al. 1995) or *Trichoderma reesei* cellobiohydrolase I (Nyyssönen et al. 1995), which is fused to the N-terminal end of a heterologous protein. With this strategy a fusion protein is produced of which the N-terminal GLA is believed to improve the secretion efficiency of the heterologous protein by facilitating translocation and subsequent folding in the endoplasmic reticulum. Further along the pathway, in most cases the fusion protein is cleaved, resulting in the secretion of the separate proteins. Cleavage has been shown to occur either by autocatalytic processing of the heterologous protein (Ward et al. 1990), an unknown fungal protease (Roberts et al. 1992; Nyyssönen et al. 1993; Nyyssönen et al. 1995) or by a KEX2-like protease, for which a recognition site had been introduced in the fusion protein (Contreras et al. 1991; Broekhuijsen et al. 1993; Ward et al. 1995).

The aim of our study is to obtain more insight in the parameters determining the efficiency of the production and secretion of heterologous proteins in filamentous fungi. Recently, we showed that production of two different heterologous proteins, guar α-galactosidase (*aglA*) and human interleukin-6 (*hil6*), was limited at the transcriptional and post-translational level, respectively (Gouka et al, 1996b). In this study the effect of a glucoamylase gene fusion with hIL6 and *aglA* on mRNA level and protein level was investigated. For our study, the filamentous fungus *Aspergillus awamori* was used, which has been shown to be an efficient producer of proteins (Finkelstein et al. 1989; Berka et al. 1991, Ward et al. 1995). Recently, a similar study was carried out in *Trichoderma reesei* for another fusion protein, cellobiohydrolase I (Nyyssönen et al. 1995). In contrast to these authors we have chosen to analyze defined single-copy strains similar as described previously (Gouka et al. 1996b), which would allow us to compare the effect of the gene fusions on the production levels of

the different target proteins more precisely.

MATERIALS AND METHODS

Strains, media and growth conditions. *A. awamori* AW15.7 (*exlA*$^-$::*uidA*$^+$, *pyrG*$^-$, Hm^R; Gouka et al. 1996a), used as a recipient strain for transformation, is a derivative of *A. awamori* ATCC 11358 (CBS 115.52). In AW15.7 the genomic 1,4-ß-endoxylanase (*exlA*) encoding sequence had been replaced by a DNA fragment containing an expression cassette with the gene encoding mature *E. coli* ß-glucuronidase under control of the *exlA* expression signals (Gouka et al. 1996a). *Escherichia coli* strain JM109 (Sambrook et al. 1989) was used for propagation of plasmids. Induction of the *A. awamori exlA* promoter was carried out as described by Gouka et al (1996a), using a transfer of mycelium from non-inducing medium containing 1% sucrose as C-source to induction medium with 5% D-xylose as inducer. D-xylose concentrations were determined as described before (Gouka et al. 1996a).

Aspergillus expression strains. For construction of the five different expression strains five expression vectors were made (Table I) consisting of (i) an expression cassette, which include transcriptional regulatory sequences and DNA sequences encoding the truncated *A. niger* GLA G2 form (514 AA), thus lacking the starch-binding domain (Broekhuijsen et al. 1993) fused to the coding sequence of hIL6 or *aglA*, (ii) a selection marker for transformation and (iii) pUC19 sequences for cloning and amplification in *E. coli*.

In four of the five expression cassettes, the sequences encoding GLA_{G2} and hIL6, AGL or AGLS were separated by a sequence encoding the amino acids NVISKR (identical to the *A. niger glaA* prosequence; further referred to as KEX2 site). The C-terminal dibasic proteolytic cleavage site KR results in cleavage of the fusion protein into both separate proteins by a KEX2-like protease (Broekhuijsen et al. 1993). All expression cassettes contained the *exlA* promoter and transcription terminator region. The DNA sequences of hIL6 or α-galactosidase encoded the mature proteins, thus without the original pre- and prosequence.

Construction of the vectors was carried out as follows (Figure 1): vectors pAWGLA1 or pAWGLA2 were used as basis for all five vector constructions. These vectors contain respectively the *A. awamori exlA* (Hessing et al. 1994) promoter fused to the truncated *A. niger glaA* gene (encoding GLA_{G2}), followed by the *A. awamori exlA* transcription terminator in pUC19. The only difference between both vectors is located between the *Nar*I and the *Afl*II site at the 3' end of the *glaA* gene (see Fig.1).

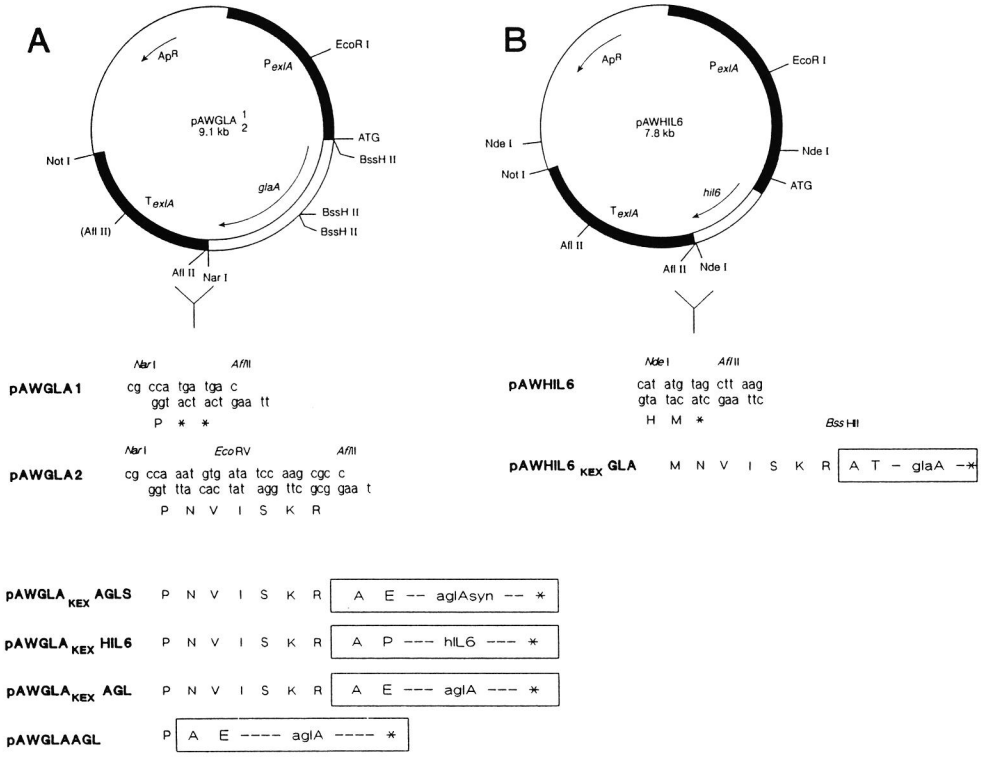


Fig.1 Maps of plasmids pAWGLA1 and 2 (Panel A) and pAWHIL6 (Panel B) that were used as starting vectors for the construction of the gene-fusions. Thin lines represent *E. coli* sequences, closed boxes represent *A. awamori exlA* 5' and 3' regulatory sequences and open boxes either *A. niger glaA* (panel A) or human IL6 (panel B) sequences. Only relevant restriction sites are shown. The *Afl*II site which is between brackets (panel A) indicates that it has been removed in pAWGLA1 and -2. Below the plasmid maps, the introduced linker sequences and the corresponding amino acid sequence are shown. The amino acid sequence of the fusion region of the constructed fusion-plasmids is indicated in the lower part. All fusions were obtained with the restriction sites shown in the figure (see M&M).

For *glaA* fusions with the human interleukin-6 gene, first in pAN56-4 (Broekhuijsen et al. 1993), an *Afl*II site was introduced downstream of the stopcodon of *hil6* by PCR, resulting in pAN56-4AflII. Simultaneously, the last two codons before the stopcodon of *hil6* (caa atg) were

changed into the sequence cat atg to introduce a *Nde*I site.

For construction of vector pAWGLA$_{KEX}$HIL6, in which *glaA* is fused to the 5' end of hIL6, the *hil6* gene was isolated from vector pAN56-4AflII as 0.6 kb *Eco*RV/*Afl*II fragment and introduced into pAWGLA2 which was also cut with *Eco*RV/*Afl*II, resulting in a correct in-frame fusion of *glaA* with hIL6, separated by the KEX2-site. For construction of vector pAWHIL6$_{KEX}$GLA, in which *glaA* is fused to the 3' end of hIL6, a partial 1.8 kb *Eco*RI/*Nde*I fragment was isolated from pAWHIL6 (Gouka et al. 1996b). This fragment contains part of the *exlA* promoter and preprosequence, followed by the *hil6* gene up to the *Nde*I site. From pAWGLA1 a partial 7.9 kb *Bss*HII/*Eco*RI fragment was isolated containing the *glaA* gene followed by the *exlA* transcription terminator, pUC19 and the remaining part of the *exlA* promoter. To obtain a correct fusion between the *Nde*I site at the 3' end of *hil6* and the *Bss*HII site at the 5' end of *glaA*, oligonucleotides were used (5' tatgaatgtgatttccaag 3' and 5'cgcgcttggaaatcacattca 3').

For the construction of the *Cyamopsis tetragonoloba* α-galactosidase fusion cassettes two versions of the *aglA* gene were used. The wild-type gene (Overbeeke et al. 1989) and a synthetic gene with an optimized yeast codon bias (*aglA*$_{syn}$; Verbakel 1991) differing 22% in nucleotide sequence from the cDNA clone. For construction of vector pAWGLA$_{KEX}$AGLS, in which *glaA* is fused to the 5' end of the *aglA*$_{syn}$ gene, first, the 5' end of the *aglA*$_{syn}$ gene was adapted in vector pAWAGLS (Gouka et al. 1996b). In this vector the *aglA*$_{syn}$ gene is under control of the *exlA* promoter and is preceded by the *exlA* preprosequence. At the 5' end of *aglA*$_{syn}$ an *Eco*RV site was introduced using an oligonucleotide linker (5' gatccgatatccaag 3' and 5' cttggatatcg 3'), which encodes part of the KEX2-site (the amino acids ISK). The missing arginine (R) is obtained from the *exlA* prosequence. The *aglA*$_{syn}$ gene was subsequently isolated from pAWAGLS as a partial 1.1 kb *Eco*RV/*Afl*II fragment and ligated with the 9.1 kb *Eco*RV/*Afl*II fragment from pAWGLA2.

The construction of pAWGLA$_{KEX}$AGL and pAWGLAAGL, in which the *glaA* gene is fused to the wild-type *aglA* gene was identical for both vectors except for the linkers used to obtain the *glaA-aglA* fusion. These linkers were inserted in vector pUR2303AflII, a derivative of pUR2303 (Verbakel 1991) in which an *Afl*II site was introduced almost immediately downstream of the stopcodon of the *aglA* gene. pUR2303AflII was digested completely with *Nar*I that cuts in the pBR322 sequence (four times) and *Nco*I that cuts after nucleotide 29 from the start of the mature encoding sequence. For a GLA$_{KEX}$AGL fusion the following oligonucleotide linker was inserted: 5' cg cca aat gtg att tcc aag cgc gct gaa aat gga cta ggc cag aca cct cc 3' and 5' cat ggg agg tgt ctg gcc tag tcc att ttc agc gcg ctt gga aat cac att tgg 3', encoding the sequence PNVISKRAENGLGQTPP). For a GLAAGL fusion the following

oligonucleotide linker was used: 5' cg cca gct gaa aat gga cta ggc cag aca cct cc 3' and 5' cat ggg agg tgt ctg gcc tag tcc att ttc agc tgg 3', encoding the sequence PAENGLGQTPP). From both vectors the *aglA* gene was isolated as 1.1 kb *Nar*I/*Afl*II fragment and ligated with a 9.1 kb *Nar*I/*Afl*II fragment from pAWGLA1, resulting in both cases in a correct in-frame fusion between *glaA* and *aglA*.

All relevant DNA sequences were verified by sequence analysis. As selection marker, a mutant *A. awamori pyrG* gene (Gouka et al. 1995) present on a 2.4 kb *Not*I fragment, was inserted in all the expression vectors. Transformants containing vectors integrated in a single-copy at the *pyrG* locus in *A. awamori* AW15.7 (*exlA*$^-$::*uidA*$^+$, *pyrG*$^-$, Hm^R) were identified by colony blot hybridization and subsequent Southern blot analysis.

To compare the effect of a glucoamylase gene fusion of the heterologous genes with the corresponding non-fused genes a number of control strains were included in the analysis (Table I). Construction of these strains, AW15.7-1 (Gouka et al. 1996a) and AWGLA, AWHIL6, AWAGL2, and AWAGLS (Gouka et al. 1996b), has been reported previously.

Recombinant DNA techniques. Standard recombinant DNA techniques were used for cloning procedures (Sambrook et al. 1989). *Aspergillus* chromosomal DNA and RNA was isolated as described by Kolar et al. (1988). Colony blot hybridization of *A. awamori* was carried out according to Kinsey (1989). Transformation of *Aspergillus* and selection of transformants was carried out as described by before (Gouka et al. 1995). For Northern blot analysis, as a probe a 400 bp DNA fragment was used that contained the 5' and 3' *exlA* untranslated sequences (Gouka et al. 1996b), which allowed a direct comparison of the amount of specific mRNA in the different expression strains. With this probe in each transformant the *uidA* mRNA provided with *exlA* expression signals could also be detected and used as a reference for the induction level.

For purification of polyadenylated mRNA Dynabeads Oligo $(dT)_{25}$ (Dynal) were used.

Quantification of protein production levels. Quantification was carried out by Western blotting (Sambrook et al. 1989). For determination of protein concentrations purified protein was used as a standard. Purified α-galactosidase and corresponding polyclonal antibodies were obtained from Unilever Research Laboratorium. Purified glucoamylase was obtained from Boehringer-Mannheim (lot 12005320-30). Monoclonal antibodies raised in mice against glucoamylase were produced in our laboratory (A. Mateo-Rosell, unpublished). Purified human interleukin-6 (200-06) and polyclonal antibodies were supplied by Sanvertech. For detection,

the ECL Western blot detection kit (Amersham) was used, based on peroxidase labelled antibodies. Concentration of proteins was carried out as described by Broekhuijsen et al (1993). ß-Glucuronidase assays were performed according to Roberts et al. (1989). α-Galactosidase activity assays were carried out essentially as described by Overbeeke et al. (1990).

For mycelial dry weight measurements 25 ml of homogenous culture samples was filtered through Miracloth. The mycelium was squeezed and dried overnight in a vacuum oven at 80°C.

Isolation of nuclei and nuclear run-on transcription assays. Isolation of nuclei and nuclear run-on transcription assays were carried out essentially as described by Schuren et al. (1993).

Reverse transcriptase PCR and cloning of a PCR fragment. For synthesis of cDNA from a mRNA template M-MLV reverse transcriptase (Gibco BRL) was used. cDNA synthesis, using poly(A)$^+$ RNA from the strains AWGLA$_{KEX}$AGL and AWAGL2, was carried out according to the protocol of the supplier (Gibco BRL; 60 minutes, 37°C), using an oligo dT primer (5' ggaattcgcggccgc(t)$_{20}$(g/a/c) 3'). Prior to PCR the samples were extracted once with phenol/chloroform to remove the M-MLV and the cDNA was ethanol precipitated. PCR on the cDNA was performed with, as an upstream primer, an oligonucleotide specific for *aglA* comprising the first 18 nucleotides of the coding sequence of the mature protein and the oligo dT primers mentioned above. A specific DNA fragment of approximately 250-260 bp was obtained which contained the *aglA* gene as was shown by Southern blot analysis (data not shown). The fragment was isolated using two successive PCR reactions, since the first RT-PCR reaction resulted in the amplification of a mixture of DNA products. The conditions of the first reaction were 1 min. 94°C, 1 min. 37°C, 1 min. 72°C, 40 cycles. In the second reaction, which was performed with a fraction of the first PCR mixture, the conditions were changed to 30 cycles and an anneal temperature of 40°C. The 250-260 bp fragment obtained with AWGLA$_{KEX}$AGL cDNA was digested with *Nco*I and *Eco*RI and the resulting fragment was subcloned in pUC19 which was digested with *Afl*III and *Eco*RI.

RESULTS

Construction of recombinant fungal strains and analysis of their growth and induction characteristics

Recently, we showed that production of guar *aglA* and human interleukin-6 was limited at the transcriptional and post-translational level, respectively (Gouka et al, 1996b). In this study *glaA* gene fusions with hIL6 and *aglA* were constructed to obtain more insight in which factors influence the expression of these genes and the production and secretion of the gene-products. Five vectors, containing expression cassettes in which *glaA* was fused to either *aglA*, $aglA_{syn}$ or hIL6 were constructed (M&M; Table I). Strain *A. awamori* AW15.7 (*exlA*$^-$::*uidA*$^+$, *pyrG*$^-$, Hm^R) was transformed with the five vectors and transformants with a single copy of the vector at the *pyrG* locus were identified by colony blot hybridization and subsequent Southern blot analysis. For each of the vectors two single copy transformants were used for further analysis. To compare the mRNA levels and the protein levels with those of strains containing the corresponding non-fused genes, also these single copy strains (Gouka et al. 1996b) were included in the analysis (Table I). As a control AW15.7-1, a *pyrG*$^+$ derivative of AW15.7, was used (Gouka et al. 1996a).

For analysis of mRNA levels and protein levels, each transformant was cultivated in duplicate according to the induction procedure described in M&M. Since mRNA and protein levels can only accurately be compared if no differences occur in growth and the level of induction, a number of control parameters were analyzed for each strain as described previously (Gouka et al. 1996b).

Growth of the cultures was checked by the parameters culture morphology, dry weight, D-xylose consumption and pH. With none of the parameters significant differences were observed, precluding any effects of them on the specific mRNA and protein level. Measurements of the pH showed that the pH remained almost neutral (6.4 - 6.7) for all strains. In previous experiments (Gouka et al. 1996b) it was shown that hIL6 and α-galactosidase were stable in this culture medium, excluding the effect of extracellular proteolytic degradation, a phenomenon that often influences the production of heterologous proteins in filamentous fungi (Roberts et al. 1992; Broekhuijsen et al. 1993).

Induction levels were verified by analysis of the expression of the ß-glucuronidase gene, present on a P_{exlA}-*uidA*-T_{exlA} reporter construct in all different recombinant strains. Also in this case no significant differences were observed.

Table I Strains and vectors used in this study

Strain/Vector[a)]	prepro-sequence	gene(-fusion)	KEX2-site	reference
AWGLA$_{KEX}$HIL6	*glaA*	*glaA*$_{G2}$-*hil6*	+	this paper
AWHIL6$_{KEX}$GLA	*exlA*	*hil6*-*glaA*$_{G2}$	+	this paper
AWGLA	*glaA*	*glaA*$_{G2}$	_	Gouka et al. 1996b
AWHIL6	*exlA*	*hil6*	_	Gouka et al. 1996b
AWGLA$_{KEX}$AGLS	*glaA*	*glaA*$_{G2}$-*aglA*$_{syn}$	+	this paper
AWGLA$_{KEX}$AGL	*glaA*	*glaA*$_{G2}$-*aglA*	+	this paper
AWGLAAGL	*glaA*	*glaA*$_{G2}$-*aglA*	_	this paper
AWAGLS	*exlA*	*aglA*$_{syn}$	_	Gouka et al. 1996b
AWAGL2	*exlA*	*aglA*	_	Gouka et al. 1996b
AW15.7-1	_	_	_	Gouka et al. 1996a

[a)] The vectors correspond with the names of the strains preceded with the letter p. For example, strain AWGLA$_{KEX}$AGL corresponds with vector pAWGLA$_{KEX}$AGL.

Previously it was shown that expression of genes driven by the *exlA* promoter is dependent on the presence of D-xylose in the medium. After approximately 48 hours of cultivation D-xylose was depleted in the medium resulting in the degradation of mRNA and cessation of protein synthesis (Gouka et al. 1996a). Therefore, medium and mycelium samples were taken 22 and 38 hours after induction. Since at 38 hours the highest protein levels were obtained, about twice the amount at 22 hours, only the results of the 38-hour samples are reported.

Analysis of protein production in hIL6 strains

The amounts of extracellular and intracellular hIL6 of the relevant recombinant strains were quantified by Western blot analysis using purified hIL6 as a standard. Only in medium samples of AWGLA$_{KEX}$HIL6 transformants, containing an expression cassette encoding GLA$_{G2}$ fused to the N-terminal end of hIL6, hIL6 could be detected, at levels of approximately 5-10 mg l^{-1} (Figure 2B). In medium samples of AWHIL6$_{KEX}$GLA transformants, containing an

expression cassette encoding GLA_{G2} fused to the C-terminal end of hIL6, and AWHIL6 transformants, no hIL6 could be detected by Western analysis of non-concentrated medium samples. After concentration of the medium approximately 0.05 mg l^{-1} of hIL6 could be determined for both strains. As expected, in medium of the control strains AWGLA and AW15.7-1, hIL6 could not be detected.

The GLA_{G2} levels in these strains were also determined by Western blot analysis (Fig 2A). All strains produced approximately 10 mg l^{-1} (after 38 hours) of endogenous GLA. Since endogenous GLA is slightly larger in size than GLA_{G2}, it could clearly be distinguished from GLA_{G2} on a protein gel. The amount of GLA_{G2} produced by $AWGLA_{KEX}HIL6$ transformants was slightly lower than the amount produced by AWGLA transformants, being 125 - 175 mg l^{-1} versus 150 - 200 mg l^{-1}. In medium of $AWHIL6_{KEX}GLA$ strains only 5-10 mg l^{-1} of GLA_{G2} was present, which could only be detected when medium samples of 10 μl instead of 1 μl were applied to the gel (not visible in Fig 2A). All protein levels are summarized in Table II.

To compare the amounts of protein independent from the molecular weight of the various proteins, the levels were calculated on molar basis and given relatively to the amount of

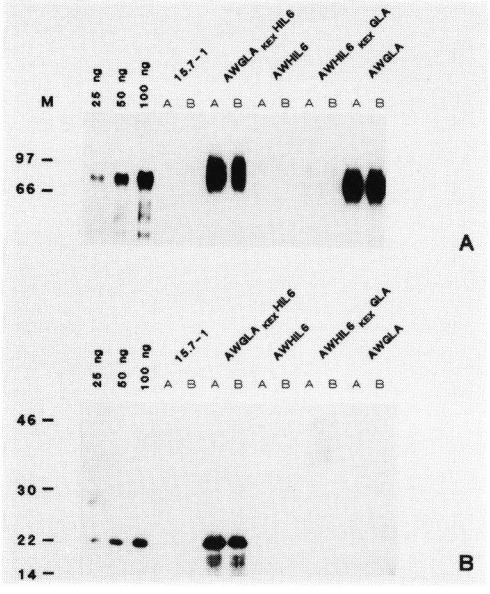


Fig.2 Western blot analysis of medium samples taken after induction with D-xylose. Blots were incubated with antisera against glucoamylase (Panel A) and human interleukin-6 (Panel B). Each panel contains resp. a concentration range of purified protein, followed by a 1 μl medium sample (Panel A) or a 15 μl medium sample (Panel B) of 38-hours cultures. The purified GLA protein is a deglycosylated GLA_{G1} sample. GLA_{G2} levels were quantified more accurately using serial dilutions (not shown). The letters A and B above the lanes indicate samples of duplicate cultures. A marker for molecular weight (kD) is indicated on the left.

GLA_{G2} produced by the AWGLA strain (Table II). From this Table it is clear that the GLA_{G2} production by $AWGLA_{KEX}$HIL6 is nearly as efficient (60-90%) as by AWGLA. However, on molar basis the secreted level of hIL6 by $AWGLA_{KEX}$HIL6 was only 10-20% of the GLA_{G2} level. Since both GLA_{G2} and hIL6 were initially produced as a fusion-protein, this result would indicate that 60-90% of the hIL6 is lost.

The hIL6 level of $AWHIL6_{KEX}$GLA was comparable with the hIL6 level observed for AWHIL6, being only 0.1%, thus a 100 to 200-fold lower as compared to the hIL6 yields observed for $AWGLA_{KEX}$HIL6. $AWHIL6_{KEX}$GLA also produced less GLA_{G2} (3-6%), which was also in this case still higher than the hIL6 level, indicating that more than 95% of hIL6 is lost.

For all strains also the presence of intracellular GLA_{G2} and hIL6 was analyzed, mainly to see whether accumulation of hIL6 (and/or GLA_{G2}) had occurred, which could explain the low levels of hIL6 in the medium. However, no significant accumulation of hIL6 or GLA_{G2} was observed (data not shown).

Comparison of steady-state mRNA levels in hIL6 strains

To investigate whether the differences in protein levels could be explained by variations in steady-state mRNA levels, Northern blot analysis was carried out (Figure 3). As a probe a DNA fragment was used containing the 5' and 3' *exlA* untranslated region (see M&M). In general, the mRNA level of a GLA_{KEX}HIL6 fusion was slightly higher than that of a $HIL6_{KEX}$GLA fusion and non-fused hIL6, but lower than non-fused GLA_{G2} mRNA (see Table II).

Analysis of protein production in α-galactosidase strains

Previous experiments had shown that expression of the wild-type guar *aglA* gene resulted in undetectable mRNA levels, possibly due to truncation of the transcripts (Gouka et al. 1996b). Full length mRNA, although at a low level, could only be observed when the wild-type guar *aglA* gene was replaced by a synthetic *aglA* gene with optimized yeast codon usage (Gouka et al. 1996b). As a result, AGL protein was only detected with strains containing the $aglA_{syn}$ gene. To further investigate these results and to see whether the limitations could be resolved, strains were constructed containing an expression cassette in which the *A. niger glaA* gene was fused to *aglA* or $aglA_{syn}$.

The amounts of extracellular and intracellular α-galactosidase of the relevant strains (Table I) were quantified by Western blot analysis using purified AGL protein as a standard. In medium samples of $AWGLA_{KEX}$AGLS approximately 10-12 mg l^{-1} of AGL protein could be detected (Figure 4B), which was 25-fold more than the amount of AGL produced by the

Table II mRNA levels and protein levels

Strains	Relative mRNA levels [a] (%)	Protein levels ($mg\ l^{-1}$)			Relative protein levels (%)	
		GLA_{G2} [b]	HIL6	AGL [c]	GLA_{G2} [d]	Heterologous protein [d]
AWGLA	100	150 - 200	-	-	(75 -) 100	-
AWHIL6	25 - 50	-	0.05	-	-	0.1
$AWGLA_{KEX}HIL6$	50 - 75	125 - 175	5 - 10	-	60 - 90	10 - 20
$AWHIL6_{KEX}GLA$	25 - 50	5 - 10	0.05	-	3 - 6	0.1
AWAGL2	< 1 [e]	-	-	ND	-	*
AWAGLS	1	-	-	0.4	-	0.4
$AWGLA_{KEX}AGLS$	20 - 25	30 - 40	-	10 - 12	15 - 20	10 - 12
$AWGLA_{KEX}AGL$	< 1 [e]	ND	-	ND	*	*
AWGLAAGL	< 1 [e]	ND	-	ND	*	*

ND Not detectable, * No relative levels known, since protein was not detected, - Not applicable

[a] mRNA levels relative to the steady-state mRNA level of GLA_{G2}

[b] GLA_{G2} levels have been determined using pure, deglycosylated GLA (G1-form) as standard

[c] The protein amounts of AGL have been corrected for background values using strain AW15.7-1 as a control

[d] Protein levels (GLA_{G2}, HIL6 and AGL) relative to GLA_{G2} level of AWGLA strain (based on MW GLA_{G2} 80 kD, MW AGL 45 kD, MW HIL6 20 kD). The GLA_{G2} level of 200 mg l^{-1} was taken as 100%.

[e] No full length mRNA detected (data based on comparison with $aglA_{syn}$ mRNA using an *aglA* probe)

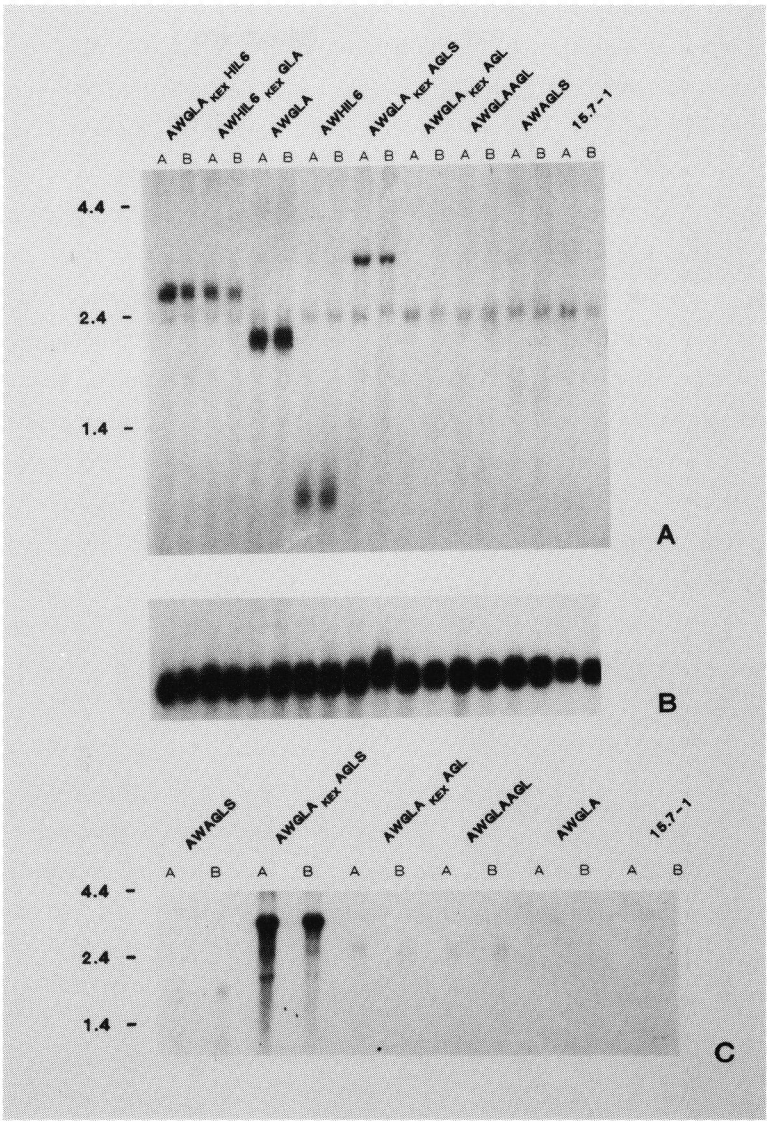


Fig.3 Northern blot analysis of total RNA isolated from all expression strains after 38 hours of induction. As probes DNA fragments were used that contained both the 5' and 3' *exlA* non-coding regions (Panel A), a 1.4 kb DNA fragment from pAN5-2, containing the *A. niger gpdA* gene (panel B) or a 1.1 kb DNA fragment from pAWAGL2 containing the *aglA* gene (panel C). The letters A and B above the lanes indicate samples of duplicate cultures. A marker for molecular size is indicated on the left.

corresponding non-fused strain AWAGLS (0.4 mg l^{-1}). In medium of this latter strain AGL was only visible after concentration of the sample (not shown in Fig 4B). AGL activity assays showed that all AGL, produced by AWGLA$_{KEX}$AGLS or AWAGLS, was present as active protein (data not shown). In medium samples of transformants containing expression cassettes with *glaA* fused to the wild-type *aglA* gene no AGL could be detected at all.

Also the GLA$_{G2}$ levels in these strains were determined by Western blot analysis (Fig 4A). All strains produced a small amount of the endogenous GLA in D-xylose medium (see previous §). In medium of AWGLA$_{KEX}$AGLS transformants about 30 - 40 mg l^{-1} of GLA$_{G2}$ could be detected. In medium of AWGLA$_{KEX}$AGL or AWGLAAGL, both containing the wild-type *aglA* gene, no GLA$_{G2}$ could be detected. The protein levels are summarized in Table II.

Also in this case, the protein levels were calculated on molar basis and given relatively to the amount of GLA$_{G2}$ to compare GLA$_{G2}$ levels with AGL levels (Table II). The amount of GLA$_{G2}$ in medium of AWGLA$_{KEX}$AGLS was about 15-20% of the GLA$_{G2}$ level of AWGLA. The level of AGL in medium of AWGLA$_{KEX}$AGLS was more or less comparable (10-12%), suggesting that almost no intracellular degradation of AGL had occurred.

Analysis of intracellular protein levels only showed very low AGL levels indicating that no accumulation had occurred (data not shown).

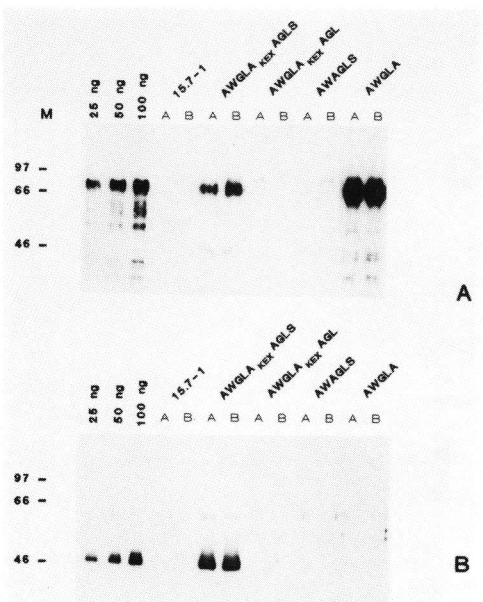


Fig.4 Western blot analysis of medium samples taken after induction with D-xylose. Blots were incubated with antisera against glucoamylase (Panel A) and α-galactosidase (Panel B). Each panel contains resp. a concentration range of purified protein, followed by a 1 μl medium sample (Panel A) or a 15 μl medium sample (Panel B) of 38-hours cultures. The purified GLA protein is a deglycosylated GLA$_{G1}$ sample. The letters A and B above the lanes indicate samples of duplicate cultures. A marker for molecular weight (kD) is indicated on the left.

Analysis of steady-state mRNA levels in α-galactosidase strains

To investigate whether fusion of *glaA* with *aglA* or $aglA_{syn}$ resulted in increased mRNA levels, Northern blot analysis was performed. The mRNA levels at both 22 and 38 hours after induction were determined using the same DNA fragment as probe as described before (M&M, Gouka et al. 1996b). From Figure 3 it is clear that a fusion of *glaA* with the $aglA_{syn}$ gene resulted in a large increase in steady-state mRNA. The level was approximately 25-fold higher than the mRNA level observed for the non-fused $aglA_{syn}$ mRNA (not clearly visible in Fig. 3). However, this mRNA level was still about 4-fold lower than that observed for non-fused *glaA* mRNA (AWGLA).

Fusions of *glaA* with the *aglA* wild-type gene, either in $AWGLA_{KEX}AGL$ or in AWGLAAGL, resulted in a smaller mRNA molecule than the expected size of 3.2 kb. Instead, a 2.2 - 2.3 kb mRNA was observed which was at the same position on a Northern blot as the *uidA* mRNA as could be shown by hybridization with an *aglA* or *glaA* specific probe (Fig. 3C). These experiments showed that the fusion mRNA was just a fraction larger than the *glaA* mRNA, indicating that a large part of the *aglA* gene was probably missing. Similar results (truncated *aglA* mRNA) had been observed for the non-fused *aglA* gene (Gouka et al. 1996b).

Analysis of truncated (glaA-)aglA mRNA

To determine which part(s) in the *glaA-aglA* mRNA was missing, Northern blot analysis was performed with probes derived from different parts of the *aglA* gene (Figure 5).

Strong hybridization was only observed when DNA fragment I, comprising the first 200 bp of the sequence encoding the mature AGL, was used as a probe. A much weaker hybridization signal was observed when fragment II (nucleotides 180-377) was used. With fragments III, IV and V, comprising sequences downstream of nucleotide 378, no hybridization with the major truncated transcript was observed. However, with fragment III a faint hybridization signal was observed with a slightly larger mRNA. This hybridizing mRNA was also visible when fragments I and II were used, although the intensity of the signal was approximately 20 - 50 fold less than the major hybridization signal. Together, these results indicate that the majority of the truncated mRNAs lacks at least more than 700 nucleotides of the 3' part of the *aglA* coding sequence.

In order to investigate whether a poly(A) tail was present in the *glaA-aglA* truncated mRNA molecule, poly(A) mRNA was isolated from strains containing *glaA* fusions with either the wild-type or the synthetic *aglA* gene and analyzed by Northern blot using a DNA fragment comprising the *aglA* gene as a probe. With both strains more than 90% of the hybridizing RNA

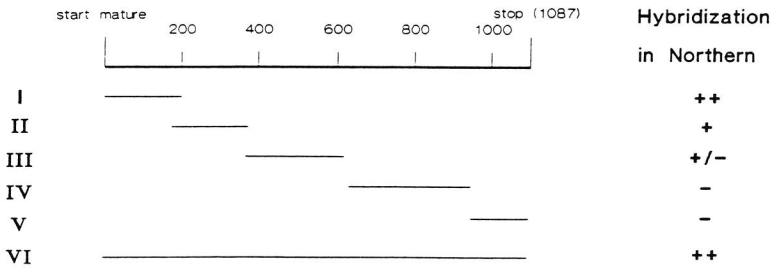


Fig.5 Schematic representation of the DNA sequence encoding mature guar α-galactosidase. Roman numbers correspond to the DNA fragment used for hybridisation on a Northern blot. ++ strong hybridisation, + weak hybridisation, +/- weak hybridization with a larger mRNA transcript, but not with the major truncated transcript, - no hybridisation

was present in the poly(A)$^+$ fraction which clearly indicated that also the truncated mRNA was polyadenylated (data not shown).

To further identify the 3' end of the truncated mRNA, RT-PCR was carried out using poly(A) mRNA isolated from multi copy strains of AWGLAAGL and AWAGL2 (see M&M). A 220 bp fragment was obtained and cloned in pUC19. Restriction enzyme analysis of the vectors showed that a small percentage (< 5%) of the inserts was slightly larger then 220 bp. Based on this restriction enzyme analysis four plasmids containing the 220 bp fragment were analyzed by DNA sequence analysis. This analysis showed that the four cDNAs contained the first 196-200 bases of the AGL encoding sequence, followed by a poly(A) tract (Figure 6). Nucleotide sequence analysis of some plasmids containing a larger cDNA insert showed that these fragments indeed corresponded with a slightly larger *aglA* mRNA.

To determine whether the truncation of the *glaA-aglA* mRNA was due to premature termination nuclear run-on transcription assays were carried out. Since the specific mRNA levels of the single copy strains were too low to carry out reliable run-on assays (data not shown), multi copy strains were used which had a higher *aglA* specific mRNA level. Different fragments of the *aglA* gene, generated by PCR, were slot blotted and hybridized with radioactively synthesized RNA. Figure 7 shows the results of run-on transcription assays of multi copy strains of AWGLAAGL, AWAGL2, AWAGLS and, as a control, AWGLA which contained a single copy of the *glaA* gene at the *pyrG* locus. Radioactive mRNA isolated from AWGLAAGL, AWAGL2 and AWAGLS hybridized with all *aglA* DNA fragments tested, indicating that in all three strains transcription proceeds at least 400 bases from the start of the mature

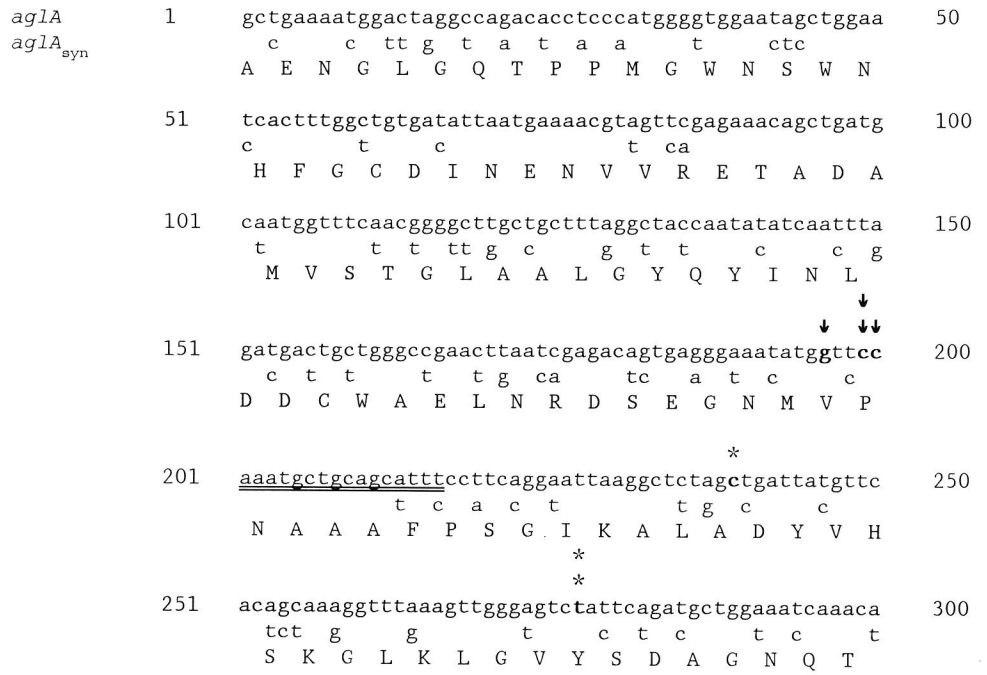


Fig.6 Part of the DNA sequence encoding mature guar α-galactosidase (start = 1). The upper row shows the wild-type *aglA* gene, the middle row the synthetic ($aglA_{syn}$) gene. Only differences are indicated. The lower row shows the amino acid sequence. Bold arrows indicate the last nucleotide prior to the polyadenylation signal of the major transcripts (> 95%). An asterisk marks the last nucleotide prior to the polyadenylation signal of some of the minor transcripts (< 5%). The 16-bp inverted repeat is double underlined.

AGL encoding sequence. Comparison of the signal intensities observed for the different hybridizing DNA fragments with those in a similar blot hybridized with a ^{32}P-labelled *aglA* DNA fragment showed an identical pattern, indicating that the differences observed were due to the decreasing length of the DNA fragments used on the blot. No *aglA* specific signal was observed when RNA from AWGLA, which did not contain the *aglA* gene, was used as a probe. From the results shown in Figure 7 it can also be seen that in AWGLAAGL a high level of *glaA* mRNA, most likely as part from the *glaA-aglA* fusion, is present, whereas this level is much lower in the strains that do not contain the *glaA* fusion. Also in AWGLA the *glaA* level was low, probably due to the presence of a single copy of the *glaA* gene only. In all strains

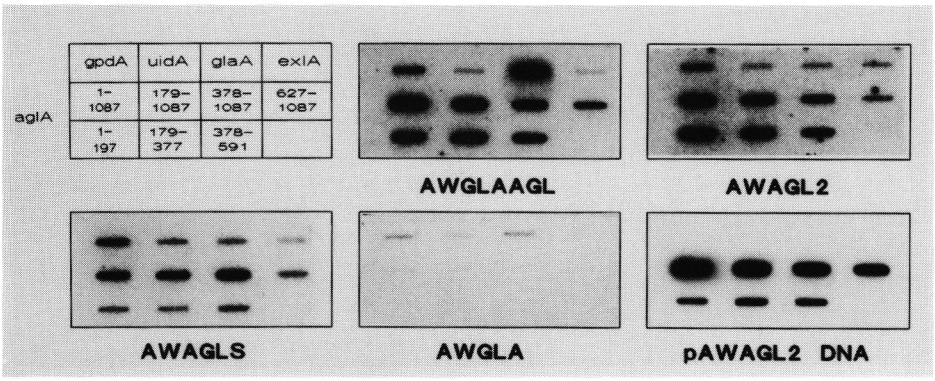


Fig.7 Run-on transcription assay with nuclei from induced cultures of the multi copy strains AWGLAAGL, AWAGL2, AWAGLS, and a single copy strain of AWGLA. 1 μg of several PCR-derived DNA fragments was slot-blotted on membranes and hybridized with labeled RNA isolated from the nuclei after run-on transcription in the presence of [^{32}P]-UTP. Numbers below *aglA* refer to the corresponding nucleotide sequence in the *aglA* DNA sequence counted from the start of the mature protein (start = 1, stop = 1087). The DNA fragments used for the genes in the upper row comprised the total coding region. Panel E shows a filter that was hybridized with a [^{32}P]-labelled DNA fragment containing the *aglA* gene.

there is a low level of *uidA* and a higher level of *gpdA* transcription. Since all strains lack the *exlA* gene, this gene was used as a negative control. In most of the strains a signal is absent, however, some faint background signal is sometimes visible.

The results of these experiments indicate that transcription elongation had occurred at least 400 bases downstream of the *aglA* mRNA truncation as was previously determined by Northern blot analysis (Fig. 5) and RT-PCR (Fig. 6). Whether transcription termination was correct could not really be observed in this way, but the similar results obtained with both *aglA* and $aglA_{syn}$ in a run-on transcription assay suggests that termination is correct.

In conclusion, the results with the nuclear run-on transcription assays, showing identical results for both wild-type and synthetic *aglA*, indicate that truncation of *aglA* mRNA is a result of incorrect processing.

DISCUSSION

The use of a gene-fusion strategy for the production of heterologous proteins has been subject of several studies. Although precise data are sparse, it is generally believed that a gene-fusion has its beneficial effect mainly on (post-)translational level, e.g. by enhanced stabilization, increased efficiency of secretion or prevention from proteolytic degradation.

The effect of gene fusion on hIL6 production

In a previous study we showed that although relatively high hIL6 mRNA levels were obtained, very low amounts of hIL6 protein were produced (Gouka et al. 1996b). Also in other members of the genus *Aspergillus* production of hIL6 was low (Carrez et al. 1990, Broekhuijsen et al. 1993). Inefficient intracellular transport and/or extracellular proteolytic degradation have been proposed as explanations for this phenomenon. As has been shown in previous experiments (Gouka et al. 1996b) hIL6 is stable in our culture medium, excluding extracellular proteolytic degradation as the reason for the absence of hIL6 in the growth medium. By fusing GLA_{G2} to the N-terminal end of hIL6 we showed that, as in *A. niger* (Broekhuijsen et al. 1993) and *A. nidulans* (Contreras et al. 1991), also in *A. awamori* hIL6 production could be highly improved. Interestingly, in our case levels of 5-10 mg l^{-1} of secreted hIL6 were obtained with a single copy strain whereas in *A. niger* and *A. nidulans* these amounts were obtained with strains containing multiple copies of the fusion construct, which illustrates that *A. awamori* can be used efficiently for the production of heterologous proteins.

The steady-state *glaA-hil6* fusion mRNA level was still lower than that of non-fused *glaA* mRNA suggesting that hIL6 contains motifs which influence the stability of the mRNA.

To investigate whether inefficient protein synthesis or secretion could be an explanation for the initial low hIL6 levels in non-fused hIL6 strains, a GLA_{G2} fusion to the C-terminus of hIL6 was constructed. With this strain not only very low levels of hIL6 but also of GLA_{G2} were obtained. Northern blot analysis showed that the hIL6-GLA steady-state mRNA levels were almost similar to the levels of GLA-hIL6 mRNA excluding that differences were located at the transcriptional level. These results indicate that the main bottle neck in hIL6 production probably occurs at the (post-)translational level, either by inefficient translation, inefficient translocation or rapid degradation which could be due to improper folding.

Although the GLA_{G2} amounts, produced by $AWGLA_{KEX}HIL6$, were almost as high as with a strain containing non-fused GLA_{G2}, much less hIL6 was produced (see Table II). Since

separation of both proteins by cleavage of the KEX2 site is thought to occur in the late Golgi, part of the hIL6 protein - 60-90% - must have been degraded (either intracellular or by mycelium associated proteases).

The effect of gene fusion on AGL production

The few data reported in literature concerning mRNA levels of fused and non-fused genes indicate that the effect of a gene-fusion on steady-state mRNA levels can either be advantageous or disadvantageous. Van Hartingsveldt et al. (1990) showed that when chymosin was fused to the first 71 AA of GLA as compared with fusions to the first 18 or 24 AA, the mRNA level dropped with a factor 2.5, suggesting a disadvantageous effect. Interestingly, the mRNA levels of non-fused chymosin were already high. A decrease was also suggested when the total coding region of GLA was fused to prochymosin (Ward et al. 1990). In contrast, slight improvements of mRNA levels were reported for a fusion of GLA to HEWL (Jeenes et al. 1994) and fusion of CBHI to Fab (Nyyssönen et al. 1995).

Our study showed that fusion of glucoamylase to the N-terminal end of guar α-galactosidase (encoded by $aglA_{syn}$) resulted in a 25-fold increase in AGL protein levels as compared to a non-fused AGL_{syn} strain. Northern blot analysis showed that, in contrast to the results obtained with hIL6, this increase was mostly due to higher steady-state mRNA levels. These results clearly show that a GLA fusion does not only have influence at (post-) translational level, but also can have a drastic effect at the transcriptional level. Since the mRNA level observed for the fusion mRNA was still about 4-fold lower than the level of non-fused GLA_{G2} mRNA, it is clear that, similar as found for fusions of hIL6 with *glaA*, the chimeric mRNA is less stable.

Strains containing expression cassettes with the wild-type *aglA* gene revealed truncated *aglA* mRNA molecules which only contained the first 200 bases of the *aglA* transcript. Nuclear run-on transcription assays showed that transcription elongation had proceeded at least 400 bases downstream of the truncation, indicating that premature termination was not the reason for the truncation.

The presence of a defined although truncated, specific mRNA signal and the presence of a poly(A) tail implies that the truncation occurs post-transcriptional, probably due to incorrect processing by the polyadenylation complex. This processing involves recognition of a specific DNA sequence. In higher eukaryotes this signal is a highly conserved hexanucleotide sequence (AAUAAA), which is present approximately 15-25 nucleotides upstream of the poly(A) addition site (Humphrey and Proudfoot 1988, Manley 1988, Wickens 1990). In

addition, a less conserved U- or GU-rich sequence is usually present downstream of the polyadenylation site (Humphrey and Proudfoot 1988, Manley 1988, Wickens 1990). In contrast, in the yeast *Saccharomyces cerevisiae* a diffuse AU-rich element is usually found instead of the mammalian consensus poly(A) signal (Proudfoot 1991). None of these motifs is sufficient to direct 3' end processing, and therefore, several structural motifs have to act in concert for efficient 3' end formation (Irniger et al. 1991; Guo and Sherman 1995).

For filamentous fungi no experimental data are available about 3' end formation of mRNAs. In some genes an AAUAAA sequence has been found, whereas in others it is absent (Gurr et al. 1988; Bruchez et al. 1993), indicating that also in fungi other motifs or combinations of motifs might be necessary to act as polyadenylation signal. When a consensus polyadenylation signal has been identified, the functionality of it has never been proven experimentally.

In the case of the truncated *aglA* gene, an AAUAAA sequence upstream of the poly(A) tract is absent. Analysis of the *aglA* sequence and subsequent comparison with the $aglA_{syn}$ gene, which does produce full length *aglA* mRNA, reveals some putative AT-rich regions upstream of the poly(A) tail which might be involved in incorrect processing. This signal is not present in the $aglA_{syn}$ gene.

Similar results have been reported for a number of heterologous genes which were expressed in yeasts. Also in that case the corresponding proteins were not detectable, which was due to the presence of putative internal yeast transcriptional terminator sequences resulting in truncated mRNA. The problem was first described with the highly AT-rich *Clostridium tetani* tetanus toxin fragment C gene in *S. cerevisiae* and was solved by synthesis of a gene with increased GC-content, from 29% to 47% (Romanos et al. 1991). Identical problems have been observed for at least four other genes (reviewed by Romanos et al. 1992). With one of these genes, encoding HIV-1 envelope protein (Scorer et al. 1993), truncated mRNA was produced in *Pichia pastoris*, whereas in *S. cerevisiae* full length mRNA was found. Although, in contrast to tetanus toxin fragment C, HIV Env is only slightly AT-rich, but contains AT-rich runs which might have caused the problem in this case. Full length mRNA could be obtained by construction of a synthetic HIV Env gene with increased GC-content, especially eliminating AT-rich runs. However, in these studies the authors performed no experiments to discriminate between premature termination of the RNA polymerase II and incorrect processing as the reason for the truncation.

Also remarkable in *C. tetragonoloba aglA* is the presence of a perfect inverted repeat of 16 bases (AAATGCTGCAGCATTT) immediately downstream of the polyadenylation site of the

aglA gene which might be involved in incorrect processing. In the $aglA_{syn}$ gene this repeat is reduced to only 8 bases.

Finally, it is interesting to note that truncation of *aglA* mRNA has also been observed when the wild-type *aglA* gene was expressed in *A. niger* and *A. nidulans* (PJ Punt, unpublished results). This would suggest that the problems observed with expression of the guar *aglA* gene are specific for *Aspergillus* species (if not for more fungal species), since in all other species reported, *Bacillus subtilis* (Overbeeke et al. 1990), *Hansenula polymorpha* (Fellinger et al. 1991), *S. cerevisiae* (Verbakel 1991) and *Kluyveromyces lactis* (Bergkamp et al. 1992), this was not the case. This might indicate that transcription termination and/or correct 3' end formation of fungal mRNAs might be different from other organisms, as was also suggested for the yeasts *S. cerevisiae* and *P. pastoris* (Scorer et al. 1993).

In conclusion, our results have clearly shown that production of heterologous proteins is limited both at the transcriptional and at the (post-)translational level. Limitations at the mRNA level are determined by the coding region of the heterologous gene, resulting in either low mRNA stability or in truncated mRNA, possibly due to AT-rich regions. At the (post-) transcriptional level, intracellular degradation is the major factor. Degradation can occur at least at two locations. First, during or immediately upon translation, as was observed for hIL6 and a hIL6-GLA fusion, preventing the synthesis of the fusion protein, and second, after cleavage of hIL6 from GLA in a GLAHIL6 fusion protein. As this cleavage is thought to occur in the Golgi, degradation occurs afterwards.

Finally, we showed that glucoamylase gene fusions can largely alleviate limitations at the transcriptional level and at the (post-)translational level.

Acknowledgements.

The research described in this paper was partially paid by Unilever Research Laboratorium, Gist-brocades, Solvay Duphar, DSM and Ministry of Economic Affairs. We wish to thank Daniëlle Remmerswaal for technical assistance, Prof. Dr. Ir. C.T. Verrips for critical reading of the manuscript and Dr. J.G.M. Hessing, Dr. W. Musters, Dr G.C.M. Selten, Dr. D. Carrez and Dr. C. Groeneveld for fruitful discussions.

Submitted

CHAPTER 6

Kinetics of mRNA and protein synthesis of genes controlled by the 1,4-ß-endoxylanase A promoter in controlled fermentations of *Aspergillus awamori*

Robin J. Gouka[1)], Hein Stam[2)], Arthur J. Fellinger[2)], Rob Muijsenberg[2)], Arjan van den Wijngaard[2)], Peter J. Punt[1)], Wouter Musters[2)], and Cees A.M.J.J. van den Hondel[1)]

[1] TNO Nutrition and Food Research Institute, Department of Molecular Genetics and Gene Technology, P.O. Box 5815, NL-2280 HV Rijswijk, the Netherlands, and [2] Department of Gene Technology and Fermentation, Unilever Research Laboratorium, Olivier van Noortlaan 120, NL-3133 AT Vlaardingen, the Netherlands

SUMMARY

Induction and repression kinetics of the expression of the *Aspergillus awamori* 1,4-ß-endoxylanase A (*exlA*) gene was analyzed at mRNA and at protein level under defined physiological conditions. Induction was analyzed by pulsing D-xylose to a sucrose-limited continuous culture of an *A. awamori* EXL overproducing strain. Directly after the D-xylose pulse *exlA* mRNA was synthesized, and reached a constant maximal level after 45-60 minutes. This level was maintained as long as a D-xylose/sucrose feed was used. When a D-xylose pulse was followed by a sucrose feed, *exlA* mRNA could no longer be detected as soon as D-xylose was depleted from the medium. Repression of *exlA* expression was analyzed by pulsing D-glucose in a D-xylose-limited continuous culture. Immediately after the glucose pulse the *exlA* mRNA level declined rapidly and reached a minimal level after 60-90 minutes. The half-life of the *exlA* mRNA was determined to be approximately 20-30 minutes. Kinetics of mRNA synthesis of the heterologous genes encoding *Thermomyces lanuginosus* lipase (*lplA*) and *E. coli* ß-glucuronidase (*uidA*), which were under control of the *exlA* promoter, was similar as observed for *exlA* mRNA. The time span between mRNA synthesis and secretion of the proteins was determined for EXL and LPL. In both cases mRNA became visible after approximately 7.5 minutes, whereas after 1 hour protein became detectable in the medium.

INTRODUCTION

Filamentous fungi, especially members of the genus *Aspergillus*, have the ability to secrete large amounts of a wide range of different enzymes into their environment. This characteristic has favoured the use of these organisms for the large scale production of commercially important enzymes (reviewed by a.o. Van den Hondel et al. 1991; Fowler and Berka, 1991; Verdoes et al. 1995).

However, unlike homologous proteins the secreted yields of most heterologous non-fungal proteins are very low and reach levels that are often 2-3 orders of magnitude lower. Therefore, in our laboratory a study was started to obtain more insight in the parameters that influence the production and secretion of heterologous proteins.

For the production and secretion of proteins in large quantities efficient gene expression systems are required (Van den Hondel et al. 1991; Fowler and Berka 1991). During the years a number of such expression systems has been developed for *Aspergilli*, most of them based on the *A. niger* or *A. awamori* glucoamylase (*glaA*) promoter (reviewed by Davies 1994). Recently, in our laboratory, an alternative expression system was developed for *A. awamori* based on the expression signals of the *A. awamori exlA* gene encoding 1,4-ß-endoxylanase A (Hessing et al. 1994; Gouka et al. 1996a).

Studies in shake flask cultures indicated that the expression of the *A. awamori exlA* gene is highly inducible with D-xylose (Gouka et al. 1996a) and that its expression control region can be used efficiently for the production of homologous (1,4-ß-endoxylanase A, Hessing et al. 1994) and heterologous (Gouka et al. 1996b) proteins in *A. awamori*. Initial analysis of shake flask cultures indicated that induction of P*exlA* driven expression was regulated at the transcriptional level by the presence of D-xylose in the culture medium (Gouka et al. 1996a). Although studies on the induction of 1,4-ß-endoxylanases have been described in a number of reports for other species, as yet nothing is known about the kinetics of P*exlA* driven expression. In fact, studies on the kinetics of mRNA and protein synthesis in fungi are sparse. To obtain more knowledge about the kinetics of mRNA and protein synthesis in fungi and in particular of P*exlA* driven expression in *A. awamori*, continuous culture fermentations were carried out with some of our recombinant strains.

Fermentation of filamentous fungi, especially in continuous culture fermentations, has shown to be very difficult due to (i) problems to acquire a constant, stable culture morphology and effluent control and (ii) accumulation of mycelium on the wall, pipes and valves of the vessel, especially in long term continuous cultures (our unpublished data). For studies on the

kinetics of mRNA synthesis or degradation, which can usually be performed in a relatively short term fermentation, the use of continuous cultures is necessary, since a large number of samples can be taken in a relatively short period without disturbing the culture. In this study we have used continuous cultures of the filamentous fungus *A. awamori* to investigate the kinetics of induction and repression of the P*exlA* driven expression on mRNA and protein level, both for homologous and heterologous genes.

MATERIALS AND METHODS

Strains and media. All *A. awamori* strains used are derived from *A. awamori* ATCC 11358 (CBS 115.52). *A. awamori* 1A (Hessing et al. 1994) is a strain containing 10-15 copies of the *A. awamori exlA* gene integrated at the *exlA* locus (unpublished data). *A. awamori* AW15.7-1, AW14A, AWLPL1 and AWLPL2 are *pyrG*$^+$ transformants of *A. awamori* AW15.7 (*exlA*$^-$::*uidA*$^+$, Hm^R, *pyrG*$^-$), containing respectively, no additional *exlA* based expression cassette (Aw15.7-1) or a single copy of the *A. awamori exlA* gene (AW14A) or the *Thermomyces lanuginosus* lipase gene (AWLPL1, *lplA* preceded by the *lplA* preprosequence; AWLPL2, *lplA* preceded by the *exlA* preprosequence) at the *pyrG* locus (Gouka et al. 1996a; Gouka et al. 1996b)

Medium for the batch phase of the fermentation consisted of 15 mM sucrose, 170 mM NH_4Cl, 7 mM KH_2PO_4, 4 mM $MgSO_4.7H_2O$, 2 mM $CaCl_2.2H_2O$ and 0.1% yeast extract. For induction D-xylose (final concentration 33 mM l^{-1}) and for repression D-glucose (final concentration 40 mM l^{-1}) was added to the vessel. Medium for the feed during continuous cultivation was similar to the batch-phase medium except that the C-source was either 15 mM sucrose, or a combination of 15 mM D-xylose/15 mM sucrose, depending on the type of fermentation.

Bioreactor cultivations. Fermentation experiments were carried out in a 3 litre Chemoferm glass fermenter with baffles, provided with a magnetically driven stirrer (8 blades Rushton impeller) on the bottom plate. The dissolved oxygen was measured by an Ingold O_2 probe, pH by an Ingold pH probe and temperature with an PT 100 sensor. Gelman bacterial filters were used to filter incoming air and outlet gasses before analysis. The outlet gasses were cooled by a condenser and led to the mass-spectrometer (MM8-80, VG) for measurements and calculations. pH was regulated by 12.5% NH_4OH and 3M H_3PO_4 using a Watson Marlow peristaltic pump.

Two types of continuous cultures were carried out:

(I) In continuous cultures of *A. awamori* 1A the working volume of 2.5 litres was maintained by using a teflon overflow tube. The feed rate was checked with a burette during addition with a Meredos SPGL70 pump. The set points for fermentation were: pH 3.0, stirrer speed 225 rpm, 25°C, airflow 1.5 $l.min^{-1}$.

(II) In all other continuous cultures the working volume (2.5 l) and feed rate was checked by pump control via a fermentation control system with the package CDAS running on a μVAX. The set points for these fermentations were: pH 4.5, stirrer speed 225 rpm, 25°C, airflow 1.5 $l.min^{-1}$.

After inoculation, the pH was regulated at the moment it reached the desired value. The fermentation was switched from batch to continuous ($D=0.1\ h^{-1}$) when the culture reached the end of the exponential phase. Simultaneously, the stirrer speed was increased to maintain the DO_2 above 35%.

Induction or repression was started when a steady-state phase was achieved which was defined as constant outlet gasses and constant biomass during at least 20 hours (type I fermentations) or 50 hours (type II fermentations).

Recombinant DNA techniques. Standard recombinant DNA techniques were used for cloning procedures (Sambrook et al. 1989). *Aspergillus* chromosomal DNA and RNA was isolated as described by Kolar et al. (1988).

Analysis of samples. After pulsing, medium and mycelium samples were taken at intervals of a few minutes in the first hour and of several hours at the later stages of the fermentation. Analysis of mRNA was carried out by Northern blotting. mRNA levels were quantified by scanning the developed films with a densitometer (Ultroscan XL, Pharmacia LKB, Uppsala, Sweden).

Quantification of protein levels was carried out by enzyme activity assays or Western blot analysis (Sambrook et al. 1989). Total xylanase activity assays were carried out using the DNS method as described by Hessing et al. (1994). ß-Glucuronidase assays were performed according to Roberts et al. (1989). For determination of protein concentrations by Western blot analysis purified protein was used as a standard. Purified 1,4-ß-endoxylanase A and lipase protein and the respective polyclonal antibodies were obtained from Unilever Research Laboratorium. For detection the ECL Western blot detection kit (Amersham) based on peroxidase labelled antibodies were used. The protein amounts on a Western blot were also

quantified densitometrically. Total protein amounts were determined using the BioRad protein assay kit. Concentration of proteins was carried out as described by Broekhuijsen et al (1993).

The extracellular metabolites D-xylose, glucose and sucrose were measured by means of high-pressure liquid chromatography analysis and/or enzymatically on the Cobas Mira auto-analyzer (Hoffmann LaRoche).

For mycelial dry weight measurements 25 ml of homogenous culture sample was filtrated through Miracloth (Calbiochem). The mycelium was squeezed and dried overnight in a vacuum oven at 80°C. This method resulted in a variation which was less than 10%.

RESULTS

Kinetics of *exlA* mRNA and protein synthesis upon induction with D-xylose

To study induction, continuous culture fermentations were carried out with the EXL overproducing strain *A. awamori* 1A (Hessing et al. 1994) under carbon-limited conditions in medium with sucrose as non-inducing C-source (Gouka et al. 1996a). To achieve a high concentration of D-xylose in the fermenter vessel at once, which would allow immediate induction of the *exlA* promoter, at t=0 hours a pulse of D-xylose was added to the medium to obtain a final concentration of 33 mM l^{-1}. In the feed 15 mM sucrose was maintained as C-source. The kinetics of *exlA* mRNA synthesis was studied by Northern blot analysis of total RNA isolated from mycelium sampled at several time points after induction. As shown in Fig.1A, a very quick response in *exlA* mRNA synthesis occurred, visible by the appearance of *exlA* mRNA already after 6 minutes. After approximately 45-60 minutes the *exlA* mRNA reached a constant maximum level. This maximum mRNA level was maintained until D-xylose was no longer detectable which resulted in a rapid decline in mRNA level (data not shown). The mRNA levels at the different time points during induction were quantified and relatively expressed to the mRNA level at t=14 hours (Fig. 1B).

To analyze the kinetics of EXL protein synthesis the total xylanase activity was determined. Although *A. awamori* contains more than one endoxylanase of which EXL has only a minor contribution to the total xylanase activity (Hessing et al. 1994, Gouka et al. 1996a), in multi copy strain 1A EXL is responsible for almost all xylanase activity (Hessing et al. 1994). EXL activity accumulated rapidly in the medium upon induction, up to approximately t=14 hours (Fig. 1B), followed by a decrease to very low levels at later time points (data not shown).

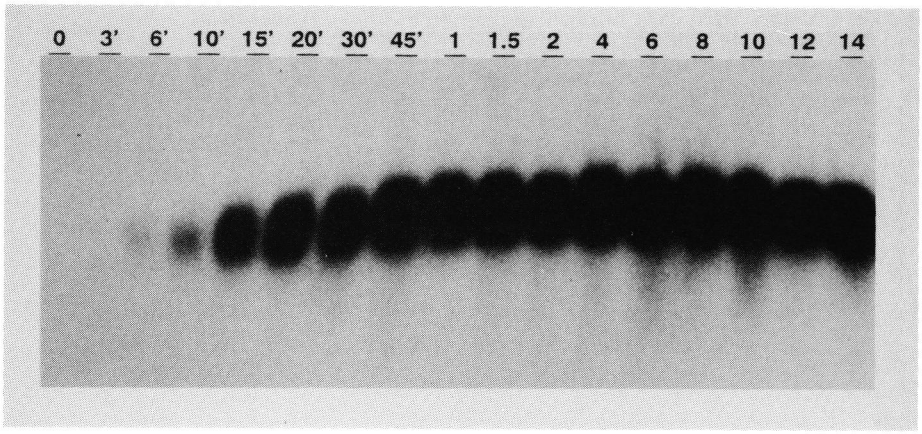


Fig. 1A Northern blot of *A. awamori* 1A containing equal amounts of total RNA isolated from mycelium samples obtained resp. 0, 3, 6, 10, 15, 20, 30, 45 minutes and 1, 1.5, 2, 4, 6, 8, 10, 12, and 14 hours after addition of D-xylose to a continuous culture grown in medium with sucrose as C-source. The blot was probed with a ^{32}P-labelled 0.7 kb *Bsp*HI/*Afl*II DNA fragment containing the *exlA* gene.

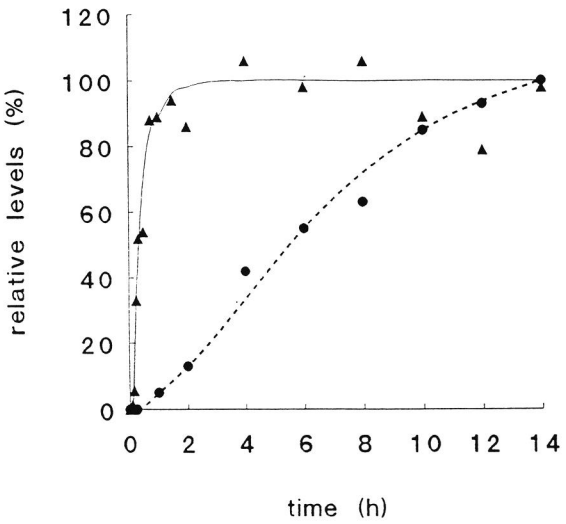


Fig. 1B Levels of *exlA* mRNA (-▲-) and EXL protein (-●-) after induction of a continuous culture of *A. awamori* 1A with D-xylose. The levels are given relatively to the maximum.

Together, these results show that induction with D-xylose results in a very quick response at the mRNA level followed by a much slower response at the protein level. EXL mRNA and protein levels declined rapidly as soon as D-xylose was depleted from the medium (data not shown).

Kinetics of *exlA* mRNA and protein degradation upon repression with D-glucose

To investigate the kinetics of glucose repression on *exlA* mRNA and protein levels, continuous cultures of *A. awamori* 1A were carried out. As a feed a combination of 15 mM D-xylose and 15 mM sucrose was used as C-source, which had previously been shown to be sufficient for induction of P*exlA* driven expression (Gouka et al. 1996a). At t=0 hours a pulse of 40 mM l^{-1} D-glucose was added, whereas the feed with D-xylose and sucrose was maintained.

At t=0 hours a high level of *exlA* mRNA was present (Fig. 2A), due to the presence of D-xylose in the feed. Very quickly after addition of glucose the amount of *exlA* mRNA decreased and reached a minimal level 60 - 90 minutes after glucose addition. After this time induction again occurred as shown by an increase of the *exlA* mRNA level at later time points, although not all glucose was consumed.

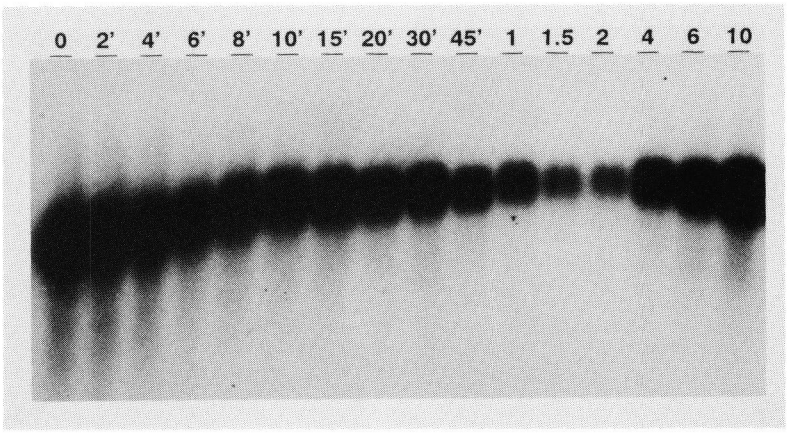


Fig. 2A Northern blot of *A. awamori* 1A containing equal amounts of total RNA isolated from mycelium samples obtained resp. 0, 2, 4, 6, 8, 10, 15, 20, 30, 45 minutes and 1, 1.5, 2, 4, 6, and 10 hours after addition of glucose to a continuous culture grown in medium with a combination of D-xylose/sucrose as C-sources. The blot was probed with a ^{32}P-labelled 0.7 kb *Bsp*HI/*Afl*II DNA fragment containing the *exlA* gene.

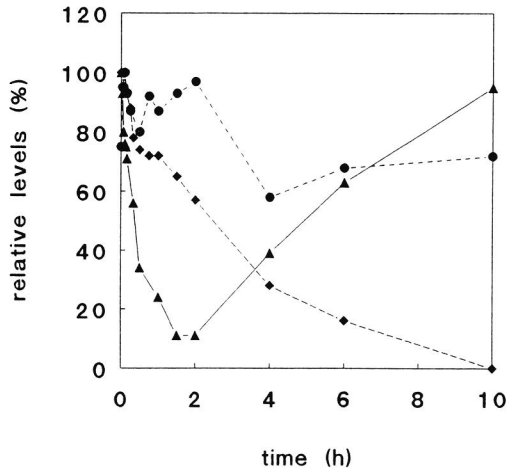


Fig. 2B Levels of *exlA* mRNA (-▲-), EXL protein (-●-) and glucose concentration (-◆-) after addition of glucose to a continuous culture of *A. awamori* 1A grown in medium with D-xylose as C-source. The levels are given relatively to the maximum level at t=0 hours.

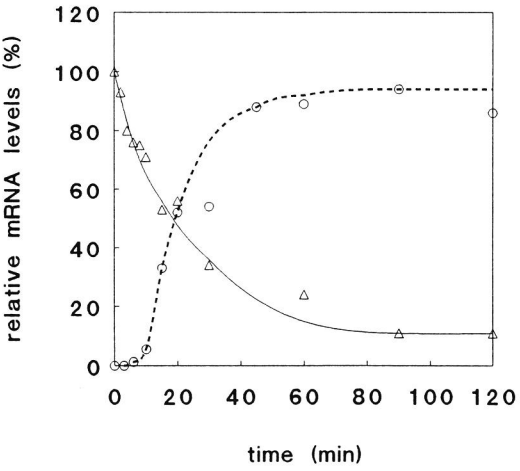


Fig. 2C Levels of *exlA* mRNA in *A. awamori* 1A after repression with glucose (-△-) or induction with D-xylose (-○-). The levels are given relatively to the maximum level at t=0 and t=14 hours, respectively.

The total xylanase activity (Fig. 2B) responded to the glucose pulse again with a lag, resulting in a drop of about 30-40% in EXL activity at t=4 hours. This relatively slight decrease is most likely caused by the short period of repression as was seen on mRNA level.

The glucose concentration (Fig. 2B) decreased from 40 mM directly after pulsing (t=0 hours), to undetectable levels, 10 hours later, indicating that glucose was consumed rapidly. Shortly after glucose addition to the fermenter, D-xylose started to accumulate, possibly due to a preferential uptake of glucose. The presence of both glucose and D-xylose in the medium might explain the increase in *exlA* mRNA level after t=2 hours. The induction rate was lower

as compared to the presence of D-xylose alone (Fig. 1A). Full induction was achieved when glucose levels became undetectable (t=10 hours).

Since repression at the mRNA level occurred almost immediately after addition of glucose to the vessel, the mRNA half-life could be determined. Based on Fig. 2B, the half-life of the *exlA* mRNA is approximately 20-30 minutes, assuming no de novo synthesis. In Fig. 2C also the induction rate at the mRNA level is shown during the first 2 hours. These results indicate that transcriptional regulation of the *exlA* promoter, either by induction or repression, is fast, allowing the fungus to respond very rapidly to environmental changes.

Comparison of kinetics of mRNA synthesis of *exlA*, *lplA* and *uidA* upon induction with D-xylose

To determine whether the use of the *exlA* expression signals for the expression of heterologous genes results in differences of mRNA kinetics as compared to *exlA*, continuous culture fermentations were carried out with *A. awamori* strains AW15.7-1, AW14A and, AWLPL1. These strains contain respectively a single copy of the *E. coli* ß-glucuronidase gene, the *A. awamori* 1,4-ß-endoxylanase A gene and the *T. lanuginosus* lipase gene under control of the *exlA* expression signals. At t=0, a pulse of D-xylose was given to sucrose-limited continuous cultures, resulting in a D-xylose concentration of 33 mM l^{-1}. The feed was changed from medium with 15 mM sucrose as single C-source to medium with both 15 mM D-xylose and 15 mM sucrose as C-sources, allowing efficient induction and maintenance of the induction levels.

Total RNA was isolated from mycelium samples taken at several intervals and analyzed by Northern blot analysis using DNA fragments containing the coding regions of the corresponding genes as probes.

To compare the induction rates, for each strain the mRNA levels were quantified densitometrically. For each strain the mRNA level at t=6 hours was defined as 100% and the other mRNA samples were given relatively to this level (Fig.3). For all strains (including multi copy strain *A. awamori* 1A), the kinetics of specific mRNA induction (*exlA*, *lplA* and *uidA*) was similar. Specific mRNA was visible 7.5 minutes after induction and reached a maximum level after approximately 60 minutes. This maximum level was maintained during cultivation by the presence of D-xylose in the feed. These results indicate that the kinetics of mRNA synthesis was not influenced by the presence of other coding sequences as *lplA* and *uidA*.

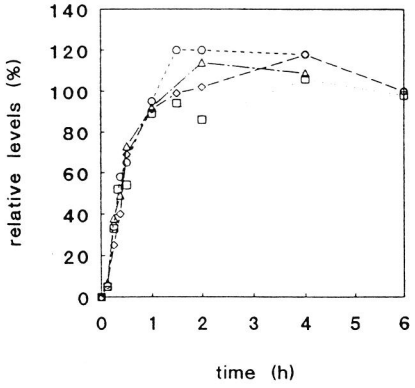


Fig. 3 Kinetics of mRNA synthesis of *A. awamori* 1A (-□-), AW14A (-◇-), AWLPL1 (-○-) and AW15.7-1 (-△-). The levels are given relatively to the mRNA level at t=6 hours in each strain.

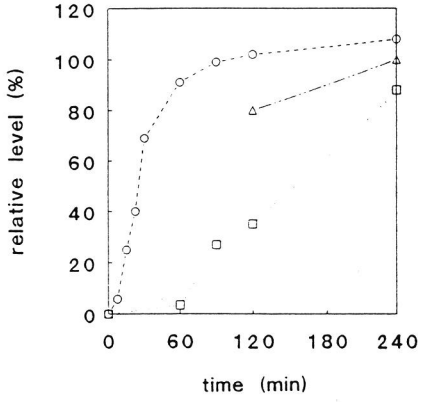


Fig. 4A Levels of *exlA* mRNA (-○-) and intra- (-△-) and extracellular (-□-) EXL protein after induction of a continuous culture of *A. awamori* AW14A with D-xylose. The levels are given relatively to the level at t=6 hours.

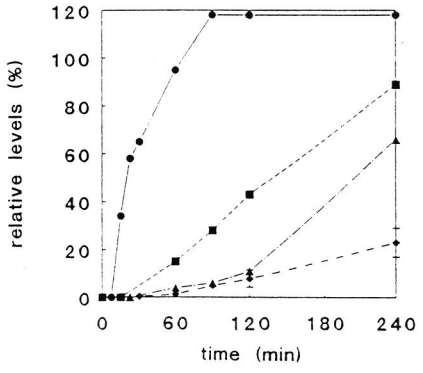


Fig. 4B Levels of lipase mRNA (-●-) and intra- (-■-) and extracellular (-▲-) LPL protein after induction of a continuous culture of *A. awamori* AWLPL1 with D-xylose and levels of intracellular GUS (-◆-) after induction of continuous cultures of *A. awamori* AW14A, AWLPL1, AWLPL2 and AW15.7-1 with D-xylose. The levels are given relatively to the level at t=6 hours.

Kinetics of protein synthesis of intra- and extracellular EXL and LPL and intracellular GUS upon induction with D-xylose

To compare the kinetics of intracellular synthesis and secretion of a homologous (EXL) and a heterologous (LPL) protein, Western blot analysis was carried out with cellular extracts and medium samples of AW14A and AWLPL1 taken at different time points. The protein amounts were quantified using a densitometer and given relatively to the level at t=6 hours, since at this time point both intra- and extracellular protein levels had more or less reached a maximum.

Intracellular EXL (Fig.4A) accumulated rapidly and a maximum was already reached between t=2 and t=4 hours. Extracellular EXL (Fig.4A) became detectable approximately 60 minutes after induction. The protein levels reached a maximum level at approximately t=6 hours.

Intracellular LPL could be observed in the cell extract at t=1 hour, and the amounts increased until a maximum was reached between t=4 and t=6 hours (Fig. 4B). Secretion of LPL also started at t=1 hour (Fig. 4B), but a constant level was obtained between 6 and 8 hours (data not shown).

The single copy strains AW14A, AWLPL1, AWLPL2 and the control strain AW15.7-1 all contain the *uidA* gene under control of the *exlA* expression signals (M&M). To determine the kinetics of synthesis of GUS, a cytoplasmic protein, in all these strains the GUS activity was determined. GUS protein became clearly detectable after 1 hour and slowly accumulated intracellularly. At t=6 hours the GUS levels were only 30-40% of the constant level which was reached at approximately t=24 hours (Fig. 4B).

DISCUSSION

Induction of the expression of the *exlA* gene in an *A. awamori* multi copy strain with D-xylose resulted in a very quick response at the mRNA level. *exlA* mRNA became visible already 6 minutes after induction and within 1 hour a maximum level was obtained. EXL protein still accumulated during the first 14 hours, indicating that a maximum protein level was not reached. At 14 hours, when D-xylose was no longer detectable, de novo mRNA synthesis stopped and *exlA* mRNA was degraded rapidly. As a result also de novo synthesis of EXL protein stopped and EXL levels declined due to wash out of the culture medium.

Addition of glucose to a D-xylose induced culture of *A. awamori* 1A resulted in rapid repression of the P*exlA* driven expression as shown by a fast decrease in *exlA* mRNA levels

from 100 to about 10% in 2 hours. Although at t=2 hours glucose was still present in the medium, induction started again as shown by an increase in mRNA, probably due to induction by D-xylose which was present in the feed and accumulated after glucose addition. The increase in *exlA* mRNA level was much slower as observed in the induction experiment (Fig. 1B) and is likely the result of the combination of D-xylose and glucose which are simultaneously present in the medium. Only when very low glucose levels were reached (t=10 hrs), full induction was achieved again.

As expected, the growth rate (μ) of the fungus will increase after pulsing D-xylose or glucose to the culture vessel. This could indeed be observed by an increase in biomass (data not shown), which results in a disturbance of the existing steady-state and the installation of a new steady-state after a certain period. The newly added C-source will therefore not only be used for induction or repression but also for growth, which might influence induction kinetics at the mRNA and protein level. A second factor that could influence the induction kinetics is the fact that induction will not be synchronous in all cells.

The induction kinetics of *exlA* mRNA synthesis and of *exlA* mRNA degradation is fast. Induction results in detectable mRNA levels after 6 minutes and repression in a mRNA half-life of approximately 20 minutes. However, the half-life might be an underestimation since in our cultivation D-xylose was still present in the feed and might have influenced the half-life by de novo synthesis of mRNA. Nevertheless, these results indicate that transcriptional regulation of P*exlA* driven expression is fast, enabling the fungus to respond rapidly to environmental changes. Data on half-lives of fungal mRNAs have not been described in detail, however, data from *S. cerevisiae* are available. Sierkstra et al (1992) showed that the half-lives of glucose repressible genes were very low. The lowest half-live, about 30-45 s, was observed for HXK1 and GAL2. GAL10 and SUC2 mRNA had a half-life of 1 minute and GAL7 and GAL1 of about 2 minutes. Others (Brown 1988) demonstrated that the half-lives of yeast mRNAs vary between 1 and 100 minutes.

Analysis of kinetics of mRNA synthesis of a homologous (*exlA*) and two heterologous genes (*lplA* and *uidA*) revealed that the rate of synthesis was identical in all cases. Shake-flask experiments, previously carried out, showed that the mRNA levels of these genes were different (Gouka et al. 1996b). Together, these results indicate that not the transcription rate is important for the differences observed between expression levels of homologous and heterologous genes. Apparently, other factors like variations in the mRNA stability, determined by the coding sequence, since all other flanking sequences are identical, are more important.

In our fermentation experiments the kinetics of production and secretion of EXL was slightly

faster than that of LPL. An explanation for this might be that translation and/or translocation of LPL is not as fast as EXL, possibly due to differences in codon usage or that LPL, which is known to stick to surfaces, is temporarily retained to or in the cell wall or cell organelles.

In conclusion, we have shown that continuous cultures of filamentous fungi, in particular *A. awamori*, are very useful to study kinetics of induction and repression of gene expression at the molecular level. Our results showed that expression of the *A. awamori* 1,4-ß-endoxylanase A gene can be induced rapidly with D-xylose or repressed with D-glucose. Induction at the transcriptional level was independent of the coding region of the gene following the *exlA* transcription control region; *exlA*, *lplA* and *uidA* mRNA displayed a similar induction pattern. Slight differences occurred in kinetics of intra- and extracellular protein accumulation, suggesting that both production and secretion of EXL was somewhat faster than LPL. Intracellular accumulation of GUS, a cytoplasmic protein, was slow.

ACKNOWLEDGEMENTS

Prof. Dr. Ir. C.T. Verrips is greatly acknowledged for discussions and critical reading of the manuscript.

CHAPTER 7

GENERAL DISCUSSION

1. INTRODUCTION

Filamentous fungi, especially members of the genus *Aspergillus*, have the ability to secrete large amounts of different enzymes into their environment, a characteristic which has favoured their use for the large scale production of commercially important enzymes (reviewed by a.o. Van den Hondel et al. 1991; Fowler and Berka, 1991; Verdoes et al. 1995). However, unlike the high yields obtained for homologous proteins the secreted yields of most heterologous proteins are 100- to 1000-fold lower. The research described in this thesis is carried out to obtain more insight in the factors that influence the production levels of heterologous proteins in filamentous fungi. Therefore, a comparative study was started in which the production levels of different proteins in defined single-copy *A. awamori* strains were analyzed. The expression of three fungal genes was studied, namely *A. awamori* 1,4-ß-endoxylanase (*exlA*), *A. niger* glucoamylase (*glaA*) and *Thermomyces lanuginosus* lipase (*lplA*) and two non-fungal genes, human interleukin-6 (*hil6*) and *Cyamopsis tetragonoloba* α-galactosidase (Chapter 4). The latter gene was available in two versions, a plant cDNA clone (*aglA*) and a synthetic gene with yeast-optimized codon bias ($aglA_{syn}$). To evaluate the effects of a recently developed glucoamylase gene-fusion approach, also the expression of *glaA*-hIL6, hIL6-*glaA* and *glaA*-*aglA* gene-fusions was studied (Chapter 5). All expression vectors were integrated in a single copy at the *pyrG* locus of *A. awamori* (Chapter 2) and all genes were placed under control of the efficient 1,4-ß-endoxylanase A promoter and transcription terminator (Chapter 3). The use of defined strains and a defined induction system excludes a number of parameters that might influence the production levels of proteins, such as gene copy number and genomic integration site, and allows a quantitative comparison of the results obtained with the different genes.

2. LIMITATIONS IN HETEROLOGOUS PROTEIN PRODUCTION BY FILAMENTOUS FUNGI

In agreement with the results obtained with other fungal proteins (reviewed by Van den Hondel et al. 1991; Verdoes et al. 1995), high extracellular protein levels were obtained with

the three fungal genes (*exlA*, *glaA* and *lplA*). Similar as generally found for non-fungal proteins (Van den Hondel et al. 1991), for both non-fungal proteins hIL6 and α-galactosidase, the protein levels were 100- to 1000-fold lower. Analysis of the steady-state mRNA levels showed that for both hIL6 and α-galactosidase production low mRNA levels accounted (partly) for the low protein levels observed. Compared to the *exlA* mRNA level, the level of hIL6 was 4-8 fold lower and of α-galactosidase about 100-fold lower. Fusion of *glaA* to the 5' end of either $aglA_{syn}$ or hIL6, increased both the mRNA levels and the yields of secreted protein, compared to the levels of the corresponding non-fused strains, whereas with a *glaA* fusion to the 3' end of hIL6 protein, levels similar as observed for a non-fused hIL6 strain were found (Chapter 5). For both fused and non-fused wild type *aglA* only truncated transcripts were observed.

From these data it is clear that at the transcriptional level, at least, two different factors can influence the level of heterologous gene expression: mRNA stability (see section 2.1.A) and incorrect processing of pre-mRNAs (see section 2.1.B).

For both non-fused and fused α-galactosidase encoded by the synthetic gene, the protein levels corresponded more or less with the amount expected on basis of the mRNA levels, indicating that for AGL no major limitations occurred at the (post-)translational level. Detailed analysis of the hIL6 protein levels obtained with the fused and non-fused genes in relation to the corresponding mRNA levels showed a discrepancy indicating that also a limitation at the (post-)translational level occurred (see section 2.2).

2.1 Limitations at the transcriptional level

2.1.A. mRNA stability

Low steady-state mRNA levels can result from either a low transcription (initiation) rate or a low mRNA stability. The presence of identical 5' and 3' expression signals derived from the *exlA* gene, however, makes differences in transcription initiation unlikely. Therefore, both hIL6 and $aglA_{syn}$ mRNA probably have a lower stability than *exlA*, *lplA* and *glaA* mRNA. From experiments carried out in mammalian cells, yeast cells and cell-free systems at least five structural components of the mRNA have shown to strongly influence its stability: (1) the 5' 7-methylguanosine triphosphate cap and, (2) the 3'-poly(A) tail; (3) the mRNA length; (4) posttranscriptional base modifications such as methylation of adenine residues or the conversion of adenines to inosines; (5) mRNA stabilizing or destabilizing sequences (see Santiago et al. 1986; Brown et al. 1988; Brown 1989; Sachs 1993; Raué 1994; Ross 1995). In particular (de)stabilizing elements have been studied and have shown to occur in each of three distinct regions: the 5' untranslated region, the coding region and the 3' untranslated

region (o.a. Raué 1994). From research in *S. cerevisiae* destabilizing elements have been identified in the coding region of HIS3 (Herrick and Jacobson 1992), STE3 (Heaton et al. 1992) and MATα1 (Parker and Jacobson 1990, Caponigro et al. 1993). Degradation of mRNAs can be influenced by the mRNA structure (i.e. polysome structure) or a high content of rare codons, which results in temporary translation arrest and subsequent degradation of the transcript (Hentze 1991; Ross 1995; Caponigro et al. 1993).

In filamentous fungi, mRNA stability has not been studied yet. Therefore, for hIL6 and α-galactosidase it is not known whether mRNA structure or codon usage is responsible for the low mRNA levels. In *S. cerevisiae* and *A. nidulans* codon usage was shown to be related to expression: highly expressed genes often have a high percentage of C's as the last base of a triplet (De Boer and Kastelein, 1986; Lloyd and Sharp, 1991). In contrast, α-galactosidase contains a relatively high amount of codons rarely used in *Aspergilli*, ending at an A, which might be responsible for the low levels of mRNA observed.

From the increased mRNA levels obtained with *glaA* fusions, we can conclude that *glaA* stabilizes the hIL6 or *aglA*$_{syn}$ transcripts, although the levels were lower than non-fused *glaA* mRNA. Apparently, *glaA* contains stabilizing sequences which partly overcome the destabilizing elements in hIL6 and *aglA*$_{syn}$. Similar results in filamentous fungi have been observed with GLA-hen egg white lysozyme fusions in *A. niger* (Jeenes et al. 1994) and CBHI-antibody fragment fusions in *T. reesei* (Nyyssönen et al. 1995), although the 25-fold increase of a *glaA-aglA*$_{syn}$ compared to non-fused *aglA*$_{syn}$ is clearly higher than observed in these other examples.

2.1.B. Incorrect processing of pre-mRNAs

The results obtained with wild type α-galactosidase showed that not only the mRNA level of the heterologous gene was lower, but that also incorrect processing of mRNA influences the levels of heterologous protein. For wild type *aglA*, a truncated transcript was observed which only contains 200 bases of the *aglA* gene. Nuclear run-on transcription assays showed that improper transcription elongation or premature termination were not the reason for the generation of truncated mRNAs. The fact that the truncated *aglA* mRNA was polyadenylated indicates that truncation has occurred by incorrect processing of the *aglA* pre-mRNA in the nucleus. Truncation of *aglA* mRNA had also been observed when the wild type *aglA* gene was expressed in *A. niger* and *A. nidulans* (PJ Punt, unpublished results). In all other species in which guar α-galactosidase has been produced (*Bacillus subtilis* (Overbeeke et al. 1990), *Hansenula polymorpha* (Fellinger et al. 1991), *S. cerevisiae* (Verbakel et al. 1991) and

Kluyveromyces lactis (Bergkamp et al. 1992) incorrect processing was not observed, indicating that the problems are specific for *Aspergillus*.

As described in Chapter 1, pre-mRNA processing in eukaryotes generally involves the recognition of an AU-rich polyadenylation signal. In the truncated *aglA* mRNA a consensus polyadenylation signal (AAUAAA) upstream of the poly(A) tract is absent, but several AU-rich regions are present which are absent in the $aglA_{syn}$ gene where incorrect processing did not occur. In yeast AU-rich sequences in the coding region of a number of heterologous genes acted as internal yeast polyadenylation sequences resulting in truncated mRNAs. When the highly AU-rich *Clostridium tetani* tetanus toxin fragment C gene was expressed in *S. cerevisiae*, full length transcripts could only be obtained by generating a synthetic gene with increased GC-content, from 29% to 47% (Romanos et al. 1991). Similar situations have been observed for other genes (Romanos et al. 1992; Scorer et al. 1993). With one of these genes, encoding HIV-1 envelope protein, a truncated mRNA was produced in *Pichia pastoris*, although in *S. cerevisiae* full length mRNA was found. In contrast to tetanus toxin fragment C, the overall sequence of HIV Env was not AU-rich, but contained AU-rich stretches. Full length mRNA could be obtained by construction of a synthetic HIV Env gene with increased GC-content, especially eliminating the AU-rich runs. Comparison of the plant *aglA* and $aglA_{syn}$ sequences indicates that also in our case less AU-rich stretches in the processing area, as in $aglA_{syn}$, might play a role in obtaining full length mRNA.

2.2 Limitations at the (post-)translational level

From the experiments described in this thesis it can be concluded that not only low mRNA levels accounted for the low levels of secreted heterologous proteins. The results obtained with the non-fused and fused genes indicate that the efficiency of secretion of heterologous proteins is also determined at the (post-)translational level.

Although for both non-fused and fused $aglA_{syn}$, a correlation was found between mRNA levels and AGL protein levels, indicating that no major limitations are occurring at the (post-)translational level. Comparison of the mRNA and protein levels obtained after *hil6*, *glaA-hil6* and *hil6-glaA* expression clearly revealed that the hIL6 protein levels do not correlate with the mRNA levels, but are much lower. Since hIL6 is not degraded in the culture medium these results indicated that the major limitation for hIL6 production must be at the (post-) translational level. This means that this limitation probably takes place during translation of mRNA and/or during subsequent translocation, folding, processing and transport of the (fusion) protein through the secretory pathway.

The secretion pathway of proteins involves a number of complex processes carried out in several cellular compartments. The secretion pathway has been studied primarily in yeasts and mammalian cells and from these organisms a number of proteins and their encoding genes involved in several steps in the secretion process have been isolated and characterized. Secretion of proteins involves the following processes:

Translation and translocation

Translation of mRNAs from secreted proteins starts in the cytoplasm, where ribosomes bind to the mRNA molecule. As soon as a small part of the protein, the signal peptide, has emerged from the ribosome, it is bound by the signal recognition particle (SRP). Next, the complex, consisting of the mRNA/ the nascent peptide chain/ ribosome and SRP is targeted to the ER membrane by an interaction with a membrane-bound SRP receptor or docking protein (see Chapter 1). The SRP is then released from the complex, which requires GTP hydrolysis, and the nascent peptide chain is transferred into the membrane (Rapoport 1992).

Based on research in yeast and mammals translocation across the ER-membrane occurs via a pore consisting of several protein components (Simon and Blobel, 1991). As soon as the translocated protein enters the ER lumen, processing of the N-terminal signal peptide occurs by the ER-membrane bound signal peptidase (YaDeau et al. 1991; Larriba, 1993) and maturation starts. Upon entry of newly synthesized proteins in the ER-lumen, the chaperone protein BiP (in yeast called KAR2, Normington et al. 1989; Rose et al. 1989) becomes associated, keeps them in an unfolded state and prevents them from aggregation before proper folding or subunit assembly has occurred (Kim et al. 1992). BiP has also been shown to associate with aberrantly folded proteins, so preventing them to leave the ER (Ng et al 1992).

Quality control in the ER

Maturation of newly synthesized proteins involves folding, glycosylation, subunit assembly and other protein modifications like phosphorylation. Folding of the nascent protein is catalyzed by two enzyme families (Gething and Sambrook, 1992), protein disulphide isomerases (PDI and other thioredoxin like proteins; reviewed by Freedman et al. 1994), which catalyzes disulphide bond formation, and peptidyl prolyl *cis-trans* isomerases (PPIases; reviewed by Schmid et al. 1993), catalyzing the isomerization of X-proline peptide bonds, where X can be any amino acid. Quality control of protein structure again involves chaperones (Kelley and Georgopoulos, 1992). Recently, Hammond and Helenius (1995)

proposed a model for an ER-network in which chaperones together with specific ER-located enzymes act as a quality control system for newly synthesized proteins. The key enzymes of the system are the ER-luminal chaperones calreticulin (recently cloned from *S. cerevisiae* by Parlati et al. 1995), its membrane bound counterpart calnexin, and UDP-glucose:glycoprotein glucosyltransferase (UDP-Glc transferase). Upon entry in the ER, newly synthesized proteins undergo cycles of glycosylation, deglycosylation and reglycosylation. UDP-Glc transferase distinguishes between correctly and incorrectly folded glycoproteins by only recognizing the latter. Correctly folded glycoproteins are recognized specifically by calnexin/calreticulin via which they leave the ER-compartment. As this only accounts for glycosylated proteins, also other types of interactions are possible as was shown for BiP which was shown to specifically bind hydrophobic regions (Georgopoulos and Welch 1993).

Modifications/processing in the Golgi compartment and secretion

After leaving the ER, the proteins are transported in vesicles to the Golgi compartment. In the Golgi compartment processes such as glycosylation and processing (e.g. of propeptides from the mature protein) occurs. Finally, the protein is transported to the cell wall, again in vesicles, and is secreted into the extracellular medium.

Although analysis of the secretion pathway in filamentous fungi has, thusfar, not been studied in detail, ultrastructural electromicroscopic studies suggest that in filamentous fungi these processes do not really differ from yeast and mammalian cells (Punt et al. 1994; Keijzer et al. in prep). A hypothetical representation of the secretion pathway in filamentous fungi is given in Figure 1 (p. 116).

With non-fused hIL6 and with a *glaA* fusion to the 3' end of hIL6, the secreted yields were more than 100-fold less as expected on basis of mRNA levels. Also with a *hil6-glaA* fusion, very low levels of both GLA_{G2} and hIL6 were obtained, which may suggest that a limitation has occurred during the translation/translocation across the ER membrane and/or quality control in the ER compartment. A limitation in translation/translocation can be caused by the presence of codons which are inefficiently used in *Aspergillus*, which can lead to a low translation rate, errors, frame-shifts and degradation of the protein, as has been observed when proteins containing rare codons were produced in *E. coli* (reviewed by Kane 1995). Due to a low translation rate, translocation of the nascent protein into the ER might be impaired, which can result in the release of the protein in the cytoplasm where it is degraded. A limitation by quality control can be triggered by incorrect folding and/or incorrect

glycosylation, after which the quality control system initially will prevent hIL6 to leave the ER, resulting in hIL6 degradation. For bovine chymosin it was shown that the introduction of a N-glycosylation site into the protein resulted in a 10-fold increase in chymosin levels, suggesting that glycosylation is indeed an important parameter (Berka, 1991). Although hIL6 contains two N-glycosylation sites, secreted hIL6 is not glycosylated in *Aspergillus* and *S. cerevisiae* (Carrez et al. 1990). One might hypothesize that N-glycosylation of hIL6 may result in incorrect folding and subsequent degradation of hIL6, whereas only the fraction of non-glycosylated hIL6 is normally secreted.

With a GLA_{G2}-hIL6 fusion higher yields of hIL6 were obtained. However, the yield was only 10-20% of the yields of GLA_{G2} originating from the same fusion protein, indicating that after proteolytic processing of the fusion protein, probably occurring in the Golgi compartment, the hIL6 part of the fusion protein will be degraded intracellularly. Degradation of hIL6 and AGLA by extracellular proteases was not observed, however, degradation by mycelium-bound proteases cannot be excluded.

At this moment, more experiments, e.g. a combination of pulse-chase labeling and cell fractionation studies, have to be carried out to precisely define where the limitations in hIL6 are located.

3. IMPROVEMENT STRATEGIES AND PERSPECTIVE

3.1 Improvements at the (post-)transcriptional level

The results presented in this thesis indicate that at the transcriptional level at least two factors can determine the level of mRNA: (1) a low mRNA stability and (2) incorrect processing of the transcripts. Both factors are determined by the coding region of the heterologous gene and result in either low mRNA levels and/or truncated mRNAs. Several strategies could lead to improvement of mRNA level and correct processing:

(1) Changing the coding sequence of a gene to optimize mRNA stability and processing.

Improvement of mRNA stability by the removal of destabilizing sequences or the introduction of stabilizing sequences is difficult unless it is known which sequences in the coding region are influencing mRNA stability. Nevertheless, in case a synthetic gene is constructed, it is advisable to adapt the codon usage to the one as used in highly expressed genes in the host organism (De Boer and Kastelein, 1986; Lloyd and Sharp, 1991). Incorrect processing can be prevented by elimination of AT-rich regions in the transcript of a heterologous gene by construction of synthetic genes as has been shown in *Saccharomyces*

(Romanos 1992 and references therein). Similarly, the data in this thesis indicate that also in *Aspergillus* expression of a synthetic gene with less AU-rich regions results in full length mRNA (Chapter 4/5).

(2) Improvement of the mRNA stability by fusion with highly expressed genes.

Gene-fusion strategies have been developed for filamentous fungi especially to increase the amount of secreted protein (Chapter 1). Fusion of *glaA* to $aglA_{syn}$ or hIL6 increased the mRNA levels up to 25-fold (Chapter 5). Improvements of mRNA levels have also been observed with a GLA-hen egg white lysozyme fusion in *A. niger* (Jeenes et al. 1994) and a CBHI-antibody fragment fusion in *T. reesei* (Nyyssönen et al. 1995).

(3) Introduction of an intron sequence into the cDNA sequence of the heterologous gene

Such a strategy was shown to stabilize heterologous mRNA in plants. For example, the introduction of a castor bean catalase intron in the 5' end of the coding sequence of the *uidA* gene resulted in a 10 to 90-fold increase in GUS activity in transgenic rice, caused by an elevated *uidA* mRNA level (Tanaka et al. 1991). An identical strategy, however, did not work in tobacco (Tanaka et al. 1991). To date, such a strategy has never been reported for filamentous fungi.

3.2 Improvements at the (post-)translational level

As described in section 2.2, low levels of secreted protein may also be caused by limitations at the (post-)translational level, either initiated by inefficient translation/translocation, incorrect glycosylation/folding, and/or after processing in the late Golgi compartment. Several strategies could result in improved levels of secreted protein:

(1) Construction of a synthetic gene with a codon usage optimized for the host organism

As described for *E. coli* (Kane 1995), the presence of rare codons might negatively influence the production levels of heterologous proteins e.g. by the occurrence of translational errors. The construction of a synthetic gene with a codon usage optimized for *Aspergillus* species (Lloyd and Sharp, 1991) might improve the translation efficiency, resulting in increased protein levels.

(2) Fusion of a highly expressed gene to the 5' end of a heterologous gene

A limitation at the early stage of the secretion pathway (ER) can be solved by fusion of a

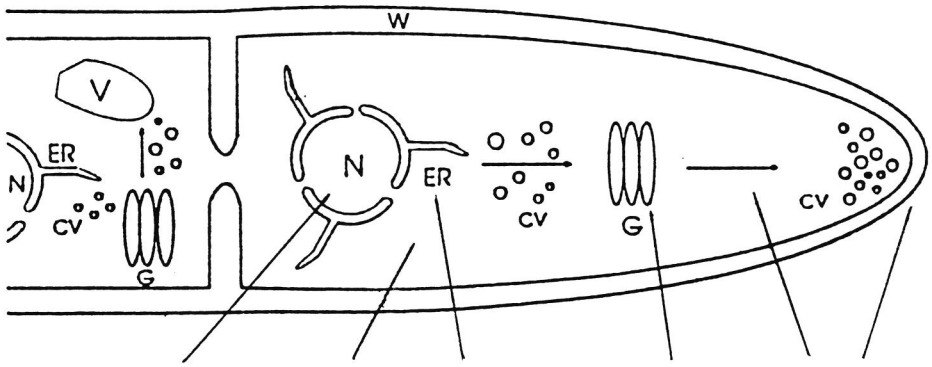
highly expressed gene to the 5' end of the heterologous gene (Chapter 1 and 5). This N-terminal protein (usually GLA or CBHI) possibly improves translocation and folding, thereby protecting the heterologous protein from recognition and degradation by the ER quality control system.

(3) Overproduction of foldases and/or chaperones

A different strategy to improve production of proteins is the overproduction of foldases and/or chaperones. Such a strategy will be useful if proteins are improperly folded. In *S. cerevisiae* secretion of intracellularly accumulated inactive heterologous proteins can be increased using strains that overproduce a foldase (PDI) or a chaperone (BiP). Overproduction of the KAR2 (BiP) protein by 4- to 5-fold resulted in a 5-fold overproduction of erythropoietin in *S. cerevisiae* (Wittrup and Robinson, 1994). Also, tenfold overproduction of BiP gave a 26-fold increase in secreted bovine prochymosin, but a similar approach did not result in elevated levels of secreted thaumatin (Harmsen 1995). Sixteen-fold overproduction of PDI increased the secretion of human platelet derived growth factor B homodimer with a factor 10, whereas *S. pombe* acid phosphatase secretion was increased fourfold (Robinson et al. 1994). Secretion of human granulocyte colony stimulating factor, however, was unaffected by PDI overexpression (Robinson et al. 1994). Overproduction of *A. niger* BiP did not result in increased secreted levels of hIL6 in an *A. niger* protease deficient strain (Punt et al. in prep). It is now interesting to also investigate whether PDI overexpression increases hIL6 secretion.

(4) Isolation and use of protease-deficient strains

The actual degradation of heterologous proteins occurs by proteases. These proteases may either be secreted and act in the extracellular medium or intracellularly, usually in the vacuole, or may be localized in the cell wall. Extracellular proteases, such as PEPA (Berka et al. 1990), are responsible for the degradation of a large number of heterologous proteins (Archer et al. 1992; Broekhuijsen et al. 1993). Fungal strains deficient for extracellular proteases have been isolated by random mutagenesis (Mattern et al. 1992; Van den Hombergh et al. 1995) or molecular genetic approaches (Berka et al. 1990) and the use of such strains has resulted in a marked improvement in the production of heterologous proteins (Berka, 1991; Roberts et al. 1992; Broekhuijsen et al. 1993). In this study it was shown that in our case the yield of secreted hIL6 was probably affected by intracellular degradation. Therefore, strains devoid of (certain) intracellular proteases may result in improved hIL6 yields.



	Nucleus	Cytoplasm	ER	Golgi	Post-golgi
Key processes	* pre-mRNA processing	* mRNA stability * translation	* signal recognition * translocation * folding * modification * glycosylation	* protein processing * glycosylation	* post-Golgi processes * cell wall passage
Putative bottlenecks	* AU-rich sequences	* mRNA destabilizing elements	* tertiary protein structure * low levels of foldases or chaperones * ER proteolytic degradation	* incorrect processing	* proteolytic degradation
Observed for	*aglA*	*aglA*, *aglA*$_{syn}$ *hil6*	*hil6*		*hil6* (*aglA*$_{syn}$)
Putative solutions	* synthetic genes with less AU-rich regions	* gene-fusion	* gene-fusion * overproduction of foldases or chaperones * protease-deficient strains	* over-production of processing enzymes	* protease-deficient strains

Figure 1. Generalized overview of the fungal secretion pathway and the putative limiting steps in heterologous protein production. C, cytoplasmic vesicles; ER, endoplasmic reticulum, G, Golgi(-like) compartment; N, nucleus; V, vacuole; W, cell wall.

The lower part shows the key processes, the putative bottlenecks and solutions in heterologous protein production and the genes for which such a bottleneck was observed (Figure adapted from Punt et al. 1994).

A strategy, similar as described for extracellular proteases, may be employed for the isolation of such intracellular protease mutant strains. For *S. cerevisiae*, strains are already available which are deficient in vacuolar proteases, and the use of these strains has resulted in increased levels of heterologous proteins (Suzuki et al. 1989; Wingfield and Dickinson 1993). Recently, genes encoding vacuolar proteases in filamentous fungi have been cloned (Jarai et al. 1994) and defined mutants can now be generated which will possibly improve heterologous protein levels.

It is clear that all the processes involved in protein synthesis and secretion are highly complicated and that each step in the secretion route can seriously influence the production of heterologous proteins. Figure 1 shows a simplified scheme of the fungal secretion pathway. This scheme summarizes the processes involved in the production and secretion of heterologous proteins together with the putative bottlenecks and solutions described for the proteins studied in this thesis. The initiation of more fundamental research towards the understanding of the processes of protein secretion will be required to eventually eliminate the limitations in secretion.

SUMMARY

Filamentous fungi are able to occupy many different environments like soil, organic waste or plant cell material. The fact that filamentous fungi can secrete a broad range of different enzymes, which enables them to use a wide variety of organic compounds as source for their nutrients, underlies this successful life style. Already more than a century ago, it became evident that fungi were able to produce and secrete these enzymes in large quantities, which has resulted in an increasing interest in studying and using filamentous fungi in industrial processes as well as in the laboratory.

The development of fungal recombinant DNA techniques (in particular that of DNA-mediated transformation) has greatly stimulated the use of fungi for the production of homologous and heterologous (especially mammalian) proteins and these techniques have partly replaced classical strain improvement strategies (such as mutagenesis). In particular species of the genera *Aspergillus* and *Trichoderma* are being used extensively for large scale production of proteins.

Production of fungal proteins by filamentous fungi is usually very efficient and production levels of grams per litre are within reach. Production levels of most non-fungal proteins, either of bacterial, avian, mammalian or plant origin, are usually two to three orders of magnitude lower than the levels of fungal proteins and do not exceed a few tens of milligrams per litre of culture medium. Unfortunately, most data reported in the literature are obtained from undefined, multi copy strains. To obtain more insight in the factors that are responsible for the low levels of heterologous protein, a study was started in which the production levels of a number of different proteins were systematically analyzed, using defined single copy strains of *A. awamori*.

To facilitate efficient selection of transformants containing a single copy of an expression vector integrated at a specific locus, a system was developed to increase the frequency of site-specific integration. This system is based on the transformation of a defined *pyrG* mutant strain, in which a specific mutation was introduced by transformation, with a vector containing a *pyrG* gene with a mutation different from the one in the genome (Chapter 2). Using this system, in 50% of the transformants the vector was integrated at the *pyrG* locus.

For the production of proteins in large quantities a new, efficient gene expression system, based on the expression signals of the 1,4-ß-endoxylanase A gene (*exlA*) was used. Induction

of the *exlA* promoter was studied both in shake-flask cultures (Chapter 3) and in continuous culture fermentations (Chapter 6). Regulation of *exlA* gene expression was shown to be at the transcriptional level. Induction was achieved with D-xylose, resulting in a rapid increase of specific mRNA. The *exlA* promoter was efficiently induced with D-xylose up to levels that were three-fold higher than the frequently used *glaA* promoter.

In Chapter 4 a systematic analysis is described to compare the production levels of a homologous protein with different heterologous - fungal and non-fungal - proteins using the transformation and expression systems described in Chapter 2 and 3. Analysis showed that large differences occurred at the steady-state mRNA levels obtained from the various genes, varying from high levels for genes of fungal origin (*A. awamori* 1,4-ß-endoxylanase A, *A. niger* glucoamylase and *Thermomyces lanuginosus* lipase) to low levels for genes of non-fungal origin (human interleukin-6 and *Cyamopsis tetragonoloba* (guar) α-galactosidase). Full length α-Galactosidase mRNA could only be obtained when a synthetic gene copy was used with an optimized yeast codon bias, but not with the wild type α-galactosidase gene. In all cases except for hIL6, the protein levels corresponded to the amount expected on basis of the mRNA levels. For hIL6, very low protein levels were observed whereas on basis of the steady-state mRNA levels higher protein amounts were expected.

In Chapter 5 experiments are described that were carried out to further investigate the problems observed with hIL6 and α-galactosidase production using a gene-fusion approach. The analyses were again carried out with defined single copy strains. This strategy enabled the comparison of the effect of a gene-fusion with the corresponding non-fused genes, both at the mRNA and at the protein level. For human IL-6, a *glaA* fusion to the 5' end of hIL6 resulted in 100 to 200-fold higher levels of secreted hIL6, whereas hIL6 production of strains with a *glaA* fusion to the 3' end remained very low. The mRNA levels of both fusions was nearly similar, indicating that differences occurred at (post-)translational level. Most likely hIL6 is incorrectly folded in the ER and subsequently degraded. After processing of the GLA-hIL6 fusion protein in the two separate proteins, less hIL6 was observed than GLA_{G2}, indicating that also after processing in the Golgi compartment a considerable fraction of hIL6 was degraded.

For guar α-galactosidase, *glaA* fusions with the wild type *aglA* gene still resulted in truncated mRNA lacking more than 80% of the *aglA* sequences. Nuclear run-on experiments indicated that transcription elongation was correct. The fact that the truncated *aglA* mRNA was polyadenylated indicated that incorrect processing of the pre-mRNA had occurred. Replacement of the wild type α-galactosidase gene for the synthetic one resulted in full length mRNA. Compared to non-fused $aglA_{syn}$ these mRNA levels were about 25-fold higher. As a

consequence also the α-galactosidase protein levels were approximately 25-fold higher.

In conclusion, our results have clearly shown that production of heterologous proteins is limited both at the transcriptional and at the (post-)translational level. Improvements at both the transcriptional level and at the (post-)translational level could be achieved with a *glaA* gene-fusion to the 5' end of the heterologous gene.

SAMENVATTING

Filamenteuze schimmels komen op de meest uiteenlopende plaatsen voor, zoals in aarde, op organisch afval of op plantecel materiaal. Deze succesvolle levenstijl is mogelijk doordat filamenteuze schimmels een diversiteit aan enzymen kunnen uitscheiden hetgeen hen in staat stelt een grote variatie aan organische verbindingen te gebruiken als voedingsbron. Al meer dan een eeuw geleden werd bekend dat schimmels deze enzymen in grote hoeveelheden kunnen maken en uitscheiden. Dit heeft geresulteerd in een toenemende interesse in het bestuderen en het gebruik van filamenteuze schimmels, zowel in industrieële processen als in het laboratorium. De ontwikkeling van recombinant DNA technieken voor schimmels (in het bijzonder transformatie met DNA) heeft het gebruik van schimmels voor de produktie van homologe en heterologe eiwitten (vooral afkomstig van zoogdieren) enorm gestimuleerd en inmiddels hebben deze technieken de klassieke stamverbeterings-strategieën zoals mutagenese gedeeltelijk vervangen. Met name soorten van het geslacht *Aspergillus* en *Trichoderma* worden veelvuldig gebruikt voor de produktie van eiwitten op grote schaal.

Produktie van schimmel eiwitten door filamenteuze schimmels is in het algemeen erg efficient en produktie niveau's van enkele grammen per liter liggen vaak binnen handbereik. De produktie niveau's van de meeste niet-schimmel eiwitten, afkomstig van bacterieën, vogels, zoogdieren of planten, liggen vaak honderd tot duizend maal lager dan de produktie niveau's van schimmel eiwitten en komen meestal niet uit boven enkele milligrammen per liter cultuur medium. Helaas zijn de meeste gegevens die in de literatuur worden vermeld verkregen met ongedefinieerde stammen waarin een verschillend aantal kopieën van het gen zijn ingebracht. Om meer inzicht te krijgen in de factoren die de produktie niveau's van heterologe eiwitten beïnvloeden, is een onderzoek gestart waarin de produktie niveau's van een aantal verschillende eiwitten systematisch zijn geanalyseerd. Hierbij werd gebruik gemaakt van gedefinieërde *A. awamori* stammen, waarin één kopie van een expressie plasmide aanwezig is op een specifieke plaats in het genoom.

Om een efficiente selectie van transformanten met één kopie van een expressie plasmide op een specifieke plaats in het genoom te vergemakkelijken, is een systeem ontwikkeld waarmee de frequentie van plaats-specifieke integratie kan worden vergroot. Dit systeem is gebaseerd op transformatie van een gedefinieerde *pyrG* mutant stam, waarin een specifieke

mutatie is geïntroduceerd door middel van transformatie, met een plasmide dat een *pyrG* gen bevat met een andere mutatie dan die in de acceptor-stam (Hoofdstuk 2). Met dit systeem werd in 50% van de gevallen transformanten verkregen waarin het plasmide in één kopie was geïntegreerd in het *pyrG* locus.

Voor de produktie van eiwitten in grote hoeveelheden werd een nieuw, efficient expressie systeem gebruikt, gebaseerd op de expressie signalen van het 1,4-ß-endoxylanase A gen (*exlA*). Inductie van de *exlA* promoter werd zowel in erlenmeyers (Hoofdstuk 2) als in continu cultures (Hoofdstuk 6) bestudeerd. De regulatie van de expressie van het *exlA* gen ligt op transcriptie niveau. Inductie werd bewerkstelligd met D-xylose, hetgeen resulteerde in een snelle toename van specifiek mRNA. Inductie van de *exlA* promoter met D-xylose resulteerde in een drie maal zo hoog niveau als inductie van de frequent gebruikte *glaA* promoter met maltodextrine.

In Hoofdstuk 4 wordt een systematische analyse beschreven om de produktie niveau's van een homoloog eiwit te vergelijken met die van verschillende heterologe - schimmel en niet-schimmel -eiwitten, waarbij het transformatie systeem en de expressie signalen werden gebruikt zoals beschreven in Hoofdstuk 2 en 3. Analyse toonde aan dat er grote verschillen bestonden tussen de mRNA niveau's van de verschillende genen, variërend van hoge niveau's voor genen afkomstig van schimmels (*A. awamori* 1,4-ß-endoxylanase A, *A. niger* glucoamylase en *Thermomyces lanuginosus* lipase) tot lage niveau's voor niet-schimmel genen (humaan interleukine-6 en *Cyamopsis tetragonoloba* α-galactosidase). Een volledig α-galactosidase transcript kon alleen worden verkregen indien een synthetisch gen werd gebruikt met een codon gebruik dat geoptimaliseerd is voor expressie in gist, maar niet met het wild type α-galactosidase gen. In alle gevallen behalve voor hIL6, correspondeerde de eiwit hoeveelheden met de hoeveelheden die werden verwacht op basis van de mRNA niveau's. Voor hIL6 werden zeer lage eiwit hoeveelheden gevonden, terwijl op basis van het mRNA niveau verwacht werd dat deze hoger zouden zijn.

In Hoofdstuk 5 worden experimenten beschreven die zijn uitgevoerd om de problemen die werden gevonden bij hIL6 en α-galactosidase produktie verder te onderzoeken, waarbij gebruik werd gemaakt van een gen-fusie aanpak. De analyses werden wederom uitgevoerd met gedefinieerde stammen waarin één kopie van de betreffende vector aanwezig was. Deze strategie maakt het mogelijk de effecten van een gen-fusie te vergelijken met de corresponderende niet-gefuseerde genen, zowel op mRNA als op eiwit niveau. Een *glaA* fusie aan de 5' kant van hIL6 resulteerde in hIL6 niveau's die 100- tot 200-maal hoger waren dan niet-gefuseerd hIL6, terwijl de hIL6 produktie van stammen met een *glaA* fusie aan de 3' kant

van hIL6 laag bleef. Het mRNA niveau van beide fusies was vrijwel gelijk, hetgeen erop duidde dat de verschillen op (post-)translationeel niveau moeten liggen. Waarschijnlijk wordt hIL6 verkeerd gevouwen in het ER, waarna het wordt afgebroken. Na klieving van het GLA-hIL6 fusie-eiwit in de twee afzonderlijke eiwitten werd tevens minder hIL6 teruggevonden dan GLA_{G2}. Ook na klieving in het Golgi compartiment verdwijnt er dus nog een redelijk grote hoeveelheid hIL6.

Fusies van *glaA* met het wild type α-galactosidase gen resulteerde in verkorte mRNAs, waarvan meer dan 80% van de α-galactosidase sequentie ontbrak. Experimenten waarin *in vitro*, met behulp van gezuiverde kernen, radioactief *aglA* mRNA werd gemaakt ("nuclear run-on transcription assays) toonde aan dat transcriptie elongatie in orde was. Het feit dat het verkorte mRNA een poly(A) staart heeft suggereert dat er incorrecte processing van het pre-mRNA is opgetreden. Vervanging van het wild type α-galactosidase gen door het synthetische gen resulteerde in volledig mRNA. Vergeleken met een stam waarin een niet-gefuseerd $aglA_{syn}$ gen aanwezig is, zijn de mRNA niveau's van het fusie gen 25 keer hoger. Als gevolg daarvan is ook de hoeveelheid α-galatosidase eiwit ongeveer 25-keer hoger.

Geconcludeerd kan worden dat de resultaten duidelijk aantonen dat produktie van heterologe eiwitten gelimiteerd kan zijn op zowel transcriptie als (post-)translatie niveau. Verbeteringen op transcriptie en (post-)translatie niveau konden worden bewerkstelligd met behulp van *glaA* gen-fusies.

REFERENCES

Akileswaran L, Alic M, Clark EK, Hornick JL, Gold MH (1993) Isolation and transformation of uracil auxotrophs of the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Curr Genet 23:351-356

Alic M, Akileswaran L, Gold MH (1993) Gene replacement in the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Gene 136:307-311

Archer DB, Roberts IN, MacKenzie DA (1990a) Heterologous protein secretion from *Aspergillus niger* in phosphate-buffered batch culture. Appl Microbiol Biotechnol 34:313-315

Archer DB, Jeenes DJ, MacKenzie DA, Brightwell G, Lambert N, Lowe G, Radford SE, Dobson CM (1990b) Hen egg white lysozyme expressed in, and secreted from *A. niger* is correctly processed and folded. Bio/Technology 8:741-745

Archer DB, MacKenzie DA, Jeenes DJ, Roberts IN (1992) Proteolytic degradation of heterologous proteins expressed in *Aspergillus niger*. Biotechnol Lett 14:357-362

Archer DB, Jeenes DJ, MacKenzie DA (1994) Strategies for improving heterologous protein production from filamentous fungi. Antonie van Leeuwenhoek 65:245-250

Atwater JA, Wisdom R, Verma IM (1990) Regulated mRNA stability. Annu Rev Genet 24:519-541

Bailey MJ, Poutanen K (1989) Production of xylanolytic enzymes by strains of *Aspergillus*. Appl Microbiol Biotechnol 30:5-10

Ballance DJ, Buxton FP, Turner G (1983) Transformation of *Aspergillus nidulans* by the orotidine-5'-phosphate decarboxylase gene of *Neurospora crassa*. Biochemical and Biophysical Research Communications 112:284-289

Ballance DJ (1991) Transformation systems for filamentous fungi and an overview of fungal gene structure. In: Molecular Industrial Mycology (eds. Leong SA, Berka RM), pp 1-29, Marcel Dekker, New York.

Baron M, Tiraby G, Calmels T, Parriche M, Durand H (1992) Efficient secretion of human lysozyme fused to the Sh *ble* phleomycin resistance protein by the fungus *Tolypocladium geodes*. J Biotech 24:253-266

Benito EP, Diaz-Minguez JM, Iturriaga EA, Campuzano V, Eslava AP (1992) Cloning and sequence analysis of the *Mucor circinelloides pyrG* gene encoding orotidine-5'-monophosphate decarboxylase: use of *pyrG* for homologous transformation. Gene 116:59-67

Bennett JW, Lasure LL (1991) Growth media. In: Bennett JW, Lasure LL (eds) More gene manipulations in fungi, Academic Press, San Diego, pp 441-458

Bergès T, Barreau C (1991) Isolation of uridine auxotrophs from *Trichoderma reesei* and efficient transformation with the cloned *ura3* and *ura5* genes. Curr Genet 19:359-365

Bergkamp RJM, Kool IM, Geerse RH, Planta RJ (1992) Multiple-copy integration of the α-galactosidase gene from *Cyamopsis tetragonoloba* into the ribosomal DNA of *Kluyveromyces lactis*. Curr Genet 21:365-370

Berka RM, Ward M, Wilson LJ, Hayenga KJ, Kodama KH, Carlomagno LP, Thompson SA (1990) Molecular cloning and deletion of the aspergillopepsin A gene from *Aspergillus awamori*. Gene 86:153-162

Berka RM, Bayliss FT, Bloebaum P, Cullen D, Dunn-Coleman NS, Kodama KH, Hayenga KJ, Hitzeman RA, Lamsa MA, Przetak MM, Rey MW, Wilson LJ, Ward M (1991) *Aspergillus niger* var. *awamori* as a host for the expression of heterologous genes. In: Applications of Enzyme Biotechnology (eds. Kelly JW, Baldwin TO), Plenum

Press, New York. pp 273-292

Berka RM, Dunn-Coleman NS, Ward M (1992) Industrial enzymes for *Aspergillus* species. In: *Aspergillus*, The biology and industrial applications (Eds. Bennet JW, Klich MA), pp 155-195, Butterwoth

Biely P (1985) Microbial xylanolytic systems. Trends Biotechnol 3:286-290

Bienroth S, Keller W, Wahle E (1993) Assembly of a processive messenger RNA polyadenylation complex. EMBO J 12:585-594

Bodie EA, Bower B, Berka RM, Dunn-Coleman NS (1994) Economically important organic acid and enzyme products. In: Martinelli SD, Kinghorn JR (Eds). *Aspergillus*: 50 years on, pp 561-602, Elsevier Science BV, Amsterdam, The Netherlands

Boeke JD, LaCroute F, Fink GR (1984) A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet 197:345-346

Boel E, Christensen T, Wöldlike HF (1987) Process for the production of protein products in *Aspergillus oryzae* and a promoter for use in *Aspergillus*. European Patent Application 0238023

Boel E, Huge-Jensen IB (1989) Recombinant *Humicola* lipase and process for the production of recombinant *Humicola* lipase. European Patent Application 0305216

Broekhuijsen MP, Mattern IE, Contreras R, Kinghorn JR, van den Hondel CAMJJ (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 of Biotechnology 31:135-145

Brown AJP, Purvis IJ, Santiago CT, Bettany AJE, Loughlin L, Moore J (1988) Messenger RNA degradation in *Saccharomyces cerevisiae*. Gene 72:151-160

Brown AJP (1989) Messenger RNA stability in yeast. Yeast 5:239-258

Bruchez JJP, Eberle J, Russo VEA (1993) Regulatory sequences in the transcription of *Neurospora crassa* genes: CAAT box, TATA box, introns, poly(A) tail formation sequences. Fungal Genetics Newsletters 40:89-96

Buxton FP, Radford A (1984) Cloning of the structural gene for orotidine 5'-phosphate carboxylase of *Neurospora crassa* by expression in *Escherichia coli*. Mol Gen Genet 190:403-405

Calmels TPG, Martin F, Durand H, Tiraby G (1991) Proteolytic events in the processing of secreted proteins in fungi. J Biotechnol 17:51-66

Caponigro G, Muhlrad D, Parker R (1993) A small segment of the MATα1 transcript promotes mRNA decay in *Saccharomyces cerevisiae*: a stimulatory role for rare codons. Mol Cel Biol 13:5141-5148

Carrez D, Janssens W, Degrave P, van den Hondel CAMJJ, Kinghorn JR, Fiers W, Contreras R (1990) Heterologous gene expression by filamentous fungi: secretion of human interleukin-6 by *Aspergillus nidulans*. Gene 94:147-154

Christensen T, Woeldike H, Boel E, Mortensen SB, Hjortshoej K, Thim L, Hansen MT (1988) High level expression of recombinant genes in *A. oryzae*. Bio/Technology 6:1419-1422

Christofori G, Keller W (1988) 3' cleavage and polyadenylation of mRNA precursors in vitro requires a poly(A) polymerase, a cleavage factor and a snRNP. Cell 54:875-889

Contreras R, Carrez D, Kinghorn JR, van den Hondel CAMJJ, Fiers W (1991) Efficient KEX2-like processing of a glucoamylase-interleukin-6 fusion protein by *A. nidulans* and secretion of mature interleukin-6. Bio/Technology 9:378-380

Coughlan MP, Hazlewood GP (1993) ß-1,4-D-xylan-degrading enzyme systems: biochemistry, molecular biology

and applications. Biotechnol Appl Biochem 17:259-289

Cullen D, Gray GL, Wilson LJ, Hayenga KJ, Lamsa MH, Rey MW, Norton S, Berka RM (1987) Controlled expression and secretion of bovine chymosin in *Aspergillus nidulans*. Bio/Technology 5:369-376

Dalbey RE, Von Heijne G (1992) Signal peptidases in prokaryotes and eukaryotes - a new protease family. TIBS 17:474-478

Davies RW (1994) Heterologous gene expression and protein secretion in *Aspergillus*. In:Martinelli SD, Kinghorn JR (Eds). *Aspergillus*: 50 years on, pp 527-560, Elsevier Science BV, Amsterdam, The Netherlands.

De Boer HA, Kastelein RA (1986) Biased codon usage: An exploration of its role in optimization of translation. In: Biotech series 9 (Eds. Reznikoff W, Gold L), pp 225-285

De Graaff LH, Vandenbroeck HC, Vanooijen AJJ, Visser J (1994) Regulation of the xylanase-encoding *xlnA* gene of *Aspergillus tubigensis*. Mol Microbiol 12:470-490

De Ruiter-Jacobs YMJT, Broekhuijsen M, Unkles SE, Campbell EI, Kinghorn JR, van den Hondel CAMJJ (1989) A gene transfer system based on the homologous *pyrG* gene and efficient expression of bacterial genes in *Aspergillus oryzae*. Curr Genet 16:159-163

Devchand M, Gwynne DI (1991) Expression of heterologous proteins in *Aspergillus*. J Biotechnol 17:3-10

Dunn-Coleman NS, Bloebaum P, Berka R, Bodie E, Robinson N, Armstrong G, Ward M, Przetak M, Carter GL, LaCost R, Wilson LJ, Kodama KH, Baliu EF, Bower B, Lamsa M, Heinsohn H (1991) Commercial levels of chymosin production by *Aspergillus*. Bio/Technology 9:976-981

Durand H, Clanet M, Tiraby G (1988) Genetic improvement of *Trichoderma reesei* for large scale cellulase production. Enzyme Microb. Technol. 10:341-345

Elliot S, Griffin J, Suggs S, Lau EP, Banks AR (1989) Secretion of glycosylated human erythropoietin from yeast directed by the α-factor leader region. Gene 79:167-180

Fellinger AJ, Verbakel JMA, Veale RA, Sudbery PE, Bom IJ, Overbeeke N, Verrips CT (1991) Expression of the α-galactosidase from *Cyamopsis tetragonoloba* (Guar) by *Hansenula polymorpha*. Yeast 7:463-473

Fernandez-Espinar MT, Ramon D, Pinaga F, Valles S (1992) Xylanase production by *Aspergillus nidulans*. FEMS Microbiology Letters 91:91-96

Finkelstein DB, Rambosek J, Crawford MS, Soliday CL, McAda PC, Leach J (1989) Protein secretion in *Aspergillus niger*. In: Genetics and Molecular Biology of Industrial Microorganisms (eds Hershberger CL, Queener SW, Hegeman G) American Society for Microbiology, Washington, DC, pp 295-300

Folz RJ, Gordon JI (1986) Deletion of the propeptide from human preproapolipoprotein A-II redirects cotranslational processing by signal peptidase. J Biol Chem 261:14752-14759

Fowler T, Berka RM, Ward M (1990) Regulation of the *glaA* gene of *Aspergillus niger*. Curr Genet 18:537-545

Fowler F, Berka RM (1991) Gene expression systems for filamentous fungi. Current Opinion in Biotechnology 2:691-697

Freedman RB, Hirst TR, Tuite MF (1994) Protein disulphide isomerase: building bridges in protein folding. TIBS 19:331-336

Frenken LGJ, Egmond MR, Batenburg AM, Bos JW, Visser C, Verrips CT (1992) Cloning of the *Pseudomonas glumae* lipase gene and determination of the active site residues. Appl Environ Microbiol 58:3787-3791

Gems D, Johnstone IL, Clutterbuck AJ (1991) An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. Gene 98:61-67

Georgopoulos C, Welch WJ (1993) Role of the major heat shock proteins as molecular chaperones. Annu Rev Cell Biol 9:601-634

Gething MJ, Sambrook J (1992) Protein folding in the cell. Nature 355:33-45

Ghosh M, Nanda G (1994) Physiological studies on xylose induction and glucose repression of xylanolytic enzymes in *Aspergillus sydowii* MG49. FEMS Microbiology Letters 117:151-156

Gilmore R (1993) Protein translocation across the endoplasmic reticulum: a tunnel with toll booths at entry and exit. Cell 75:589-592

Goossen T, Bloemheuvel G, Gysler C, de Bie DA, van den Broek HWJ, Swart K (1987) Transformation of *Aspergillus niger* using the homologous orotidine-5'-phosphate-decarboxylase gene. Curr Genet 11:499-503

Gouka RJ, van Hartingsveldt W, Bovenberg RAL, van den Hondel CAMJJ, van Gorcom RFM (1991) Cloning of the nitrate - nitrite reductase gene cluster of *Penicillium chrysogenum* and use of the *niaD* gene as a homologous selection marker. J of Biotechnology 20:189-200

Gouka RJ, Van Hartingsveldt W, Bovenberg RAL, Van Zeijl CMJ, Van den Hondel CAMJJ, Van Gorcom RFM (1993) Development of a new transformant selection system for *Penicillium chrysogenum*: isolation and characterization of the *P. chrysogenum* acetyl-coenzyme A synthetase gene (*facA*) and its use as a homologous selection marker. Appl Microbiol Biotechnol 38:514-519

Gouka RJ, Hessing JGM, Stam H, Musters W, van den Hondel CAMJJ (1995) A novel strategy for the isolation of defined *pyrG* mutants and the development of a site-specific integration system for *Aspergillus awamori*. Curr Genet 27:536-540

Gouka RJ, Hessing JGM, Punt PJ, Stam H, Van den Hondel CAMJJ (1996a) An expression system based on the promotor region of the 1,4-ß-endoxylanase A gene of *Aspergillus awamori*. Appl Microbiol Biotechnol, in press

Gouka RJ, Punt PJ, Hessing JGM, Van den Hondel CAMJJ (1996b) Heterologous gene expression in *Aspergillus awamori*: Comparison of expression levels of genes from different origins, Appl Environ Microbiol, submitted

Gruber F, Visser J, Kubicek CP, de Graaff LH (1990) Cloning of the *Trichoderma reesei pyrG* gene and its use as a homologous marker for a high-frequency transformation system. Curr Genet 18:447-451

Guo Z, Sherman F (1995) 3'-End-forming signals of yeast mRNA. Mol Cell Biol 15:5983-5990

Gurr SJ, Unkles SE, Kinghorn JR (1988) The structure and organization of nuclear genes of filamentous fungi. In: Gene structure in Eukaryotic Microbes (Ed. Kinghorn JR), SGM Spec. Publ. Vol. 23, pp 93-139, IRL Press, Oxford

Gwynne DI, Buxton FP, Williams SA, Garven S, Davies RW (1987) Genetically engineered secretion of active human interferon and a bacterial endoglucanase from *Aspergillus nidulans*. Bio/Technology 5:713-719

Gwynne DI, Buxton FP, Williams SA, Sills AM, Johnstone JA, Buch JK, Guo Z, Drake D, Westphal M, Davies RW (1989) Development of an expression system in *Aspergillus nidulans*. Biochemical Society Transactions 17:338-340

Gwynne DI, Devchand M (1992) Expression of foreign proteins in the genus *Aspergillus*. In: *Aspergillus*, The biology and industrial applications (Eds. Bennet JW, Klich MA), pp 203-214, Butterwoth

Haas IG, Wabl M (1983) Immunoglobulin heavy chain binding protein. Nature 306:387-389

Hahm YT, Batt CA (1990) Expression and secretion of thaumatin from *Aspergillus oryzae*. Agric Biol Chem 54:2513-2520

Hamer JE, Timberlake WE (1987) Functional organization of the *Aspergillus nidulans trpC* promoter. Mol Cell Biol 7:2352-2359

Hammond C, Helenius A (1995) Quality control in the secretory pathway. Current Opinion in Cell Biology 7:523-529

Harkki A, Uusitalo J, Bailey M, Pentilla M, Knowles JKC (1989) A novel fungal expression system: secretion of active calf chymosin from the filamentous fungus *Trichoderma reesei*. Bio/Technology 7:596-603

Harmsen MM (1995) Heterologous protein secretion by the yeast *Saccharomyces cerevisiae*. Thesis, Free University Amsterdam, the Netherlands

Heaton B, Decker C, Muhlrad D, Donahue J, Jacobson A, Parker R (1992) Analysis of chimeric mRNAs derived from the *STE13* mRNA identifies multiple regions within yeast mRNAs that modulate mRNA decay. Nucl Acids Res 20:5365-5373

Hentze M W (1991) Determinants and regulation of cytoplasmic mRNA stability in eukaryotic cells. Biochimica and Biophysica Acta 1090:281-292

Herrick D, Jacobson A (1992) A coding region segment is necessary, but not sufficient for rapid decay of the *HIS3* mRNA in yeast. Gene 114:35-41

Herzog P, Törrönen A, Harkki A, Kubicek CP (1992) Mechanism by which xylan and cellulose trigger the biosynthesis of endo-xylanase I by *Trichoderma reesei*. In: Xylans and xylanases (Eds. Visser J, Beltman G, Kusters-van Someren MA, Voragen AGJ) Prog Biotechnol 7, pp 289-293, Elsevier, Amsterdam

Hessing JGM, van Rotterdam C, Verbakel JMA, Roza M, Maat J, van Gorcom RFM, Van den Hondel CAMJJ (1994) Isolation and characterization of a 1,4-beta-endoxylanase gene of *A. awamori*. Curr Genet 26:228-232

Hintz WE, Kalsner I, Plawinski E, Guo Z, Lagosky PA (1995) Improved gene expression in *Aspergillus nidulans*. Can J Bot 73:876-884

Hrmova M, Biely P, M Vrsanska (1986) Specificity of cellulase and ß-xylanase induction in *Trichoderma reesei* QM 9414. Arch Microbiol 144:307-311

Hrmova M, Petrakova E, Biely P (1991) Induction of cellulose- and xylan-degrading enzyme systems in *Aspergillus niger* by homo- and heterodisaccharides composed of glucose and xylose. J of Gen Microbiol 137:541-547

Humphrey T, Proudfoot NJ (1988) A beginning to the biochemistry of polyadenylation. Trends in Genetics 4:243-245

Irniger S, Egli CM, Braus GH (1991) Different classes of polyadenylation sites in the yeast *Saccharomyces cerevisiae*. Mol Cell Biol 11:3060-3069

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

Jarai G, Van den Hombergh H, Buxton FP (1994b) Cloning and characterization of the *pepE* gene of *Aspergillus niger* encoding a new aspartic protease and regulation of *pepE* and *pepC*. Gene 145:171-178

Jeenes DJ, Mackenzie DA, Roberts IN, Archer DB (1991) Heterologous protein production by filamentous fungi. Biotechnol Genet Eng Rev 9:327-367

Jeenes DJ, Marczinke B, MacKenzie DA, Archer DB (1993) A truncated glucoamylase gene fusion for heterologous protein secretion from *Aspergillus niger*. FEMS-Microbiol Lett 107:267-271

Jeenes DJ, MacKenzie DA, Archer DB (1994) Transcriptional and post-transcriptional events affect the production of secreted hen egg white lysozyme by *Aspergillus niger*. Transgenic Research 3:297-303

Joseleau JP, Comtat J, Ruel K (1992) Chemical structure of xylans and their interaction in the plant cell walls. In: Xylans and xylanases (Eds. Visser J, Beltman G, Kusters-van Someren MA, Voragen AGJ) Prog Biotechnol 7, pp , Elsevier, Amsterdam

Kelley WL, Georgopoulos (1992) Chaperones and protein folding. Current Opinion in Cell Biology 4:984-991

Keränen S, Pentillä M (1995) Production of recombinant proteins in the filamentous fungus *Trichoderma reesei*. Current Opinion in Biotechnology 6:534-537

Kikuchi M, Ikehara M (1991) Conformational features of signal sequences and folding of secretory proteins in yeast. Trends Biotechnol 9:208-211

Kim PS, Bole D, Arvan P (1992) Transcient aggregation of nascent thyroglobulin in the endoplasmic reticulum: relationship to the molecular chaperone BiP. J Cell Biol 118:541-549

Kinsey JA (1989) A simple colony blot procedure for *Neurospora*. Fungal Genetics Newsletter 36:45-46

Knittler MR, Dirks S, Haas IG (1995) Molecular chaperones involved in protein degradation in the endoplasmic reticulum: Quantitative interaction of the heat shock cognate protein BiP with partially folded immunoglobulin light chains that are degraded in the endoplasmic reticulum. Proc Natl Acad Sci 92:1764-1768

Kolar M, Punt PJ, van den Hondel CAMJJ, Schwab H (1988) Transformation of *Penicillium chrysogenum* using dominant selection markers and expression of an *Escherichia coli lacZ* fusion gene. Gene 62:127-134

Korman DR, Bayliss FT, Barnett CC, Carmona CL, Kodama KH, Royer TJ, Thompson SA, Ward M, Wilson LJ, Berka RM (1990) Cloning, characterization, and expression of two α-amylase genes from *Aspergillus niger* var. *awamori*. Curr Genet 17:203-212

Kormelink FJM, Searle-Van Leeuwen MJF, Wood TM, Voragen AGJ (1993) Purification and characterization of three endo-(1,4)-beta-xylanases and one beta-xylosidase from *Aspergillus awamori*. J Biotechnol 27:249-265

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

Lloyd AT, Sharp PM (1991) Codon usage in *Aspergillus nidulans*. Mol Gen Genet 230:288-294

Manley JL (1988) Polyadenylation of mRNA precursors. Biochimica et Biophysica Acta 950:1-12

Mattern IE, Van Noort JM, Van den Berg P, Archer DB, Roberts IN, Van den Hondel CAMJJ (1992) Isolation and characterization of mutants of *Aspergillus niger* deficient in extracellular proteases. Mol Gen Genet 234:332-336

Metwally M, el Sayed M, Osman M, Hanegraaf PPF, Stouthamer AH, van Verseveld HW (1991) Bioenergetic consequences of glucoamylase production in carbon-limited chemostat cultures of *Aspergillus niger*. Antonie van Leeuwenhoek 59:35-43

Miller BL, Miller KY, Roberti KA, Timberlake WE (1987) Position-dependent and -independent mechanisms regulate cell-specific expression of the SpoC1 gene cluster of *Aspergillus nidulans*. Mol Cell Biol 7:427-434

Ng DTW, Watowich SS, Lamb RA (1992) Analysis in vivo of GRP78-BiP/ substrate interactions and their role in induction of the *GRP78-BiP* gene. Mol Biol Cell 3:143-155

Normington K, Kohno K, Kozutsumi Y, Gething M-J, Sambrook J (1989) *Saccharomyces cerevisiae* encodes an essential protein homologous in sequence and function to mammalian BiP. Cell 57:1223-1236

Nunberg JH, Meade JH, Cole G, Lawyer FC, McCabe P, Schweickart V, Tal R, Witmann VP, Flatgaard JE, Innis MA (1984) Molecular cloning and characterization of the glucoamylase gene of *Aspergillus awamori*. Mol Cell Biol 4:2306-2315

Nyyssönen E, Penttilä M, Harkki A, Saloheimo A, Knowles JKC, Keränen S (1993) Efficient production of

antibody fragments by the filamentous fungus *Trichoderma reesei*. Bio/Technology 11:591-595

Nyyssönen E, Keränen S (1995) Multiple roles of the cellulase CBHI in enhancing production of fusion antibodies by the filamentous fungus *Trichoderma reesei*. Curr Genet 28:71-79

Oakley BR, Rinehart BL, Oakley CE, Carmona C, Gray GL, May GS (1987) Cloning, mapping and molecular analysis of the *pyrG* (orotidine-5'-phosphate decarboxylase) gene of *Aspergillus nidulans*. Gene 61:385-399

Overbeeke N, Fellinger AJ, Toonen MY, van Wassenaar D, Verrips CT (1989) Cloning and nucleotide sequence of the α-galactosidase cDNA from *Cyamopsis tetragonoloba* (guar). Plant Molecular Biology 13:541-550

Overbeeke N, Termorshuizen GHM, Giuseppin MLF, Underwood DR, Verrips CT (1990) Secretion of the α-galactosidase from *Cyamopsis tetragonoloba* (guar) by *Bacillus subtilis*. Appl Environ Microbiol 56:1429-1434

Parker R, Jacobson A (1990) Translation and a 42-nucleotide segment within the coding region of the mRNA encoded by the MATα1 gene are involved in promoting rapid mRNA decay in yeast. Proc Natl Acad Sci USA 87:2780-2784

Parlati F, Dominguez M, Bergeron JJM, Thomas DY (1995) *Saccharomyces cerevisiae CNE1* encodes an endoplasmic reticulum (ER) membrane protein with sequence similarity to calnexin and calreticulin and functions as a constituent of the ER quality control apparatus. J Biol Chem 270:244-253

Pinaga F, Fernandez-Espinar MT, Valles S, Ramon D (1994) Xylanase production in *Aspergillus nidulans*: induction and carbon catabolite repression

Pontecorvo G, Roper JA, Hemmons LM, MacDonald KD, Bufton AWJ (1953) The genetics of *Aspergillus nidulans*. Advances in Genetics 5:141-239

Proudfoot NJ (1991) Poly(A) signals. Cell 64:671-674

Punt PJ, Dingemanse MA, Kuyvenhoven J, Soede RDM, Pouwels PH, Van den Hondel CAMJJ (1990) Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene, encoding glyceraldehyde-3-phosphate dehydrogenase. Gene 93:101-109

Punt PJ, Zegers ND, Busscher M, Pouwels PH, Van den Hondel CAMJJ (1991) Intracellular and extracellular production of proteins in *Aspergillus* under control of expression signals of the highly expressed *Aspergillus nidulans gpdA* gene. J of Biotech 17:19-34

Punt PJ, van den Hondel CAMJJ (1993) Transformation of filamentous fungi based on hygromycin B and phleomycin resistance markers. Methods in Enzymology 216:447-457

Punt PJ, Veldhuisen G, Van den Hondel CAMJJ (1994) Protein targeting and secretion in filamentous fungi. Antonie van Leeuwenhoek 65:211-216

Rapoport TA (1990) Protein transport across the ER membrane. TIBS 15:355-358

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

Raué HA (1994) Metabolic stability of mRNA in yeast - a potential target for modulating productivity. TIBTECH 12:444-449

Roberts IN, Oliver RP, Punt PJ, van den Hondel CAMJJ (1989) Expression of the *Escherichia coli* ß-glucuronidase gene in industrial and phytopathogenic filamentous fungi. Curr Genet 15:177-180

Roberts IN, Jeenes DJ, MacKenzie DA, Wilkinson AP, Sumner IG, Archer DB (1992) Heterologous gene expression in *A. niger*: a glucoamylase-porcine pancreatic prophospholipase A_2 fusion protein is secreted and processed to yield mature enzyme. Gene 122:155-161

Robinson AS, Hines V, Wittrup KD (1994) Protein disulphide isomerase overexpression increases secretion of

foreign proteins in *Saccharomyces cerevisiae*. Biotechnology 12:381-384

Romanos MA, Makoff AJ, Fairweather NF, Beesley KM, Slater DE, Rayment FB, Payne MM, Clare JJ (1991) Expression of tetanus toxin fragment C in yeast: gene synthesis is required to eliminate fortuitous polyadenylation sites in AT-rich DNA. Nucleic Acids Res 19:1461-1467

Romanos MA, Scorer CA, Clare JJ (1992) Foreign gene expression in yeast: a review. Yeast 8:423-488

Ross J (1995) mRNA stability in mammalian cells. Microbiol Rev 59:423-450

Rothman JH, Yamashiro CT, Kane PM, Stevens TH (1989) Protein targeting to the yeast vacuole. TIBS 14:347-350

Russo P, Li WZ, Hampsey DM, Zaret KS, Sherman F (1991) Distinct cis-acting signals enhance 3' endpoint formation of CYC1 mRNA in the yeast *Saccharomyces cerevisiae*. EMBO J 10:563-571

Russo P, Li WZ, Guo Z, Sherman F (1993) Signals that produce 3' termini in CYC1 mRNA of yeast *Saccharomyces cerevisiae*. Mol Cell Biol 13:7836-7849

Sachs AB (1993) Messenger RNA degradation in eukaryotes. Cell 74:413-421

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning - a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Santiago CT, Purvis IJ, Bettany AJ, Brown AJP (1986) The relationship between mRNA stability and length in *Saccharomyces cerevisiae*. Nucl Acids Res 14:8347-8360

Schmid FX, Mayr LM, Mucke M, Schonbrunner ER (1993) Prolyl isomerases: role in protein folding. Adv Prot Chem 44:25-66

Schuren FHJ, van der Lende TR, Wessels JGH (1993) Fruiting genes of *Schizophyllum commune* are transcriptionally regulated. Mycol Res 97:538-542

Scorer CA, Buckholz RG, Clare JJ, Romanos MA (1993) The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast *Pichia pastoris*. Gene 136:111-119

Shamu CE, Cox JS, Walter P (1994) The unfolded-protein-response pathway in yeast. Trends in Cell Biology 4:56-60

Sheets MD, Wickens M (1989) Two phases in the addition of a poly(A) tail. Genes Dev 3:1401-1412

Shibuya I, Tsuchiya K, Tamura G, Ishikawa T, Hara S (1992) Overproduction of an α-amylase/glucoamylase fusion protein in *Aspergillus oryzae* using a high expression vector. Biosci Biotech Biochem 56:1674-1675

Sierkstra LN, Nouwen NP, Verbakel JMA, Verrips CT (1992) Analysis of glucose repression in *Saccharomyces cerevisiae* by pulsing glucose to a galactose-limited continuous culture. Yeast 8:1077-1087

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

Skory CD, Horng JS, Pestka JJ, Linz JE (1990) Transformation of *Aspergillus parasiticus* with a homologous gene (*pyrG*) involved in pyrimidine biosynthesis. Appl Environ Microbiol 56:3315-3320

Steiner DF, Smeekens SP, Ohagi S, Chan SJ (1992) The new enzymology of precursor processing endoproteases. J Biol Chem 267:23435-23438

Suzuki K, Ichikawa K, Jigami Y (1989) Yeast mutants with enhanced ability to secrete human lysozyme: isolation and identification of a protease-deficient mutant. Mol Gen Genet 219:58-64

Tanaka A, Mita S, Ohta S, Kyozuka J, Shimamoto K, Nakamura K (1990) Enhancement of foreign gene expression by a dicot intron in rice but not in tobacco is correlated with an increased level of mRNA and an efficient splicing of the intron. Nucl Acids Res 18:6767-6770

Thomson JA (1993) Molecular biology of xylan degradation. FEMS Microbiology Reviews 104:65-82

Tilburn J, Scazzocchio C, Taylor GG, Zabicky-Zissman JH, Lockington RA, Davies RW (1983) Transformation by integration in *Aspergillus nidulans*. Gene 26:205-221

Tsuchiya K, Tada S, Gomi K, Kitamoto K, Kumagai C, Jigami Y, Tamura G (1992) High level expression of the synthetic human lysozyme gene in *Aspergillus oryzae*. Appl Microbiol Biotechnol 38:109-114

Tsuchiya K, Nagashima T, Yamamoto Y, Gomi K, Kitamoto K, Kumagai C (1994) High level secretion of calf chymosin using a glucoamylase-prochymosin fusion gene in *Aspergillus oryzae*. Biosci Biotechnol Biochem 58:895-899

Turnbull IF, Rand K, Willetts NS, Hynes MJ (1989) Expression of the *E. coli* enterotoxin subunit B gene in *A. nidulans* directed by the *amdS* promotor. Bio/Technology 7:169-174

Turner G (1994) Vectors for genetic manipulation. In: Martinelli SD, Kinghorn JR (Eds). *Aspergillus*: 50 years on, pp 641-665, Elsevier Science BV, Amsterdam, The Netherlands

Upshall A, Kumar AA, Bailey MC, Parker MD, Favreau MA, Lewison KP, Joeseph ML, Maraganore M, McKnight GL (1987) Secretion of active human tissue plasminogen from the fungus *Aspergillus nidulans*. Bio/Technology 5:1301-1304

Upshall A, Kumar AA, Kaushansky K, McKnight GL (1991) Molecular manipulation of and heterologous protein secretion from filamentous fungi. In: Molecular Industrial Mycology: systems and applications for filamentous fungi (Eds. S.A. Leong and R.M. Berka), pp 31-44. Marcel Dekker: New York

Van Gemeren I, Beijersbergen A, Musters W, Gouka RJ, van den Hondel CAMJJ, Verrips CT (1995) Heterologous expression of the *Fusarium solani pisi* cutinase gene in *Aspergillus awamori*. Appl Microbiol Biotechnol, in press

Van Gorcom RFM, Pouwels PH, Goosen T, Visser J, Van den Broek HWJ, Hamer JE, Timberlake WE, Van den Hondel CAMJJ (1985) Expression of an *Escherichia coli* ß-galactosidase fusion gene in *Aspergillus nidulans*. Gene 40:99-106

Van Gorcom RFM, Van den Hondel CAMJJ (1988) Expression analysis vectors for *Aspergillus niger*. Nucl Acids Res 16:9052 **Van Hartingsveldt W, Mattern IE, van Zeijl CMJ, Pouwels PH, van den Hondel CAMJJ** (1987) Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. Mol Gen Genet 206:71-75

Van Hartingsveldt W, Van Zeijl CMJ, Veenstra AE, Van den Berg JA, Pouwels PH, Van Gorcom RFM, Van den Hondel CAMJJ (1990) Heterologous gene expression in *Aspergillus*: analysis of chymosin production in single-copy transformants of *A. niger*. In: Heslot H, Bavies J, Florent J, Bobichon L, Durand G, Penasse L (eds) Proceedings of the 6th International Symposium on Genetics of Industrial Microorganisms. Société Francaise de Microbiologie, Strassbourg, France, pp 107-116

Van den Hombergh JPTW, van de Vondervoort PJI, van der Heijden NCBA, Visser J (1995) New protease mutants in *Aspergillus niger* result in strongly reducced in vitro degradation of target proteins; genetical and biochemical characterization of seven complementation groups. Curr Genet 28:299-308

Van den Hondel CAMJJ, Punt PJ (1991) Gene transfer systems and vector development for filamentous fungi. In: Applied Molecular Genetics of Fungi (Peberdy JF, Caten CE, Ogden JE, Bennett JW, Eds) pp 1-28, Cambridge University Press, Cambridge

Van den Hondel CAMJJ, Punt PJ, Van Gorcom RFM (1991) Heterologous gene expression in filamentous fungi.

In: More gene manipulations in fungi (eds Bennet JW and Lasure LL) pp 396-428. Academic Press, San Diego

Van Snick J, Cayphas S, Vink A, Uyttenhove C, Coulie PG, Rubira MR, Simpson RJ (1986) Purification and NH_2-terminal amino acid sequence of a T-cell-derived lymphokine with growth factor activity for B-cell hybridomas. Proc Natl Acad Sci 83:9679-9683

Verbakel JMA (1991) Heterologous gene expression in the yeast *Saccharomyces cerevisiae*. PhD thesis, State University of Utrecht, the Netherlands

Verdoes JC, Punt PJ, Stouthamer AH, van den Hondel CAMJJ (1994a) The effect of multiple copies of the upstream region on expression of the Aspergillus niger glucoamylase-encoding gene. Gene 145:179-187

Verdoes JC, Calil MR, Punt PJ, Debets F, Swart K, Stouthamer AH, van den Hondel CAMJJ (1994b) The complete karyotype of Aspergillus niger: The use of introduced electrophoretic mobility variation of chromosomes for gene assignment studies. Mol Gen Genet 244:75-80

Verdoes JC, Punt PJ, Van den Hondel CAMJJ (1995) Molecular genetic strain improvement for the overproduction of fungal proteins by filamentous fungi. Appl Microbiol Biotechnol 43:195-205

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

Wahle E (1991) A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. Cell 66:759-768

Wahle E, Keller W (1992) The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors. Annu Rev Biochem 61:419-440

Wang P, Ali S, Mason JC, Sims PFG, Broda P (1992) Xylanases from Streptomyces cyaneus. In: Xylans and xylanases (Eds. Visser J, Beltman G, Kusters-van Someren MA, Voragen AGJ) Prog Biotechnol 7, pp 225-234, Elsevier, Amsterdam

Ward M, Wilson LJ, Kodama KH, Rey MW, Berka RM (1990) Improved production of chymosin in *Aspergillus* by expression as a glucoamylase-chymosin fusion. Bio/Technology 8:435-440

Ward PP, May GS, Headon DR, Conneely OM (1992a) An inducible expression system for the production of human lactoferrin in *Aspergillus nidulans*. Gene 122:219-23

Ward PP, Lo JY, Duke M, May GS, Headon DR, Conneely OM (1992b) Production of biologically active recombinant human lactoferrin in *Aspergillus oryzae*. Bio/Technology 10:784-789

Ward PP, Piddington CS, Cunningham GA, Zhou X, Wyatt RD, Conneely OM (1995) A system for production of commercial quantities of human lactoferrin: a broad spectrum natural antibiotic. Bio/Technology 13:498-503

Wickens M (1990) How the messenger got its tail: addition of poly(A) in the nucleus. Trends Biochem Sci 15:277-281

Wingfield JM, Dickinson JR (1993) Increased activity of a model heterologous protein in *Saccharomyces cerevisiae* strains with reduced vacuolar proteinases. Appl Microbiol Biotechnol 39:211-215

Wittrup KD, Robinson AS (1994) Methods for increasing secretion of overexpressed proteins. WO 94/08012

YaDeau JT, Klein C, Blobel G (1991) Yeast signal peptidase contains a glycoprotein and the *sec11* gene product. Proc Natl Acad Sci USA 88:517-521

ABBREVIATIONS

A.	*Aspergillus*
AA	amino acid(s)
AGL	α-galactosidase
AGL_{syn}	synthetic α-galactosidase
aglA	α-galactosidase encoding gene
$aglA_{syn}$	synthetic α-galactosidase encoding gene
BiP	heavy chain binding protein
bp	base pairs
C.	*Cyamopsis*
cbh1	cellobiohydrolase encoding gene
E.	*Escherichia*
ER	endoplasmic reticulum
exlA	1,4-ß-endoxylanse A encoding gene
EXL	1,4-ß-endoxylanase
5-FOA	5-fluoro-orotic acid
glaA	glucoamylase encoding gene
GLA	glucoamylase
GLA_{G2}	glucoamylase G2 form (AA 1-514)
GUS	ß-glucuronidase
hIL6	human interleukin-6
Hm^R	gene conferring resistance to hygromycin
kb	kilobase(s) or 1000 bp
kD	kiloDalton
LPL	lipase
lplA	lipase encoding gene
MM	minimal medium
MW	molecular weight
nt	nucleotide(s)
PCR	polymerase chain reaction
pyrG	orotidine-5'-monophosphate decarboxylase encoding gene
T. (lanuginosus)	*Thermomyces (lanuginosus)*
T. (reesei)	*Trichoderma (reesei)*
tcr	transcription control region
uidA	ß-glucuronidase encoding gene
wt	wild type

CURRICULUM VITAE

Robin Gouka werd op 29 maart 1965 geboren te Schiedam. Na het behalen van het V.W.O. diploma aan de Scholengemeenschap Westland-Zuid te Vlaardingen in 1983 begon hij in hetzelfde jaar aan de studie Biologie aan de Rijksuniversiteit Leiden. De studiefase omvatte deelname aan het onderzoek binnen (i) de vakgroep Organismale Zoölogie, afdeling Neuro-anatomie (Prof. Dr. J.L. Dubbeldam) van de Rijksuniversiteit Leiden, en (ii) de sectie Recombinant DNA van het Medisch Biologisch Laboratorium TNO te Rijswijk (Prof. Dr. P.H.Pouwels en Dr. Ir. B.E. Valk). Het doctoraal examen werd op 31 januari 1989 behaald.

Van september 1988 tot juli 1989 was hij werkzaam aan het onderzoek naar fytase-overproducerende *Aspergillus* stammen binnen de sectie recombinant DNA van het Medisch Biologisch Laboratorium (Dr. C.A.M.J.J. van den Hondel en Drs. W. van Hartingsveldt). Van 1 augustus 1989 tot 1 juni 1991 werkte hij in dienst van TNO aan de ontwikkeling van transformatie systemen voor de schimmel *Penicillium*, als contract research voor Gist-brocades.

Van 1 juni 1991 tot 15 november 1995 was hij verbonden aan de Dr. Daniel den Hoed Kliniek. Het onderzoek werd uitgevoerd binnen de afdeling Moleculaire Genetica en Gentechnologie bij het Medisch Biologisch Laboratorium TNO (nu TNO Voeding) te Rijswijk, onder leiding van Dr. C.A.M.J.J. van den Hondel. De resultaten van dit onderzoek zijn beschreven in dit proefschrift.

Vanaf 1 januari 1996 is hij als toegevoegd onderzoeker van de Rijksuniversiteit Utrecht (vakgroep Moleculaire Celbiologie, projectgroep Toegepaste Moleculaire Biologie) werkzaam bij het Unilever Research Laboratorium te Vlaardingen (Prof. Dr. C.T. Verrips).

LIST OF PUBLICATIONS

Gouka RJ, Van Hartingsveldt W, Bovenberg RAL, van den Hondel CAMJJ, van Gorcom RFM (1991) Cloning of the nitrate - nitrite reductase gene cluster of *Penicillium chrysogenum* and use of the *niaD* gene as a homologous selection marker. J of Biotechnology 20:189-200

Van Hartingsveldt W, Van Gorcom R,FM, Gouka RJ, Bovenberg RAL (1992) Transformation selection marker system adapted for *Penicillium*. EPA 0487118A1

Van Hartingsveldt W, Van Zeijl CMJ, Harteveld M, Gouka RJ, Suykerbuyk MEG, Luiten RGM, Van Paridon PA, Selten GCM, Veenstra AE, Van Gorcom RFM, Van den Hondel CAMJJ (1993) Cloning, characterization and overexpression of the phytase-encoding gene (*phyA*) of *Aspergillus niger*. Gene 127:87-94

Gouka RJ, Van Hartingsveldt W, Bovenberg RAL, Van Zeijl CMJ, Van den Hondel CAMJJ, Van Gorcom RFM (1993) Development of a new transformant selection system for *Penicillium chrysogenum*: isolation and characterization of the *P. chrysogenum* acetyl-coenzyme A synthetase gene (*facA*) and its use as a homologous selection marker. Appl Microbiol Biotechnol 38:514-519

Gouka RJ, Van den Hondel CAMJJ, Musters W, Stam H, Verbakel JMA (1993) Process for producing/secreting a protein by a transformed mould using expression/secretion regulating regions derived from an *Aspergillus* endoxylanase II gene. WO-A-93/12237

Gouka RJ, Hessing JGM, Stam H, Musters W, van den Hondel CAMJJ (1995) A novel strategy for the isolation of defined *pyrG* mutants and the development of a site-specific integration system for *Aspergillus awamori*. Curr Genet 27:536-540

Gouka RJ, Hessing JGM, Punt PJ, Stam H, Musters W, Van den Hondel CAMJJ (1996) An expression system based on the promotor region of the 1,4-ß-endoxylanase A gene of *Aspergillus awamori*. Appl Microbiol Biotechnol, in press

Gouka RJ, Punt PJ, Hessing JGM, Van den Hondel CAMJJ (1996) Analysis of heterologous protein production in defined recombinant *Aspergillus awamori* strains: Comparison of expression levels of genes from fungal and non-fungal origin, Appl Env Microbiol, in press

Gouka RJ, Punt PJ, Van den Hondel CAMJJ (1996) Glucoamylase gene fusions alleviate limitations for protein production in *Aspergillus awamori* at the transcriptional and (post-)translational level, submitted

Gouka RJ, Fellinger A, Stam H, Punt PJ, Musters W, van den Hondel CAMJJ (1996) Kinetics of mRNA and protein synthesis of genes controlled by the 1,4-ß-endoxylanase A promoter in controlled fermentations of *Aspergillus awamori*, submitted

Van Gemeren I, Beijersbergen A, Musters W, Gouka RJ, van den Hondel CAMJJ, Verrips CT (1996) Heterologous expression of the *Fusarium solani pisi* cutinase gene in *Aspergillus awamori*. Appl Microbiol Biotechnol, in press

DANKWOORD

Tot slot wil ik iedereen bedanken die een bijdrage aan dit proefschrift heeft geleverd. Het schrijven van een dankwoord lijkt een peulenschil na het voltooien van een proefschrift, maar blijkt toch nog veel tijd te kosten daar ik weet dat het dankwoord ook het meest gelezen gedeelte is.

De volgende mensen wil ik graag bedanken:

In de eerste plaats mijn promotor Prof. Dr. Ir. Verrips; Beste Theo, ondanks dat je niet vanaf het begin bij mijn onderzoek betrokken bent geweest, ben ik blij dat je mijn promotor wilde zijn. Ik wil je bedanken voor je enthousiasme en interesse die je tijdens mijn onderzoek hebt getoond. Tevens wil ik je bedanken voor de mogelijkheid die je me hebt geboden om bij Unilever mijn kennis op het gebied van genexpressie en eiwitproduktie in schimmels te verbreden.

Mijn co-promotor Dr. van den Hondel; Beste Kees, jouw ideeën hebben een belangrijke rol gespeeld bij het tot stand komen van dit proefschrift en bij mijn ontwikkeling als zelfstandig onderzoeker in het algemeen. Daarnaast was je ook altijd bereid om over zaken te spreken die niet direct met mijn onderzoek te maken hadden. Bedankt voor je enthousiasme en belangstelling binnen en buiten mijn onderzoek en voor je inzet de resultaten van het onderzoek tot publikaties te verwerken.

Dr. Punt; Beste Peter, aan jou ben ik speciale dank verschuldigd bij het schrijven van alle publikaties. Niet alleen voor de snelheid waarmee je ze hebt doorgewerkt maar ook voor al je waardevolle suggesties, die een belangrijke bijdrage vormden voor het tot stand komen van mijn publikaties en mijn proefschrift in het geheel.

In het speciaal Cora van Zeijl. Beste Cora, van de acht jaar die ik bij TNO heb gewerkt hebben wij 7 jaar op dezelfde kamer(s) doorgebracht. Eerst als Gb-ers, samen met Wim, daarna werkend voor verschillende projecten. Ik heb deze tijd altijd als zeer gezellig ervaren. Ik ben blij dat je bereid bent mijn paranimf te zijn.

Hanny Hessing, voor je inbreng in dit onderzoek en voor de autotochtjes naar Vlaardingen. Ik hoop wel dat je voortaan zorgt dat er voldoende benzine in je tank zit. Daniëlle Remmerswaal, voor je assistentie in de laatste fase van mijn onderzoek.

Alle (ex-)medewerkers van de afdeling MGG in het bijzonder en verder alle mensen van het ex-MBL van TNO voor de fijne en gezellige tijd die ik tijdens mijn verblijf heb gehad. Ik

wens jullie veel plezier in Zeist en ik hoop dat ze bij TNO weten dat voeding nog steeds een primaire levensbehoefte is.

De mensen van de ondersteunende diensten, die vaak ongenoemd blijven maar toch een belangrijke bijdrage leveren: Michel, Jeanne en Herbert. Ook Jan Jore ben ik zeer erkentelijk voor de synthese van een groot aantal oligonucleotiden. Zonder anderen te kort te doen, moet ik toch nog iemand noemen. Beste Hans, jarenlang heb je gezeurd om bij mij in de acknowledgements te komen, hetgeen je niet is gelukt, ondanks verschillende sabotage pogingen wanneer ik ingelogt was en niet achter de PC zat. Maar, zoals je misschien nog zult merken moet een publikatie enige vorm van niveau hebben. Met name de laatste jaren hebben we elkaars werk kritisch gevolgd (en benaderd), hetgeen misschien op sommige mensen anders overkwam. Om enig leed te verzachten heb ik je daarom genoemd in mijn dankwoord.

Ook wens ik mijn lotgenoten Hans, Anders, Christien en Robert (in chronologische volgorde ?) veel succes met het afronden van hun proefschrift.

Buiten TNO, vooral de mensen van de "cumpany" die verantwoordelijk zijn voor de uitvoering en de know-how van de schimmel fermentaties, hetgeen niet de meest makkelijke experimenten waren zoals ik zelf een keer heb mogen ondervinden: Arthur Fellinger, Rob Muijsenberg, Wouter Musters, Hein Stam en Arjan van den Wijngaard. Tevens wil ik jullie bedanken voor de plezierige samenwerking.

Mijn ouders, voor de mogelijkheid die ze mij hebben gegeven om te gaan studeren, zonder daar ooit enige druk achter te zetten.

Jackie, voor je steun tijdens mijn promotie onderzoek, waarbij je af en toe zelfs hebt geholpen, voor je geduld wanneer ik weer eens 3 uur bezig was in plaats van het beloofde half uur en voor je bijdrage aan de goede uitkomst van ons laatste "experiment". De geboorte van Linda realiseert je dat het hebben van een kind iets speciaals is.

Lieve Linda, in je nog zeer jonge bestaan heb je, onbewust, een bijdrage kunnen leveren aan dit proefschrift door me wakker te houden in tijden dat ik werd overvallen door slaap.

Tenslotte hoop ik dat ik iedereen persoonlijk kan bedanken op mijn promotiefeest.