FUNCTIONAL ELEMENTS IN THE PROMOTER REGION OF THE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE GENE OF *ASPERGILLUS NIDULANS*

Gedrukt door: Drukkerij Haveka B.V., Alblasserdam

Illustratie voorpagina: *Aspergillus nidulans*
(fotografie; M. J. M. Boermans)

FUNCTIONAL ELEMENTS IN THE PROMOTER REGION OF THE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE GENE OF *ASPERGILLUS NIDULANS*

FUNCTIONELE ELEMENTEN IN HET PROMOTOR GEBIED VAN HET GLYCERALDEHYDE-3-FOSFAAT DEHYDROGENASE GEN VAN *ASPERGILLUS NIDULANS*

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam,
op gezag van de Rector Magnificus
Prof. Dr. P. W. M. de Meijer,
hoogleraar aan de faculteit der letteren,
in het openbaar te verdedigen in de Aula der Universiteit
(Oude Lutherse Kerk, ingang Singel 411, hoek Spui),
op dinsdag 23 juni 1992 te 12.00 uur

door Peter Jan Punt

geboren te Rotterdam

(verbonden aan de Faculteit Biologie
van de Universiteit van Amsterdam en
het Medisch Biologisch Laboratorium
van de Organisatie voor Toegepast
Natuurwetenschappelijk Onderzoek (TNO))

Copromotor: Dr. C. A. M. J. J. van den Hondel
(verbonden aan het Medisch Biologisch
Laboratorium van de Organisatie voor
Toegepast Natuurwetenschappelijk
Onderzoek (TNO))

Het onderzoek beschreven in dit proefschrift is uitgevoerd in de afdeling Biochemie, sectie Moleculaire Genetica en Gen-technologie, van het Medisch Biologisch Laboratorium van de Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (TNO) onder leiding van Dr. C. A. M. J. J. van den Hondel en Prof. Dr. P. H. Pouwels.

Uitgave van dit proefschrift is financieel ondersteund door een bijdrage van het Medisch Biologisch Laboratorium.

Uw woord is een lamp voor mijn voet en een licht op mijn pad

Psalm 119: 105

Voor mijn vader en moeder,
voor Eline

CONTENTS

OUTLINE

ıtline of this thesis

he study of gene organisation and regulation of gene expression is one of the ıtral topics in molecular biological research. Many research groups in the field nolecular biology, working with either prokaryotic or eukaryotic organisms, have used their attention on this kind of research for purely scientific but also more plied" reasons. For detailed analysis of regulation of gene expression *in vivo* it mportant that (modified) genetic information can be introduced and expressed in organism of interest. Therefore, in molecular biological research much effort was : in developing gene-transfer systems. Having these systems available, attention ı be directed towards an understanding of the molecular mechanisms of complex logical processes, such as the regulation of eukaryotic gene expression. To carry : research on biological processes also the availability of classical genetic proaches is of considerable importance. Whereas this kind of approach is not sible in most eukaryotic organisms, excellent classical genetic strategies have en available for several filamentous fungi for several decades. The "one gene- e enzyme" hypothesis, which may be seen as the basis of research on gene ıction, was founded on studies carried out by Beadle and Tatum with the bread ıgus *Neurospora* fifty years ago. Furthermore, very good examples of detailed netic maps for eukaryotic organisms have been established in *Aspergillus nidulans* d *N. crassa*. Classical studies on genetic recombination have also been carried t in fungi, in particular *Sordaria macrospora* and *Ascobolus immersus*.

Filamentous fungi, in particular *A. nidulans* and *N. crassa*, have a number of ditional properties which make them very attractive for molecular biological studies eukaryotic gene organisation and regulation of gene expression; (1) they show llular differentiation of the vegetative mycelium and a relatively complex life cycle, ıich distinguishes them from unicellular yeasts like, *Saccharomyces cerevisiae*; (2) ey have a broad metabolic versatility and many genetic and biochemical data are ailable for various metabolic pathways; (3) finally, they can be cultivated in simple owth media, which makes working with them in the laboratory as easy as with *cherichia coli* or *S. cerevisiae*. With the development of versatile gene-transfer chniques for filamentous fungi (first reported for *N. crassa* in the late seventies) olecular genetic studies concerning the regulation of gene expression became asible in these organisms.

The work described in this thesis is aimed at the understanding of the process transcriptional regulation of gene expression in filamentous fungi. At the time th work was started only *A. nidulans* and *N. crassa* were amenable to this kind studies. In this thesis the attention is focused on expression of an efficien expressed gene of *A. nidulans*, namely *gpdA*, encoding glyceraldehyde-3-phospha dehydrogenase. The use of an efficiently expressed gene was expected to facilita the development of the methodology for *in vivo* analysis of gene expression.

In Chapter 1 of this thesis an overview is given of the literature concerni research on the structure and organisation of fungal transcription control sequence Furthermore, the results obtained with this research are discussed.

In Chapter 2 the isolation and characterisation of the *A. nidulans* gene, *gpdA* described. The complete nucleotide sequence of the coding region of the ge together with that of the 5'- and 3'-flanking sequences is presented. Based sequence comparison with the upstream sequences of a few other fungal gen several putative 5' transcription control sequences are indicated.

The ability of the 5' upstream region of the *gpdA* gene to drive gene expressi is described in Chapter 3. Fusion of the *gpdA* upstream region to the coding regi of the *E. coli* hygromycin B phosphotransferase gene (*hph*) resulted in a plasm pAN7-1, which, when introduced into various fungal host strains, gives rise expression of the *hph* gene. Expression of this gene can be monitored easily as t *hph* gene product confers resistance to the antibiotic hygromycin B in fungal strair

In Chapter 4 the *in vivo* analysis of the *gpdA* upstream region is presente Various deletion derivatives of the upstream region, obtained by either random site directed mutagenesis, are analyzed by means of the *lacZ*-fusion strategy their capacity to drive gene expression in the fungal host. Based on the resu obtained with this analysis several elements are indicated which are involved transcription initiation and transcription activation.

One of the elements identified by the above mentioned *in vivo* strategy, w subjected to more detailed analysis. In Chapter 5 the results of this analysis a presented. It is shown that this element, the so-called *gpd* box, is capable increasing gene expression out of its normal context. Introduction of the *gpd* b into the upstream region of the highly regulated acetamidase-encoding gene (*amd* of *A. nidulans*, which is involved in carbon and nitrogen metabolism, results in marked increase of the level of gene expression driven by this upstream regic Furthermore, evidence is presented which indicates that the *gpd* box does n interfere with normal transcriptional regulation of the *amdS* gene. From the da

sented in this chapter it is suggested that the *gpd* box comprises more than one ctional element. This result would indicate that its function is more complex than ıy other elements identified in fungal transcription control regions identified sofar.

Molecular Signals in Plant-Microbe
Communications, D. P. S. Verma. ed., CRC Pre
Boca Raton, Florida, 1992, pp. 29-48

CHAPTER 1

ıalysis of transcription control sequences of fungal genes

ter J. Punt and Cees A. M. J. J. van den Hondel

TABLE OF CONTENTS

Filamentous fungi, in particular the genetically and biochemically well-characterize species *Aspergillus nidulans* and *Neurospora crassa*, have a number of propertie which make them very attractive for molecular biological studies of eukaryotic gen organization and regulation of gene expression. First, filamentous fungi (especial many plant pathogenic fungi) show distinct cellular differentiation of the vegetativ mycelium and a complex life cycle,[1] which makes them clearly distinct fro taxonomically related unicellular yeasts, such as *Saccharomyces cerevisiae* an *Schizosaccharomyces pombe*. Second, fungi generally have enormous metabol versatility and a wealth of genetic and biochemical data is available for mar biosynthetic and catabolic pathways in various fungal species.[2-4] Finally, mar species can be cultivated in simple growth media, which makes them easy to wo with in the laboratory .

The development of genetic transformation techniques was a major breakthroug for molecular biological research in filamentous fungi. The first report about genet transformation of a filamentous fungus dates from the late 1970s, when Case et al transformed a *N. crassa qa-2* mutant with a vector containing the cloned *qa-2* ger encoding dehydroquinase. Since then a large number of reports have bee published describing genetic transformation of more than 50 different fungal specie using different auxotrophic and dominant selection markers.[6] The availability transformation techniques made possible the development of various genet manipulation techniques required for further molecular biological research on th regulation of gene expression.

In the last few years many fungal genes have been isolated and characterize based on these techniques. Extensive data have been accumulated about th primary structure of these genes including 5'- and 3'-flanking sequences. In th chapter we will focus on an analysis of 5'-flanking sequences and their role regulation of gene expression, in particular, transcription control. Obviously, 3 flanking sequences can also be involved in the regulation of gene expression, e. by determining the site of transcription termination and polyadenylation. Howeve no experimental data concerning the role of 3'- flanking sequences of fungal gene in gene expression are yet available.

The first step in identifying sequences involved in transcription control is general the comparison of 5'-flanking sequences, with 5'-flanking sequences of genes of *cerevisiae* and higher eukaryotes.

For *S. cerevisiae* and higher eukaryotes, a considerable amount of data regarding the organization of expression signals already exists.[7,8] A generalized scheme of yeast transcription control regions is presented in Figure 1. Three types of sequences are indicated; 1) upstream activating/repressing sequences (UAS/URS); 2) TATA sequences; and 3) transcription initiation sequences.

UAS/URS - These are short DNA sequences of 10 to 30 nt located at various distances upstream of the mRNA start site. These sequences are involved in the regulation of gene expression, in most cases as (putative) target sites for transacting regulatory proteins.

TATA sequences - These are located at 40 to 120 nt upstream of the mRNA start site. They are involved in determining the efficiency of transcription initiation in many yeast genes. Furthermore, in several genes the TATA sequence is also involved in determining correct transcription initiation.

Transcription initiation sequences - Many different sequences can function as the initiation site. In the case of efficiently expressed yeast genes, initiation predominantly occurs at PyAAG sequences. In general, initiation sequences are not involved in determining transcription efficiency.

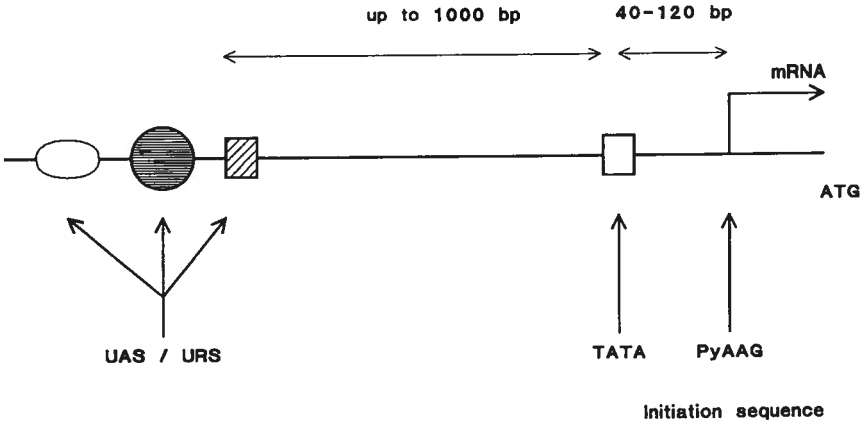


Figure 1. Scheme of yeast transcription control sequences

A similar organization is also found for transcription control regions in high eukaryotes, although, in this case, the TATA sequence is located at a more or le fixed position about 30 nt upstream of the mRNA start site. Furthermore, additional, conserved sequence, the CAAT-box, located at 70 to 90 nt upstream the mRNA start site is observed in a number of genes.

Many fungal genes lack either TATA-, CAAT-, or PyAAG-like sequences, indicating that filamentous fungi differ in the organization of their transcription contr sequences from *S. cerevisiae* (and related yeasts) and higher eukaryotes. A furth indication that filamentous fungi and yeasts, although taxonomically related, differ their transcription control sequences, is that successful use of yeast transcripti control sequences in filamentous fungi has never been reported. Moreover, in or a few cases were fungal transcription control sequences functional in *S. cerevisiae*

An overview about approaches for analysis of fungal transcription contr sequences will be provided in this chapter. The results obtained with the approaches will be discussed.

II. ANALYSIS OF TRANSCRIPTION CONTROL SEQUENCES

(a) Sequence analysis

The most simple approach to analyze (cloned) transcription control regions is comparison of the 5'-flanking DNA sequences of different genes. For examp comparison of the 5'-flanking sequences of coregulated genes may be used identify regions of similar sequence. Such regions could constitute binding sites f either general or specific regulatory proteins. The usefulness of this approach w demonstrated by Gwynne et al.,[12] in their study of the genes involved in ethar metabolism in *A. nidulans*. From genetical and biochemical data it was alrea known that the *alcR* gene product was involved in the regulation of expression both *alcA* (encoding alcohol dehydrogenase) and *aldA* (encoding aldehy dehydrogenase).[13] Comparison of the 5'-flanking sequences of both genes result in the identification of putative binding sites for the regulatory protein encoded the *alcR* gene.[12] Also, in the study of the gene cluster involved in quinic ac metabolism in *N. crassa*[14,15] and *A. nidulans*,[16] sequence comparison of 5'-flankir sequences from coregulated genes led to the identification of putative regulatc sequences. Recently, a number of elements of similar sequence has also be identified in the nitrate gene cluster of *A. nidulans* at both sides of the interger

igion between the coregulated *niiA* and *niaD* genes, encoding nitrite and nitrate iductase, respectively.[17]

Sequence comparison of the 5'-flanking regions of genes encoding the same rotein, in related fungal species, can also lead to the identification of putative anscription control sequences, as can be illustrated for the *A. niger* and *A. nidulans* pdA and *oliC* genes, encoding glyceraldehyde-3-phosphate dehydrogenase and an TPase subunit, respectively. The 5'-flanking sequences of both *gpdA* genes show n overall similarity of about 70%. However, a 50-nt region of much higher similarity 85%) was identified about 250 nt upstream of the transcription initiation site (Figure A).[18,19] In the 5'-flanking sequences of both *oliC* genes, which show about 50% verall similarity, a 30-nt region of about 85% similarity was identified about 60 nt pstream of the (major) transcription initiation site (Figure 2B).[20] Indications that iese regions are functional transcription control sequences have been obtained by *vivo* analysis (see section **(c)**2. below).

A.

```
TCCAAATATCGTGCCTCTCCTGCTTTGCCCGGTGTATGAAACCGGAAAGG -235 A. nidulans gpdA
************  ***********************************
TCCAAATATCGTGAGTCTCCTGCTTTGCCCGGTGTATGAAACCGGAAAGG -283 A. niger gpdA
```

B.

```
GGTGAAAAAAGGGCGAAAATTAAGCGGGAGA  -60     A. nidulans oliC
**  ******  *********  *********
GGGCAAAAAACCGCGAAAATTTAGCGGGAGA  -50/60  A. niger oliC
```

Figure 2. Elements of similar sequence in the upstream region of the (A) *gpdA* and (B) *oliC* genes of *A. nidulans* and *A. niger.* The distance from the transcription initiation site is given (in nt). In the case of the *A. niger oliC* gene, the exact position of the transcription initiation site is unknown.

c) Protein DNA-binding analysis

In general, regulation of gene expression at the transcriptional level is based on ie action of regulatory proteins.[7,8,21-24] The most simple representation for this type f regulation is binding of a regulatory protein to sequences in the 5'-flanking egions of the gene of interest and subsequent interaction of the bound protein with ie transcription initiation complex. Fungal DNA sequences which bind to regulatory

TABLE I.
Cloned regulatory genes of filamentous fungi

Regulatory gene	Regulatory function	Functional characterization			References
		DNA binding motif[1]	Functional analysis[2] of DNA binding motif	*In vitro* DNA binding[3] (target sequence)	
Neurospora crassa					
cpc-1	amino acid catabolism	L	(+)+	+ (ATGACTCAT)	26-29
cys-3	sulfur metabolism	L	(+)	+ ($ATGN_{1-4}CAT,TTC^{T}/_{G}C^{T}/_{G}$)	30,31
nit-2	nitrogen catabolism	Z	+	+ (repeated TATCTA)*	32-34
nit-4	nitrate assimilation	Z	NT	NT	35,36
nmr-1	nitrogen catabolism	?	NT	NT	37-39
nuc-1	phosphorus metabolism	HLH	NT	NT	40
qa-1F	quinate catabolism	Z	(+)	+ ($GG^{A}/_{G}TAA^{A}/_{G}{}^{C}/_{T}{}^{A}/_{G}{}^{C}/_{T}TTATCC$)	41,42
qa-1S	quinate catabolism	?	NT	NT	42

[1], Based on sequence comparison, various DNA binding motifs were indicated; Z, zinc finger (different types); L, leucin zipper (including basic region); HLH, helix-lo helix motif; ?, no homology to either of the known motifs; _, no sequence data available

[2], Functionality of the putative DNA-binding motif was tested by functional (*in vivo*) analysis of the products of mutant genes, which were constructed by deletion- and directed mutagenesis, indicated by +. In some cases, functionality of the DNA-binding motif was indicated by sequence analysis of various mutant alleles, indicated by (NT, not tested.

[3], *In vitro* DNA binding was analysed by bandshift analysis. If a target sequence is indicated, footprint analysis was also carried out. *, in this case footprint analysis carried out with only a part of the Nit2 protein. NT, not tested.

: I (continued)

ıtory gene	Regulatory function	Functional characterization			References
		DNA binding motif[1]	Functional analysis[2] of DNA binding motif	In vitro DNA binding[3] (target sequence)	
ɟillus nidulans					
	conidiospore development	L	NT	NT	1,43
	ethanol metabolism	Z	NT	+	44,45
	amide, ω-aminoacid and lactam catabolism	Z	+	+	46-48
	nitrogen catabolism	Z	(+)+	NT	49-51
	conidiospore development	Z	+	NT	43,52,53
	carbon catabolism	–	–	–	54,55
	amide and acetate metabolism	–	–	–	56
	nitrate assimilation	Z	NT	NT	45
	proline catabolism	Z	NT	NT	45,57
	quinate catabolism	Z	NT	NT	58
	quinate catabolism	?	NT	NT	59
	purine catabolism	Z	NT	NT	45
	conidiospore development	–	–	–	1,43

ɜe legend on p. 18.

In most cases, the strategy in identifying the sequences involved in protein binding comprises the following steps: (1) detection of specific binding of (regulatory) protein(s) to particular transcription control sequences by so-called bandshift experiments. In these experiments, DNA fragments containing putative control sequences are incubated with (nuclear) protein preparations or (partially) purified regulatory protein; (2) identification of the specific sequences involved in DNA binding by detailed footprint analyses, such as DNAse I and methylation protection assays. In most cases, the latter techniques require the availability of purified regulatory protein.

Extensive biochemical and genetical studies about the regulation of expression of the genes of several metabolic and differentiation pathways in filamentous fungi have resulted in the isolation and characterization of a number of regulatory mutants.[2-4] Based on these mutants, several of the corresponding regulatory genes have been cloned and the products of these genes have been characterized (Table I).[25] The availability of these genes permits isolation of sufficiently large amounts of the regulatory protein by overexpression of the cloned gene to identify DNA-binding sequences in the transcription control sequences of the (cloned) genes of the various pathways. As summarized in Table I, in most cases, binding of the regulatory protein to transcription control sequences is suggested by the presence of DNA binding motifs (leucine zipper, zinc finger, helix-loop-helix) in the regulatory proteins.[22] For some regulatory proteins, the role of these motifs in regulation of gene expression was tested by analysis of the activity of mutant proteins from which these DNA-binding motifs have been deleted. In all cases tested, these mutant proteins were no longer functional (Table I; column 3; indicated with +).[29,34,47,49,53] In some cases, the functionality of the DNA-binding motif was indicated by sequence alterations in the DNA-binding motif of mutant alleles of the regulatory gene [Table I; column 3; indicated with (+)].[27,30,31,41,42,49,50,51] Currently, only for a few regulatory proteins specific binding to transcription control regions of genes regulated by these proteins was demonstrated by bandshift analysis (Table I; column 4).[28-31,33,41,42,45,48] In four cases (all *N. crassa*), the DNA sequences involved in binding were further identified by footprint analysis (Table I; column 4).[29,31,33,41,42] Surprisingly, the "leucine zipper" regulatory proteins encoded by *cpc-1* and *cys-3*, which are involved in regulation of two distinct metabolic pathways, bind to a similar target sequence (Table I; ATGactCAT).[29,31]

Even if the regulatory DNA-binding protein is not available in a purified form or if

CT-TGATTTG -170/180 *niaD*
* ***** *
CGCTGATTCG -168 *niiA*
** **** *
CG-TGATCGG -225 *crnA*

Figure 3. Protein DNA-binding analysis of a sequence element in the upstream region of the *A. nidulans niaD* gene. **(A).** Elements of similar sequence in the upstream region of all three genes of the nitrate genecluster. The distance from the transcription initiation site is given (in nt). In the case of the *niaD* gene the exact position of the transcription initiation site is unknown. **(B).** Bandshift analysis with a double-stranded oligonucleotide containing the *niaD* element and nuclear extracts from *A. nidulans* cultures grown in medium containing NO_3^- (induced [I]), glutamate or proline (noninduced [NI]) or NH_4^+ (repressed [R]) as sole nitrogen source.

identity is unknown, protein DNA-binding analysis may lead to identification of anscription control sequences. Frederick et al.[60] have used bandshift experiments identify protein binding sequences in the 5'-flanking region of the *N. crassa am* ene, encoding glutamate dehydrogenase. Indications that these binding sequences e also functional *in vivo* was obtained by *in vivo* analysis (see below, section)2.).[60,73] Binding of a protein factor to sequences of the 5'-flanking region of the utinase gene from *Fusarium solani* (*cutA*) was also observed.[61] Preliminary results om bandshift analyses showed that several sequence elements present at both des of the intergenic region of the coregulated *A. nidulans niiA* and *niaD* genes e bound by protein(s).[17] With one of these elements, a copy of which is present about 170 nt upstream of the transcription initiation site of each gene (Figure 3A), nding of one or more proteins of a nuclear extract from nitrate-induced cultures

uninduced (or ammonium repressed) cultures (Figure 3B). This result suggests th this sequence is involved in the induction of the *niiA* and *niaD* genes by nitrate. similar sequence element is also present at 225 nt upstream of the *crnA* ger encoding nitrate permease (Figure 3A).[62] This gene, which is clustered with *niiA* ar *niaD*, is also induced by nitrate, further indicating a role for the identified sequen element in nitrate induction.

It is important to note that bandshift, DNAse I protection, and related types assays do not give information about the *in vivo* function of the identified sequence Therefore, additional *in vivo* analysis (see section **(c)** below) is necessary to obta conclusive data about the sequences which are involved in regulation of ge expression at transcriptional level (and the mechanisms by which regulatory protei work).

(c) *In vivo* analysis

1. Titration analysis

In general, introduction of multiple copies of a gene into a host strain leads to a increase in the amount of corresponding protein. However, if the expression of th gene is regulated through the action of regulatory protein(s) and the genes encodi these proteins are not concomitantly amplified, shortage of the regulatory protein(could result. This so-called titration of regulatory proteins could result in a no linear relation between the level of gene expression and the copy number. Titrati of a positively acting regulatory protein would lead to a decrease in the amount gene product per gene copy. In the case of a negatively acting regulatory prote an increase of the amount of gene product per gene copy is to be expecte Evidence for titration of regulatory proteins was reported in expression studies of t *A. nidulans qutE*,[63,64] *alcA*,[65] and *amdS*[66] genes. In the case of the *qutE* ge (encoding dehydroquinase), transformants containing multiple copies of this ge showed wildtype (wt) expression levels, suggesting (very tight) titration.[63] Expressi of the interferon α2 gene driven by the transcription control region of the *alcA* ge and, thus, controlled by the *alcR* gene product, resulted in production of interfer α2 in *A. nidulans*. However, in multicopy transformants a decrease in *alcA* (and *ald* which is also controlled by the *alcR* gene product) expression, compared untransformed *A. nidulans* was observed, indicating that the amount of regulato protein (*alcR* gene product) was limiting through titration.[65]

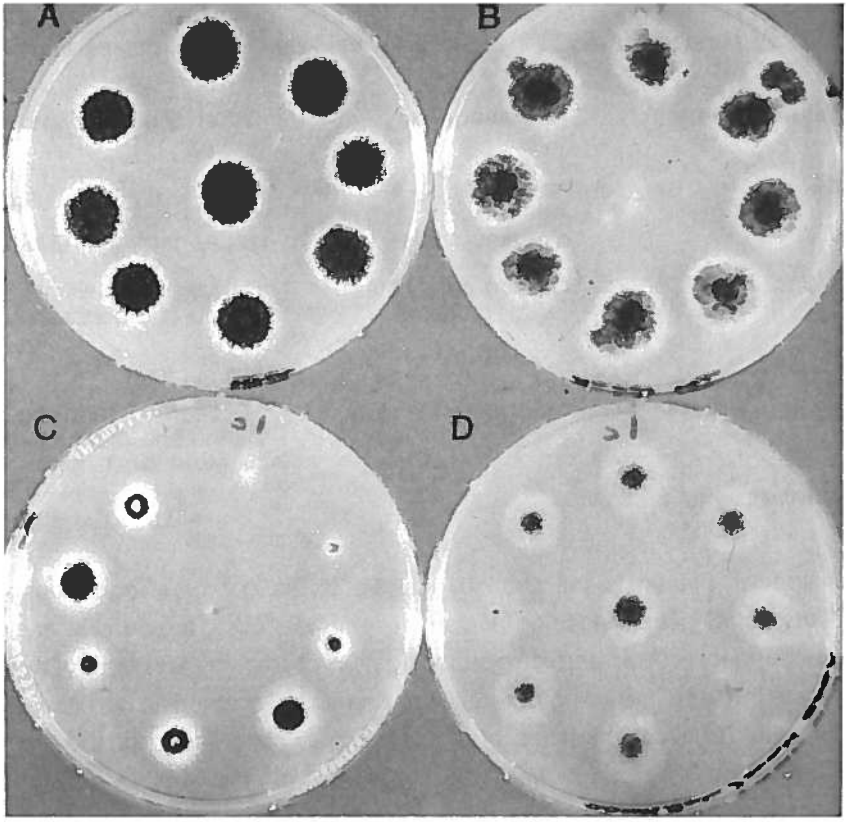
TABLE II.
Regulatory genes involved in expression of the *Aspergillus nidulans amdS* gene

Regulatory gene	Regulation by	Other regulated genes/pathways
amdA	Acetate	*aciA*
amdR	ω-Amino acids	*gabA, gatA, lamA, lamB*
creA/B/C	Carbon catabolites (e. g., glucose)	Carbon catabolism (e. g., ethanol and acetate metabolism)
facB	Acetate	*acuD, acuE, facA*
areA	Nitrogen metabolites (NH_4^+, glutamine)	Nitrogen catabolism (e. g., purine and proline catabolism, nitrate assimilation)

From references 48 and 66.

Of special interest is the titration analysis of the complex regulation of the *A. nidulans amdS* gene, encoding acetamidase. In this case, the titration analysis has led to the identification and precise localization of transcription control sequences, which are involved in binding of regulatory proteins. Genetic data had already shown that several regulatory proteins are involved in the expression of the *amdS* gene. These proteins are also involved in the regulation of several other genes (Table II). The presence of multiple copies of the *amdS* gene or *amdS* upstream sequences in an *A. nidulans* strain resulted in a change in growth properties of this strain, indicating titration of the regulatory proteins encoded by *amdR*, *facB*, *amdA* and possibly *areA*.[56,67-69] Introduction into *A. nidulans* of defined regions of the 5'-flanking sequences of the *amdS* gene led to the localization of the FacB, AmdA and AmdR target sequences.[68] Due to only weak titration effects, the AreA target sequence(s) were not precisely localized.[69]

Similar titration effects through *amdS* sequences are also observed in fungal species related to *A. nidulans*.[70,71] *Aspergillus niger* grows very poorly on agar plates containing acetamide as the sole nitrogen source. However, transformation of *A. niger* with a vector containing the *A. nidulans amdS* gene results in transformants which show strong growth on agar plates containing acetamide.[70] Transformants with multiple copies of the *amdS* gene can be selected by their strong growth on agar plates with acrylamide as nitrogen source (B13, B38; Figure 4).[70,71] These multicopy



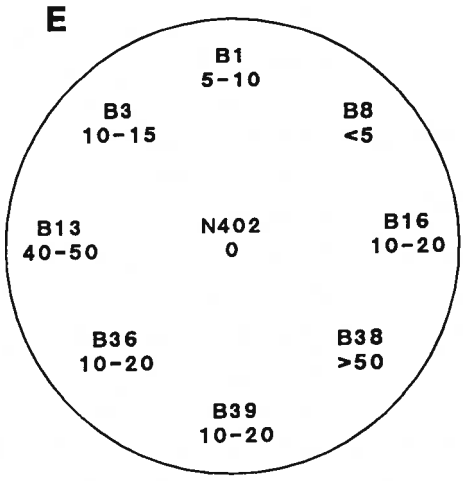


Figure 4. *Aspergillus niger* AmdS$^+$ transformants plated on **(A)** non-selective medium, **(B)** acetamide medium, **(C)** acrylamide medium, **(D)** γ-aminobutyric acid medium. **(E)** From all transformants (B1 to B39) the number of *amdS* gene copies was determined by Southern analysis. *Aspergillus niger* N402 is the recipient strain without any *amdS* gene copy.

ıminobutyric acid, as sole carbon source (Figure 4), implying that the expression genes involved in degradation of ω-amino acids is impaired.[72] This result indicates ation of a regulatory protein from *A. niger* with similar characteristics as AmdR of *nidulans*.

Strong support for titration of regulatory proteins was obtained from so-called tititration experiments. In these experiments, multiple copies of the gene encoding ə regulatory protein were introduced into strains already containing multiple copies the gene of interest. In all three cases mentioned in this section (*qutE, alcA, ıdS*), introduction of multiple copies of the relevant regulatory gene resulted in ɔreased expression of the genes of interest.[46,56,64,65]

In conclusion, the results described in this section indicate that with the aid of ation analysis, 5'-upstream sequences which are involved in binding of regulatory oteins can be localized.

2. Mutation analysis

In this type of analysis, the effects of mutations in a specific upstream region on ınscription efficiency/regulation and transcription initiation are studied. In all cases, ə amount of gene product (either of the gene corresponding to the upstream gion[60,73,74] or of a reporter gene[19,48,75-85]) is used as a measure for transcription ficiency. Transcription initiation is analyzed by nuclease S1 or primer extension ıalyses. From the data obtained by these analyses, transcription control sequences ın be identified, which are involved in the efficiency, regulation or initiation of ınscription.

The first objective in many "promoter" analysis studies is to define the region volved in transcription control. For this purpose, unidirectional deletion of upstream ɔquences is carried out and the effect of the resulting deletion mutants on the level gene expression is analysed. The results of two examples of this type of analysis e presented in Figure 5. Deletion mutants of the upstream sequences of both the *nidulans gpdA* and *oliC* gene were fused to the *E. coli lacZ* gene. Subsequently, *nidulans* transformants containing a single copy of the fusion genes integrated a specific site in the genome (*argB* locus) were generated and the amount of β-ılactosidase activity was determined in mycelial extracts from these transformants. ɜ can be seen in Figure 5, sequences as far upstream as 700 to 1000 nt from the ınscription initiation site contribute to the level of gene expression.[19,86] In several

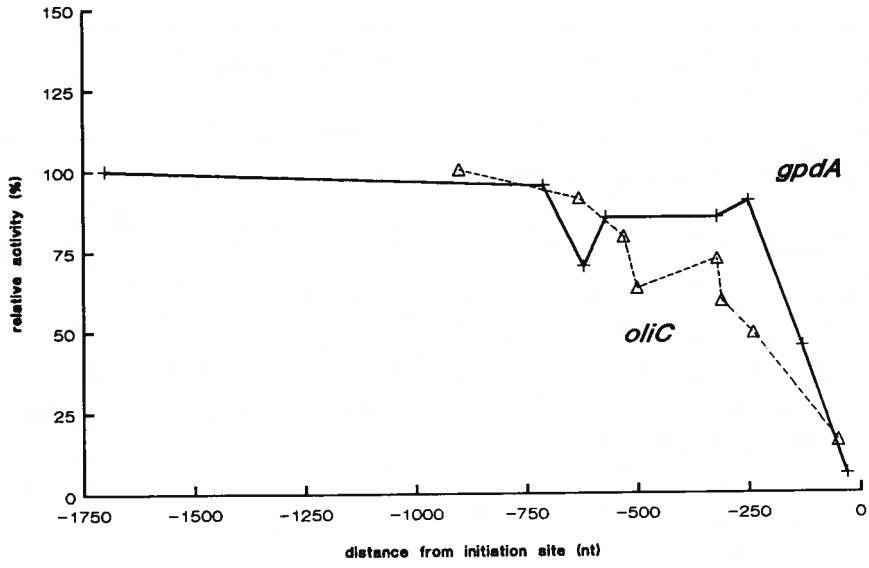


Figure 5. Relative β-galactosidase activity of *A. nidulans* transformants containing *gpdA* or *oliC* upstream sequences of various length fused to the *E. coli lacZ* gene. The distance of the 5' endpoint of the various upstream sequences from the transcription initiation site is given (in nt).

transcription control.[73,76,83] However, only very small regions, even less than 50 b upstream of the transcription initiation site, are capable of driving significa expression, as was concluded from deletion analysis of the transcription contr region of the *A. nidulans trpC, argB, oliC, abaA*, and *gpdA* genes (see also Figu 5).[19,74,76,79,82]

a. Characterization of sequences involved in transcription initiatio

For *S. cerevisiae*, TATA-like elements, PyAAG sequences, and in some case pyrimidine-rich regions in the vicinity of the transcription initiation site are involve in transcription initiation.[9,10]

Only a few studies with filamentous fungi have been reported in which this typ of sequences were analysed by deletion analysis.[19,76,79,82]

In the transcription control region of the *A. nidulans gpdA* and *oliC* gene pyrimidine-rich regions, so-called *ct* boxes, were shown to be involved in determinir

the site of transcription initiation. Deletion of one of the *ct* boxes abolished transcription initiation from the site directly downstream of this box (compare d1 with d896 in Figure 6a, and d0 with d10/d104 in Figure 6b).[19,79] Transcription initiation in that case occurred at sites downstream of other *ct* boxes. For transcription initiation of the *A. nidulans trpC* gene (encoding a trifunctional protein involved in tryptophan biosynthesis) and *abaA* gene, deletion analysis also indicated the involvement of a pyrimidine-rich region in transcription initiation.[76,82]

Deletion of the TATA box in the transcription control region of the *oliC* gene resulted in a change of the wt pattern of transcription initiation sites. Two of the initiation sites used in the wt upstream region were not used when the TATA box was deleted (compare d0 and d101 in Figure 6b).[79] In the *A. nidulans gpdA* and *abaA* genes, deletion of TATA-like sequences did not result in any change in transcription initiation sites.[19,82]

Based on the results available at present, we conclude that pyrimidine-rich regions are clearly involved in determining transcription initiation sites, whereas TATA-like sequences are not or only to a minor extent.

b. Characterization of sequences involved in transcription efficiency and regulation of gene expression

Unidirectional deletion experiments, to define the DNA region involved in transcription control, also provide information about the identification of specific sequences involved in transcription control. The regions involved in regulation of expression of the *A. nidulans argB* gene (encoding ornithine transcarbamoyl transferase)[74] and the nitrogen metabolite repression of expression of the *Penicillium chrysogenum pcbC* gene (encoding isopenicillin N synthetase)[83] were localized by this type of analysis.

An approach to localize such upstream elements more precisely is the analysis of mutant upstream sequences obtained by deletion of small specific sequences. Putative control sequences, indicated by either sequence comparison or titration analysis, can be precisely deleted from the 5'-flanking region of the gene of interest with the aid of *in vitro* mutagenesis protocols, including PCR methods.[87,88] Using this approach, functional elements in the flanking region of the *gpdA* and *oliC* gene of *A. nidulans* have been identified.[19,89] A mutant *A. nidulans gpdA* flanking region, which missed a conserved sequence present in the upstream region of the wt *gpdA*

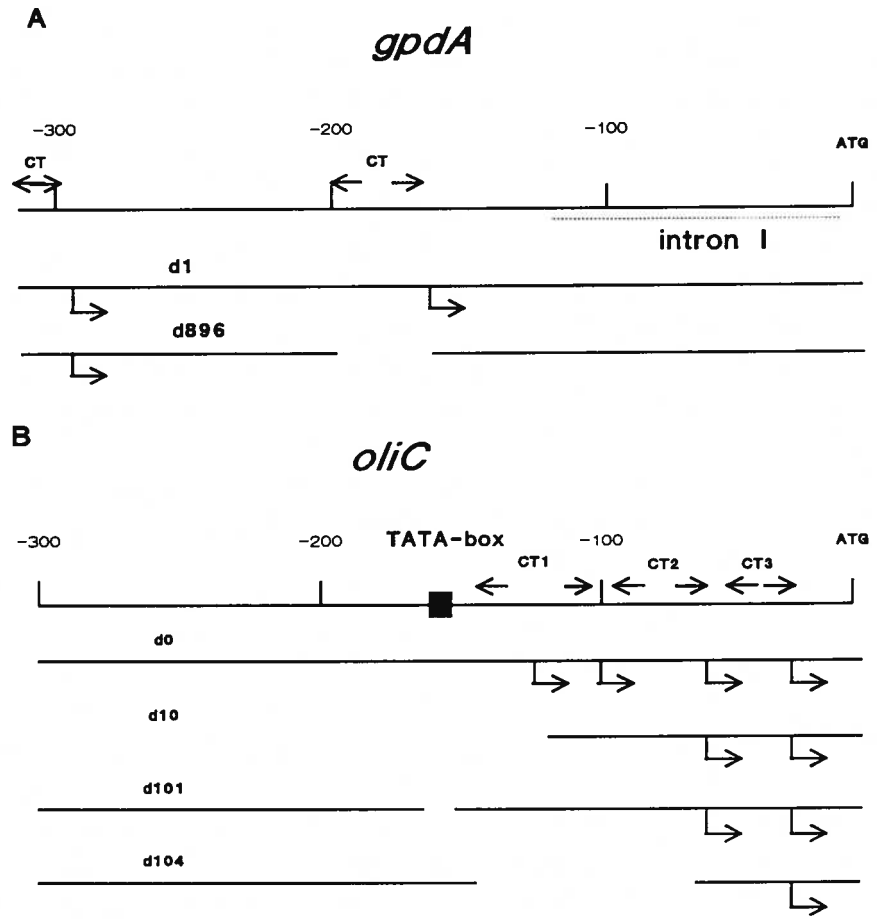


Figure 6. Transcription initiation analysis of total RNA from *A. nidulans* transformants containing mutant *gpdA* **(A)** or *oliC* **(B)** upstream sequences fused to the *E. coli lacZ* gene. Transcription initiation sites, indicated with └—>, were determined by primer extension analysis. CT, pyrimidine-rich region (*ct* box). The distance from the ATG codon is given (in nt). In **(A)** the position of an intron in the untranslated region of the *gpdA* transcript is indicated.

E. coli lacZ gene, resulted in lower expression levels (50%) compared to the levels obtained with the wt flanking region. This result indicates that this box contains (part of) a functional element.[19] Also in *A. niger* a lower expression level (30%) was found with the *A. nidulans* mutant flanking region, supporting the idea that the *gpd* box contains a functional element.[90]

To determine whether an element identified by deletion analysis contains all sequences essential for transcription activation/regulation, it is necessary to demonstrate that this element per se is functional. This can be done by introduction of the element into a 5'-flanking region of another gene and subsequent analysis of the effects on gene expression. The activity of a putative repressing element of the *abaA* gene was verified by this approach after introduction of this element into the transcription control region of the *trpC* gene. A significant decrease of the expression of the *lacZ* gene regulated by the *trpC::abaA* control sequences was observed, compared to the expression of the *lacZ* gene driven by the *trpC* control sequences.[82] Similarly, introduction of the *gpd* box into the *A. nidulans amdS* flanking region fused to *lacZ* resulted in increased levels of expression (up to 30-fold).[91] In another study, sequence elements, indicated by titration analysis, which are present in the upstream regions of the *amdS* and *gatA* gene, were introduced into the 5'-flanking region of the *amdS* gene from which most transcription control sequences were deleted. Expression from the mutant transcription control region, when fused to the *E. coli lacZ* gene, was considerably decreased compared to the expression from the wildtype *amdS* upstream region. Introduction of the *amdS* and *gatA* sequence elements (partially) restored expression of the fusion gene.[48,80]

As already mentioned, in all cases described, the amount of gene product is used as a measure for transcription efficiency. It is important to note that only when transcripts derived from different (mutant) upstream regions from a particular gene are identical, the amount of gene product faithfully represents the transcription efficiency. In several cases, transcription initiation analysis was carried out to analyze the transcripts derived from different mutant upstream regions. In general, deletion of a sequence from an upstream region, which resulted in a change in transcription efficiency, did not result in any changes in transcription initiation,[19,76,79,82] If both transcription efficiency and transcription initiation are changed as a consequence of specific deletion, direct analysis of the efficiency of transcription initiation, without interference of differences in stability and translation efficiency of the different transcripts, has to be carried out to be able to draw reliable conclusions. To our knowledge this type of experiment has not been carried out for filamentous fungi.

The approaches for identification and characterization of transcription control sequences, described thus far, are mostly based on *in vitro* generated mutations and subsequent *in vivo* analysis of the effects of these mutations. However, from genetical research on the regulation of metabolic pathways, several regulation

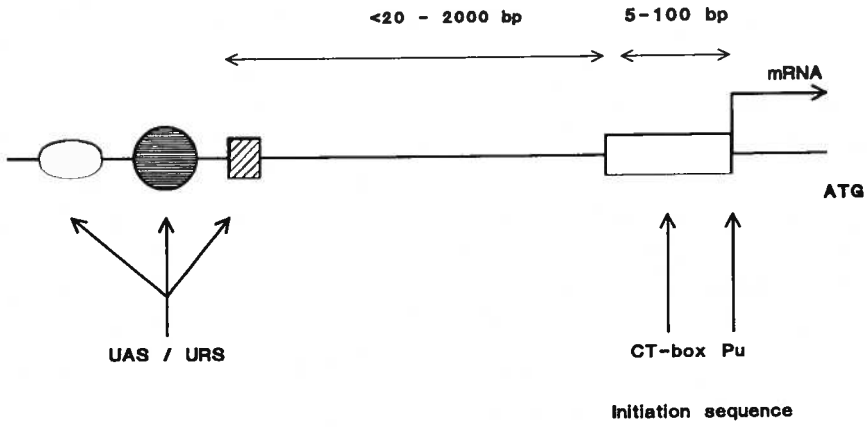


Figure 7. Scheme of fungal transcription control sequences. In most cases transcription initiation occurs at the first purine (Pu) base downstream of pyrimidine-rich regions.

5'-flanking region of the relevant genes. From these results, it was concluded tha altered regulation of expression of these genes was due to mutations in the transcription control sequences. Comparison of the upstream sequences of the *qa* 2 gene of *N. crassa*,[92] the *amdS*[3,72,93,94] and *uapA* (encoding uric acid-xanthin permease)[3,95] genes of *A. nidulans*, isolated from wildtype and mutant strains respectively, confirmed this conclusion. At the same time, this comparison als revealed the putative site of action of the different regulatory proteins involved i expression of these genes.[42,93,95]

III. CONCLUSIONS AND FUTURE PROSPECTS

In this chapter we have described a number of approaches for the analysis c transcription control sequences. From the results reviewed, it is clear that, to date only a limited amount of data concerning fungal transcription control sequences i available.

Until now, research has primarily been focused on the organization of thes sequences. A few interesting points concerning the organization of funga transcription control sequences emerge from these data. First, the region involve in transcription control can be large, extending more than 500 nt upstream of th major start site, as indicated by deletion analysis.[19,73,76,83,86] Protein DNA-bindin

analysis also indicated putative transcription control sequences at far upstream positions.[31,33,41,60] However, at least some regulatory sequences, identified by *in vitro* or *in vivo* analysis, are localized much closer to the transcription initiation site.[12,15,17,41,42,72,74,76,82] Second, pyrimidine-rich sequences are clearly involved in determining transcription initiation sites.[19,72,76,82] Authentic transcription initiation was observed from transcription control sequences consisting of only a pyrimidine-rich sequence.[19,82] The involvement of a TATA-like element in this process is indicated in only one instance,[79] whereas in a few other cases no involvement was observed.[19,82] A scheme of the 5'-flanking sequences of a fungal gene is depicted in Figure 7. Clearly, the overall structure resembles the generalized structure of eukaryotic (including yeast) transcription control regions.[7,8]

Obviously, further research on transcription regulation will not only be aimed at the elucidation of the organization of transcription control sequences, but particularly at a better understanding of the mechanisms of gene expression. Of interest in this respect, are complex regulatory circuits as, for example, those governing carbon and nitrogen metabolism, and genetic regulation of highly complex phenomena, such as, differentiation and host-pathogen interactions by pathogenic fungi.

We feel that only an integrated approach consisting of both *in vitro* and *in vivo* analysis, as described in this chapter, may lead to a further understanding of the mechanisms of gene expression. An example that clearly illustrates the necessity to combine *in vitro* and *in vivo* approaches is taken from the work of Hynes and co-workers. Sequence comparison of the 5'-flanking sequences of the *gatA* gene[80] suggested as many as four AmdR target sites. However, only one of these sequences gave positive results in titration analysis.[96] On the other hand, functional transcription control sequences may also be missed by sequence comparison, as demonstrated for FacB target sites in the *amdS* upstream sequences. Titration analysis identified a FacB site in the upstream region of the *amdS* gene not identified by sequence comparison.[96]

During the last few years molecular biological research in filamentous fungi has extended into two very interesting fields, namely, biotechnology and plant pathology.

Filamentous fungi have been used for several decades in fermentation industry for the production of proteins and primary and secondary metabolites. Based on the knowledge accumulated from molecular biological research on gene expression, the potential of fungi, especially black *Aspergillus* species and *A. oryzae*, to produce heterologous (mammalian) proteins is being investigated in several laboratories

control sequences, as described in this chapter, have already been used for optimization of the production of heterologous proteins.[65]

In the area of plant pathology, molecular biological research will focus on unravelling the very complex fungus-host interactions. The first steps in this direction are being made.[61,98] We expect that also in this area the approaches as described in this chapter will be very useful.

ACKNOWLEDGEMENTS

We wish to acknowledge all fungal scientists who communicated their unpublished data to us. We especially wish to thank Jan Verdoes, Cas Kramer, Jim Kinghorn and Geoff Turner for their contribution to this paper. Furthermore, Peter Pouwels is acknowledged for critical reading of the manuscript and stimulating discussions.

REFERENCES

1. Boylan, M. T., Mirabito, P. M., Willett, C. E., Zimmerman, C. R. and Timberlake, W. E., Isolatior and physical characterization of three essential conidiation genes from *Aspergillus nidulans* *Mol. Cell. Biol.*, 7, 3113, 1987.
2. Cove, D. J., Genetic studies of nitrate assimilation in *Aspergillus nidulans*, *Biol. Rev.*, 54, 291 1979.
3. Arst, H. N. and Scazzochio, C., Formal genetics and molecular biology of the control of gene expression in *Aspergillus nidulans*, in *Gene Manipulations in Fungi*, Bennett, J. W. and Lasure L. L., Eds., Acedemic Press, New York, 1985, 309.
4. Marzluf, G. A. and Fu, Y.-H., Molecular analyses of the nitrogen and the sulfur regulatory circuits of *Neurospora crassa*, in *Genetics and Molecular Biology of Industrial Microorganisms* Hershberger, C. L., Queener, S. W. and Hegeman, G., Eds., American Society for Microbiology Washington D. C., 1989, 279.
5. Case, M. E., Schweizer, M., Kushner, S. R. and Giles, N. H., Efficient transformation o *Neurospora crassa* by utilizing hybrid plasmid DNA, *Proc. Natl. Acad. Sci. U. S. A.*, 76, 5259 1979.
6. Van den Hondel, C. A. M. J. J., Punt, P. J., Gene-transfer systems and vector development fo filamentous fungi, in BSPI Symposium, Vol. 18, *Applied Molecular Genetics of Fungi*, Peberdy J. F., Caten, C. E., Ogden, J. E. and Bennett, J. W. Eds., British Mycological Society, Cambridge University Press, Cambridge, UK, 1991, 1.
7. Struhl, K., Yeast promoters, in, *Maximizing Gene Expression*, Biotechnol. Ser. 9, Reznikoff, W and Gold, L., Eds., Open University Press, Milton Keynes, 1986, 35.
8. Wasylik, B., Protein coding genes of higher eukaryotes: promoter elements and *trans*-acting factors, in *Maximizing Gene Expression*, Biotechnol. Ser. 9, Reznikoff, W. and Gold, L., Eds. Open University Press, Milton Keynes, 1986, 79.

CHAPTER 1

Ballance, D. J., Sequences important for gene expression in filamentous fungi, *Yeast*, 2, 229, 1986.

Gurr, S. J., Unkles, S. E. and Kinghorn, J. R., The structure and organisation of nuclear genes of filamentous fungi, in *Gene Structure in Eukaryotic Microbes*, SGM Spec. Publ. Vol. 23, Kinghorn, J.R. Ed., IRL Press, Oxford, 1988, 93.

Rambosek, J. and Leach, J., Recombinant DNA in filamentous fungi: Progress and prospects, *CRC Crit. Rev. Biotechnol.*, 6, 357, 1987.

Gwynne, D. I., Buxton, F. P., Sibley, S., Davies, R. W., Lockington, R. A., Scazzochio, C. and Sealy-Lewis, H. M., Comparison of the *cis*-acting control regions of two coordinately controlled genes involved in ethanol utilization in *Aspergillus nidulans*, *Gene*, 51, 205, 1987.

Lockington, R. A., Sealy-Lewis, H. M., Scazzochio, C. and Davies, R. W., Cloning and characterization of the ethanol utilization regulon in *Aspergillus nidulans*, *Gene*, 33, 137, 1985.

Alton, N. K., Buxton, F., Patel, V., Giles, N. H. and Vapnek, D., 5'-Untranslated sequences of two structural genes in the *qa* gene cluster of *Neurospora crassa, Proc. Natl. Acad. Sci. U. S. A.*, 79, 1955, 1982.

Hawkins, A. R., Lamb, H. K., Smith, M., Keyte, J. W. and Roberts, C. F., Molecular organisation of the quinic acid utilization (*QUT*) gene cluster of *Aspergillus nidulans*, *Mol. Gen. Genet.* 214, 224, 1988.

Baum, J. A. and Giles N. R., DNAse I hypersensitive sites within the inducible *qa* gene cluster of *Neurospora crassa, Proc. Natl. Acad. Sci. U. S. A.*, 83, 6533, 1986.

Kinghorn, J. R., personal communications, 1990.

Punt, P. J., Dingemanse, M. A., Jacobs-Meijsing, B. J. M., Pouwels, P. H. and Van den Hondel, C. A. M. J. J., Isolation and characterization of the glyceraldehyde-3-phosphate dehydrogenase gene of *Aspergillus nidulans*, *Gene* 69, 49, 1988.

Punt, P. J., Dingemanse M. A., Kuyvenhoven, A., Soede, R. D. M., Pouwels, P. H and Van den Hondel, C. A. M. J. J., Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene, encoding glyceraldehyde-3-phosphate dehydrogenase, *Gene*, 93, 101, 1990.

Ward, M., Wilson, L. J., Carmona, C. L. and Turner G., The *oliC3* gene of *Aspergillus niger*; isolation, sequence and use as a selectable marker for transformation, *Curr. Genet.*, 14, 37, 1988.

Guarente, L., Regulatory proteins in yeast, *Annu. Rev. Genet.*, 21, 425, 1987.

Johnson, P. F. and McKnight, S. L., Eukaryotic transcriptional regulatory proteins, *Annu. Rev. Biochem.*, 58, 799, 1989.

Struhl, K., Molecular mechanisms of transcriptional regulation in yeast, *Ann. Rev. Biochem.* 58, 1051, 1989.

Polyanovsky, O. L. and Stepchenko, A. G., Eukaryotic transcription factors, *BioEssays*, 12, 205, 1990.

Davis, M. A. and Hynes, M. J., Regulatory genes in *Aspergillus nidulans*, *Trends Genet.*, 5, 14, 1989.

Paluh, J. L., Orbach, M. J., Legerton, T. L., and Yanofsky, C., The cross-pathway control gene of *Neurospora crassa*, *cpc-1*, encodes a protein similar to GCN4 of yeast and the DNA-binding domain of the oncogene *v-jun*-encoded protein, *Proc. Natl. Acad. Sci. U. S. A.*, 85, 3728, 1988.

27. Paluh, J. L., Plamann, M., Kruger, D., Bartholomew, B. D., Yanofsky, C., and [illegible], Determination of the inactivating alterations in two mutant alleles of the *Neurospora crassa* cros pathway control gene *cpc-1*, *Genetics*, 124, 599, 1990.
28. Ebbole, D. J., Paluh, J. L., Plamann, M., Sachs, M. S, and Yanofsky, C., *Cpc-1*, the gene regulatory gene for genes of amino acid biosynthesis in *Neurospora crassa*, is differentia expressed during the asexual life cycle, *Mol. Cell. Biol.*, 11, 928, 1991.
29. Paluh, J. L., and Yanofsky, C., Characterization of *Neurospora cpc-1*, a *b*ZIP DNA bindi protein that does not require aligned heptad leucines for dimerization, *Mol. Cell. Biol.*, 11, 9 1991.
30. Fu, Y.-H., Paietta, J. V., Mannix, D. G., and Marzluf, G. A., *Cys-3*, the positive-acting sul regulatory gene of *Neurospora crassa*, encodes a protein with a putative leucine zipper DN binding element, *Mol. Cell. Biol.*, 9, 1120, 1989.
31. Fu, Y.-H., and Marzluf, G. A., *Cys-3*, the positive-acting sulfur regulatory gene of *Neurospo crassa*, encodes a sequence-specific DNA-binding protein, *J. Biol. Chem.*, 265, 11942, 199C
32. Fu, Y.-H., and Marzluf, G. A., *Nit-2*, the major nitrogen regulatory gene of *Neurospora cras* encodes a protein with a putative zinc finger DNA-binding domain, *Mol. Cell. Biol.*, 10, 10 1990.
33. Fu, Y.-H., and Marzluf, G. A., *Nit-2*, the major positive-acting nitrogen regulatory gene *Neurospora crassa*, encodes a sequence-specific DNA-binding protein, *Proc. Natl. Acad. S U. S. A.*, 87, 5331, 1990.
34. Fu, Y.-H., and Marzluf, G. A., Site-directed mutagenesis of the 'zinc-finger' DNA-binding dom of the nitrogen-regulatory protein NIT2 of *Neurospora*, *Mol. Microbiol.*, 4, 1847, 1990.
35. Fu, Y.-H., Kneesi, J. Y., and Marzluf, G. A., Isolation of *nit-4*, the minor nitrogen regulatory ge which mediated nitrate induction in *Neurospora crassa*, *J. Bacteriol.* 171, 4067, 1989.
36. Yuan, G. F., and Marzluf, G. A., personal communication, 1990.
37. Fu, Y.-H., Young, J. L., and Marzluf, G. A., Molecular cloning and characterization of a negati acting nitrogen regulatory gene of *Neurospora crassa*, *Mol. Gen. Genet.*, 214, 74, 1988.
38. Sorger, G. J., Brown, D., Farzannejad, M., Guerra, A., Jonathan, M., Knight, S., and Sharda, Isolation of a gene that down-regulates nitrate assimilation and influences another regulat gene in the same system, *Mol. Cell. Biol.*, 9, 4113, 1989.
39. Young, J. L., Jarai, G., Fu, Y.-H., and Marzluf, G. A., Nucleotide sequence and analysis of *NM* a negative-acting regulatory gene in the nitrogen circuit of *Neurospora crassa*, *Mol. Gen. Gen* 222, 120, 1990.
40. Kang, S., and Metzenberg, R. L., Molecular analysis of *nuc-1*$^+$, a gene controlling phosphor acquisition in *Neurospora crassa, Mol. Cell. Biol.*, 10, 5839, 1990.
41. Baum, J. A., Geever, R., and Giles, N. H., Expression of qa-1F activator protein: identificati of upstream binding sites in the *qa* gene cluster and localization of the DNA-binding doma *Mol. Cell. Biol.*, 7, 1256, 1987.
42. Geever, R. F., Huiet, L., Baum, J. A., Tyler, B. M., Patel, V. B., Rutledge, B. J., Case, M. E., a Giles, N. H., DNA sequence, organization and regulation of the *qa* gene cluster of *Neurospo crassa*, *J. Mol. Biol.*, 207, 15, 1989.

43. Mirabito, P. M., Adams, T. H., and Timberlake, W. E., Interactions of three sequentially expressed genes control temporal and spatial specificity in *Aspergillus* development, *Cell*, 57, 859, 1989.
44. Felenbok, B., Sequeval, D., Mathieu, M., Sibley, S., Gwynne, D. I., and Davies, R. W., The ethanol regulon in *Aspergillus nidulans*: characterization and sequence of the positive regulatory gene *alcR*, *Gene,* 73, 385, 1988.
45. Scazzochio, C., and Felenbok, B, personal communication, 1990
46. Andrianopoulos, A., and Hynes, M. J., Cloning and analysis of the positively acting regulatory gene *amdR* from *Aspergillus nidulans, Mol. Cell. Biol.*, 8, 3532, 1988.
47. Andrianopoulos, A., and Hynes, M. J., Sequence and functional analysis of the positively acting regulatory gene *amdR* from *Aspergillus nidulans, Mol. Cell. Biol.*, 10, 3194, 1990.
48. Hynes, M. J., Andrianopoulos, A., Davis, M. A., Van Heeswijck, R., Katz, M. E., Littlejohn, T. G., Richardson, I. B., and Saleeba, J. A., Multiple circuits regulating the *amdS* gene of *Aspergillus nidulans*, in *Proceedings of the EMBO-Alko workshop on Molecular Biology of Filamentous Fungi*, Nevalainen, H., and Penttilä, M., Eds., Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, 1989, 63.
49. Kudla, B., Caddick. M. X., Langdon, T., Martinez-Rossi, N. M., Bennett, C. F., Sibley, S., Davies, R. W., and Arst, H. N., The regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus nidulans*. Mutations affecting specificity of gene activation alter a loop residue of a putative zinc finger, *EMBO J.*, 9, 1355, 1990.
50. Scazzochio, C., Of moulds and men, or two fingers are not better than one, *Trends Genet.*, 6, 311, 1990.
51. Caddick, M. X., and Arst, H. N., Nitrogen regulation in *Aspergillus*: are two fingers better than one?, *Gene*, 95, 123, 1990.
52. Adams, T. H., Boylan, M. T., and Timberlake, W. E., *BrlA* is necessary and sufficient to direct conidiophore development in *Aspergillus nidulans, Cell*, 54, 353, 1988.
53. Adams, T. H., Deising, H., and Timberlake, W. E., *BrlA* requires both zinc fingers to induce development, *Mol. Cell. Biol.*, 10, 1815, 1990.
54. Dowzer, C. A. E., and Kelly, J. M., Cloning of the *creA* gene from *Aspergillus nidulans*: a gene involved in carbon catabolite repression, *Curr. Genet.*, 15, 457, 1989.
55. Felenbok, B., Sophianopoulou, V., Mathieu, M., Sequeval, D., Kulmburg, P., Diallinas, G., and Scazzochio, C., Regulation of genes involved in the utilization of carbon sources in *Aspergillus nidulans*, in *Proceedings of the EMBO-Alko workshop on Molecular Biology of Filamentous Fungi*, Nevalainen, H., and Penttilä, M., Eds., Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, 1989, 73.
56. Katz, M. E., and Hynes, M. J., Isolation and analysis of the acetate regulatory gene, *facB,* from *Aspergillus nidulans, Mol. Cell. Biol.*, 9, 5696, 1989.
57. Hull, E. P., Green, P. M., Arst, H. N., and Scazzocchio, C., Cloning and physical characterization of the L-proline catabolism gene cluster of *Aspergillus nidulans*, *Mol. Microbiol.*, 3, 553, 1989.
58. Beri, R. K., Whittington, H., Roberts, C. F., and Hawkins, A. R., Isolation and characterization of the positively acting regulatory gene *qutA* from *Aspergillus nidulans, Nucleic Acids Res.*, 15, 7991, 1987.

59. Lamb, H. K., Hawkins, A. R., Smith, M., Harvey, I. J., Brown, J., Turner, G., and ... Spatial and biological characterisation of the complete quinic acid utilisation gene cluster in *Aspergillus nidulans, Mol. Gen. Genet.*, 223, 17, 1990.
60. Frederick, G. D., and Kinsey, J. A., Nucleotide sequence and nuclear protein binding of the two regulatory sequences upstream of the *am* (GDH) gene in *Neurospora*, *Mol. Gen. Genet.*, 221, 148, 1990.
61. Podila, G. K., Dickman, M. B., Rogers, L. M., and Kolattuduky, P. E., Regulation of expression of fungal genes by plant signals, in *Proceedings of the EMBO-Alko workshop on Molecular Biology of Filamentous Fungi*, Nevalainen, H., and Penttilä, M., Eds., Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, 1989, 129.
62. Unkles, S. E., Hawker, K. L., Grieve, C., Campbell, E. I., and Kinghorn, J. R., *CrnA* encodes a eukaryotic nitrate transporter, *Proc. Natl. Acad. Sci. U. S. A.*, 88, 204, 1991.
63. Da Silva, A. J. F., Whittington, H., Clements, J., Roberts, C., and Hawkins, A. R., Sequence analysis and transformation by the catabolic 3-dehydroquinase (*qutE*) gene from *Aspergillus nidulans*, *Biochem. J.*, 240, 481, 1986.
64. Beri, R. K., Grant, S., Roberts, C. F., Smith, M., and Hawkins, A. R., Selective overexpression of the *qutE* gene encoding catabolic 3-dehydroquinase in multicopy transformants of *Aspergillus nidulans, Biochem. J.*, 265, 337, 1990.
65. Gwynne, D. I., Buxton, F. P., Gleeson, M. A. and Davies, R. W. Genetically engineered secretion of foreign proteins from *Aspergillus* specie, in *Protein Purification: Micro to Macro*, Burgess, R. Ed., Alan R. Liss, New York, 1987, 355.
66. Hynes, M. J., and Davis, M. A., The *amdS* gene of *Aspergillus nidulans*: control by multiple regulatory signals, *BioEssays,* 5, 123, 1986.
67. Kelly, J. M., and Hynes, M. J., Multiple copies of the *AmdS* gene of *Aspergillus nidulans* cause titration of *trans*-acting regulatory proteins, *Curr. Genet.,* 12, 21, 1987.
68. Hynes, M. J., Corrick, C. M., Kelly, J. M., Littlejohn, T. G., Identification of the sites of action for regulatory genes controlling the *amdS* gene of *Aspergillus nidulans, Mol. Cell. Biol.*, 8, 2589, 1988.
69. Hynes, M. J., and Andrianopoulos, A., Transformation studies of gene regulation in filamentous fungi, in *Genetics and Molecular Biology of Industrial Microorganisms*, Hershberger, C. L., Queener, S. W. and Hegeman, G., Eds., American Society for Microbiology, Washington, D. C., 1989, 304.
70. Kelly, J. M., and Hynes, M. J., Transformation of *Aspergillus niger* by the *amdS* gene of *Aspergillus nidulans*, *EMBO J.*, 4, 475, 1985.
71. Verdoes, J., and Van den Hondel, C. A. M. J. J., unpublished results, 1990.
72. Verdoes, J., personal communication, 1990.
73. Frederick, G. D., and Kinsey, J. A., Distant upstream regulatory sequences control the level of expression of the *am* (GDH) locus of *Neurospora crassa, Curr. Genet.*, 18, 53, 1990.
74. Goc, A., and Weglenski, P., Regulatory region of the *Aspergillus nidulans argB* gene, *Curr. Genet.* 14, 425, 1988.
75. Soliday, C. L., Dickman, M. B. and Kolattukudy, P. E., Structure of the cutinase gene and detection of promoter activity in the 5'-flanking region by fungal transformation, *J. Bacteriol.*, 171, 1942, 1989.

CHAPTER 1

. Hamer, J. E. and Timberlake, W. E., Functional organization of the *Aspergillus nidulans trpC* promoter, *Mol. Cell. Biol.*, 7, 2352, 1987.

'. Van Gorcom, R. F. M., Punt, P. J., Pouwels, P. H., and Van den Hondel, C. A. M. J. J., A system for the analysis of expression signals in *Aspergillus*, *Gene* 48, 211, 1986.

. Davis, M. A., Cobbett, C. S. and Hynes, M. J., An *amdS-lacZ* fusion for studying gene regulation in *Aspergillus*, *Gene* 63, 199, 1988.

. Turner, G., Brown, J., Kerry-Williams, S., Bailey, A. M., Ward, M., Punt. P. J., and Van den Hondel, C. A. M. J. J., Analysis of the *oliC* promoter of *Aspergillus nidulans*, in *Proceedings of the EMBO-Alko workshop on Molecular Biology of Filamentous Fungi*, Nevalainen, H., and Penttilä, M., Eds., Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, 1989, 101.

. Richardson, I. B., Hurley, S. K., and Hynes, M. J., Cloning and molecular characterisation of the *amdR* controlled *gatA* gene of *Aspergillus nidulans*, *Mol. Gen. Genet.*, 217, 118, 1989.

. Gomez-Pardo, E., and Penalva, M. A., The upstream region of the *IPNS* gene determines expression during secondary metabolism in *Aspergillus nidulans*, *Gene,* 89, 109, 1990.

:. Adams, T. H., and Timberlake, W. E., Upstream elements repress premature expression of an *Aspergillus* developmental regulatory gene, *Mol. Cell. Biol.*, 10, 4912, 1990.

. Kolar, M., Holzmann, K., Weber, G., Leitner, E., and Schwab, H., Molecular characterization and functional analysis in *Aspergillus nidulans* of the 5'-region of the *Pennicillium chrysogenum* isopenicillin N synthetase gene, *J. Biotechn.*, 17, 67, 1991.

. Greaves, P. A., and Kinghorn, J. R., personal communication, 1990.

. Sachs, M. S., and Ebbole, D., The use of *lacZ* gene fusions in *Neurospora crassa*, Fungal *Genet. Newsl.* 37, 39, 1990.

. Punt, P. J., unpublished results, 1990.

'. Kunkel, T. A., Rapid and efficient site-specific mutagenesis without phenotypic selection, *Proc. Natl. Acad. Sci. U. S. A.* 82, 488, 1985.

. PCR Protocols. A guide to methods and applications, Innes, M. A., Gelfand, D. H., Sninsky. J. J., and White, T. J., Eds., Academic Press, London, 1990.

. Turner, G., personal communication, 1990.

. Van Gorcom, R. F. M., personal communication, 1990.

. Punt, P. J., and Kramer, I., unpublished results, 1990.

:. Tyler, B. M., Geever, R. F., Case, M. E., and Giles, N. H., *Cis*-acting and *trans*-acting regulatory mutations define two types of promoters controlled by the *qa-1F* gene of *Neurospora, Cell,* 36, 493, 1984.

. Corrick, C. M., Twomey, A. P., and Hynes, M. J., The nucleotide sequence of the *amdS* gene of *Aspergillus nidulans* and the molecular characterization of 5'-mutations, *Gene*, 53, 63, 1987.

. Katz, M. E., Saleeba, J. A., Sapats, S. I., and Hynes, M. J., A mutation affecting *amdS* expression in *Aspergillus nidulans* contains a triplication of *cis*-acting regulatory sequence, *Mol. Gen. Genet,* 220, 373, 1990.

. Diallinas G., and Scazzochio, C., A gene coding for uric acid-xanthine permease of *Aspergillus nidulans*: inactivational cloning, characterization, and sequence of a *cis*-acting mutation, *Genetics* 122, 341, 1989.

Hynes, M. J., personal communication, 1990.

97. Van den Hondel, C. A. M. J. J., Punt, P. J., and Van Gorcom, R. F. M., Heterologous ge expression in filamentous fungi, in *More Gene Manipulation in Fungi*, Bennett, J. W., and Lasu L. L., Eds., Academic Press, San Diego, 1991, 396.
98. De Wit, P. J. G. M., and Oliver, R. P., The interaction between *Cladosporium fulvum* (syn. *Fuh fulva*) and tomato: A model system in molecular plant pathology, in *Proceedings of the EMB Alko Workshop on Molecular Biology of Filamentous Fungi*, Nevalainen, H., and Penttilä, I Eds., Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, 1989, 227.

Gene, 69 (1988) 49-57

CHAPTER 2

lation and characterization of the glyceraldehyde-3-phosphate ıydrogenase gene of *Aspergillus nidulans.*

combinant DNA; heterologous hybridization; gene amplification; intron; nucleotide uencing; cDNA cloning; gene libraries; phage λ vector)

er J. Punt, Maria A. Dingemanse, Brigit J. M. Jacobs-Meijsing, Peter H. ıwels and Cees A. M. J. J. van den Hondel

) *Medical Biological Laboratory, Rijswijk (The Netherlands)*

MMARY

he isolation and characterization of the highly expressed glyceraldehyde-3-)sphate dehydrogenase (GPD)-coding gene (*gpdA*) of *Aspergillus nidulans* is scribed. The gene was isolated from an *A. nidulans* λ gene library with a *:charomyces cerevisiae* GPD-coding gene as a probe. Unlike many other caryotes, *A. nidulans* contains only one GPD-coding gene. At the amino acid level, nology with other GPD enzymes is extensive. The *A. nidulans* gene contains ren introns, one of which is positioned in the 5'-untranslated part of the gene. The jor transcription start point is found at 172 bp upstream from the start codon. mparison of 5' and 3' flanking sequences with flanking sequences of other highly)ressed (glycolytic) genes shows several regions of similar sequence.

TRODUCTION

Glyceraldehyde-3-phosphate dehydrogenase (GPD; EC 1.2.1.12) plays a central e in glycolysis and gluconeogenesis. In glycolysis it converts glyceraldehyde-3-osphate into biphosphoglycerate, in gluconeogenesis it catalyses the reserve ıction. Much is known about the structure of the enzyme in different organisms arris and Waters, 1976; Skarzynski et al., 1987). Also, the nucleotide sequence of GPD-coding genes of many prokaryotes and eukaryotes has been determined

(e. g., Holland and Holland, 1980; Stone et al., 1985a; Tso et al., 1985b; Michels al., 1986; Yarbrough et al., 1987). Comparison of the structure of the different G enzymes and the nucleotide sequence of their genes shows a high degree conservation among different species (Fothergill-Gilmore, 1986; Yarbrough et 1987).

In several species multiple (in higher eukaryotes up to several hundreds) Gl coding genes are present in the genome (Tso et al., 1985a; Michels et al., 19 Yarbrough et al., 1987). In some cases presumably only one of these genes transcriptionally active, the other copies being pseudogenes (Hanauer and Manc 1984; Fort et al., 1985). In several other cases GPD is synthesized from multi genes which sometimes are differentially expressed (McAlister and Holland, 19 Tso et al., 1985b).

In *S. cerevisiae* and rat muscle up to 5% of the total amount of cellular prot consists of GPD (Krebs et al., 1953; Piechaczyk et al., 1984). This implies that expression signals of the GPD-coding gene(s) are strong, as was demonstrated Edens et al. (1984).

In our research on gene expression and gene regulation in filamentous fungi expression signals of *A. nidulans* genes are being analyzed (Van Gorcom et 1986). Both structural and functional features of these expression signals are un research. In this paper the isolation of the *A. nidulans gpdA* gene is described. T complete nucleotide sequence of the gene and its 5' and 3' flanking regions w determined. Furthermore, the nucleotide sequence of the messenger RNA w determined using cDNA clones and poly(A)$^+$RNA as templates.

MATERIALS AND METHODS

(a) Strains and plasmids

Escherichia coli K-12 JM109 (Yanisch-Perron et al., 1985) was used for t construction and propagation of vector molecules. *A. nidulans* strain FGS (Glasgow wild-type; Clutterbuck, 1986) was used for the construction of the *nidulans* cDNA library. Poly(A)$^+$RNA from *A. nidulans* strain MH1277[pAN45-1A was used as a template for mRNA sequence determination. This strain was obtain by transformation of *A. nidulans* MH1277 (*biA*1, *amdS*320, *amdI*18, *amdA*7, *niiA* Hynes et al., 1983) with plasmid pAN45-1A, which contains the *A. nidulans gp* gene and the *amdS* selection marker. The strain contains about nine copies of t

gpdA gene (J. Dekker, in preparation). Plasmid pFL1-33 containing the *S. cerevisiae* GPD-coding gene (*gap63/tdh2;* Holland and Holland, 1980) was obtained from Dr. L. E. Edens.

(b) Gene libraries

An *A. nidulans* λCharon 4a gene library was kindly provided by Dr. W. E. Timberlake (Orr and Timberlake, 1982). A partial cDNA library was constructed using poly(A)$^+$RNA isolated from a culture of *A. nidulans* FGSC4 cultivated in minimal medium (Pontecorvo et al., 1953) with 2% galactose as a carbon source. The ds cDNA was prepared and cloned as described by Teeri et al. (1987).

(c) DNA/RNA manipulations

A. nidulans poly(A)$^+$RNA was isolated as described by Teeri et al. (1987). Primer extension experiments were performed as described for first-strand cDNA synthesis (Teeri et al., 1987). Heterologous hybridization experiments were carried out at 56°C with final washes in 3 x SSC, 0.1% SDS, 0.1% Na.pyrophosphate at 56°C, as described by Van Hartingsveldt et al. (1987). All other DNA/RNA manipulations were carried out as described in Maniatis et al. (1982).

RESULTS AND DISCUSSIONS

(a) Isolation of the *Aspergillus nidulans gpdA* gene

The *A. nidulans gpdA* gene was isolated from an *A. nidulans* FGSC4 λ library by heterologous hybridizations with a DNA fragment containing one of the *S. cerevisiae* GPD-coding genes (*gap63/tdh2*; Holland and Holland, 1980) as a probe. From 25000 λ clones screened, five positive clones were obtained. Restriction enzyme analysis revealed that the inserts in these clones had a 4.0-kb *Bgl*II-*Hin*dIII fragment in common. Southern-blot analysis showed that only this fragment hybridized with the *S. cerevisiae* probe under heterologous hybridization conditions (results not shown), suggesting that a complete *A. nidulans gpd* gene is located on the fragment.

Two lines of evidence confirmed that the complete *A. nidulans gpdA* gene had

of a part of the *Bgl*II-*Hin*dIII fragment revealed a clear similarity with *S. cerevisiae tdh2* and other *gpd* sequences (for details see RESULTS AND DISCUSSION, section **b**). Second, introduction of several copies of the putative *gpdA* gene in *A. nidulans* resulted in an increase GPD enzyme activity. The level of GPD enzyme activity appeared to correlate with the number of gene copies (J. Dekker, in preparation). This indicates that the *Bgl*II-*Hin*dIII fragment contains a functional copy of the *gpdA* gene together with its expression signals.

The 4.0-kb *Bgl*II-*Hin*dIII fragment of one of the λ clones was subcloned into a pBR322 derivative, containing a polylinker with a unique *Bgl*II and *Hin*dIII site, resulting in plasmid pAN5-22 (Fig. 1). Southern-blot analysis of genomic *A. nidulans* DNA digested with *Bgl*II + *Hin*dIII showed that a single band of 4.0 kb hybridized with pAN5-22 under stringent conditions (results not shown), indicating that the *A. nidulans* genome, unlike *S. cerevisiae* and many other eukaryotic genomes, contains only one GPD-coding gene.

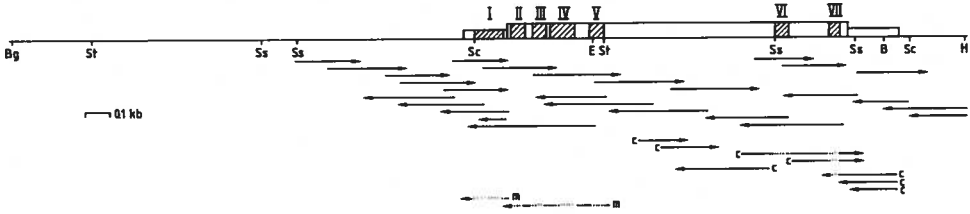


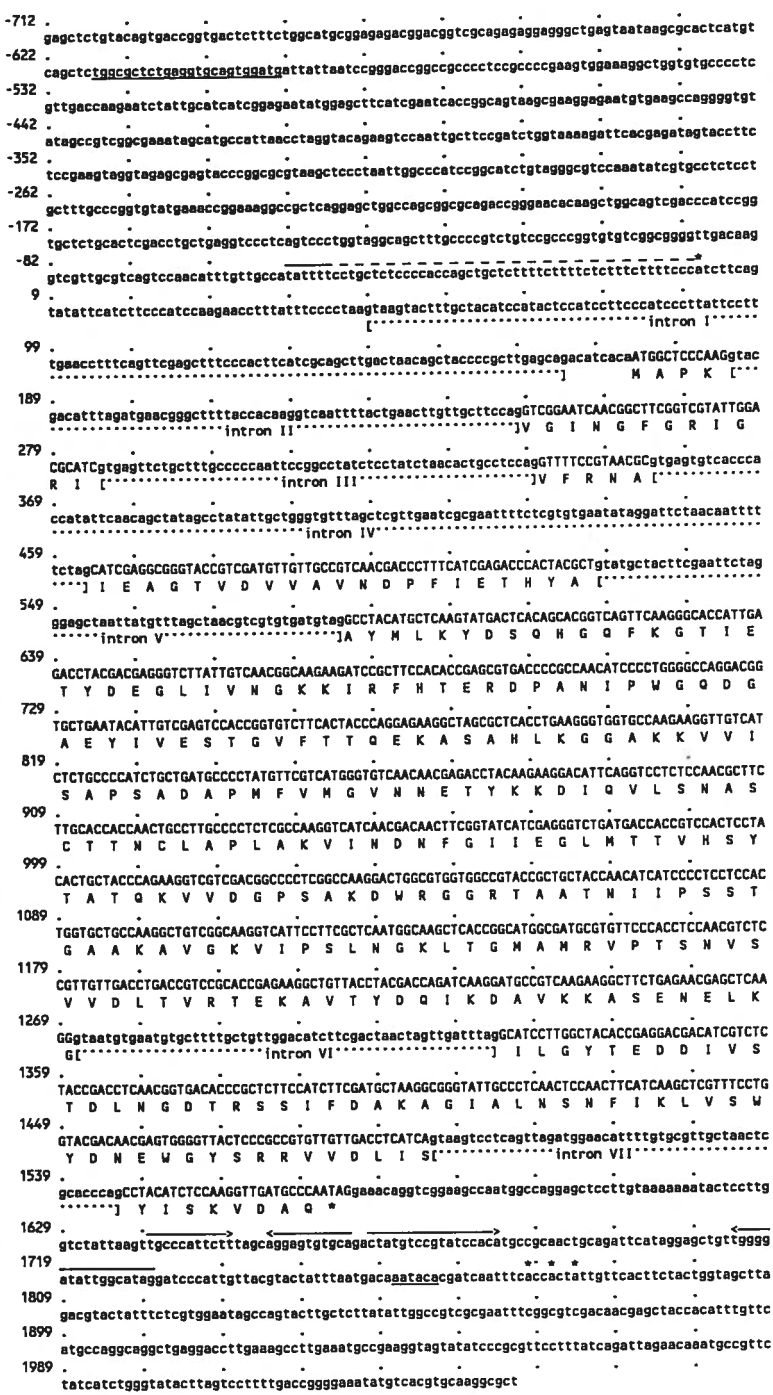
Fig. 1. Nucleotide sequencing strategy for the *A. nidulans gpdA* gene. Nucleotide sequence data were obtained using dideoxy chain-termination methods with ss-DNA (Sanger et al., 1977), ds-DNA (Cheng and Seeburg, 1985) and poly(A)$^+$RNA (Johanningmeier et al., 1987) as templates. Nucleotide sequences were analysed using UWGCG analysis programs (Deveraux et al., 1984). The top line represents the map of the 4.0-kb *Bgl*II-*Hin*dIII insert in pAN5-22. The coding region of the *gpdA* gene is indicated by open wide bars, the 5'- and 3'- noncoding regions by narrow bars. The introns are indicated by hatched bars numbered with roman numbers. Relevant restriction enzyme sites are indicated as follows: B, *Bam*HI;Bg, *BglII*; E, *Eco*RI; H, *Hin*dIII; Sc, *Sca*I; Ss, *Sst*I; St, *Stu*I. Arrows indicate the position, length and direction of the nucleotide sequence determined. Arrows marked with 'c' represent sequences obtained from cDNA clones, those marked with 'm' sequences obtained from poly(A)$^+$RNA. All other sequences were obtained from subclones of pAN5-22.

(b) Structure of the *gpdA* gene and GPD enzyme

The sequence strategy used to determine the nucleotide sequence of the *gpdA* gene and its flanking regions is given in Fig. 1. The sequence is shown in Fig. 2. The major part of the nucleotide sequence of the *gpdA* mRNA was also determined, using incomplete cDNA clones and poly(A)$^+$RNA as templates. By comparison of the genomic and the mRNA sequences the presence of five introns could be established in the 5' part of the transcribed regions of the *A. nidulans gpdA* gene (Fig. 3). Comparison of the genomic sequence with that of the cDNA clones revealed the presence of two additional introns in the 3' part of the transcribed region (results not shown). As can be seen in Fig. 1, a small part of the mRNA sequence was not investigated. No obvious intron donor or acceptor sites were found in the corresponding part of the nucleotide sequence. The features of the introns are described in RESULTS AND DISCUSSION, section **c**.

The similarity between the predicted amino acid sequence of the GPD polypeptide of *A. nidulans* (Fig. 2) and the sequences of *Nicotiana tabacum* (cytosolic) (Shih et al., 1986), *Drosophila melanogaster* (Tso et al, 1985b), chicken (Stone et al., 1985a), man (Tso et al., 1985a), rat (Tso et al., 1985a), *S. cerevisiae* (Holland and Holland, 1980) and *E. coli* (Branlant and Branlant, 1985) is 65-70%. Between the *A. nidulans* GPD polypeptide and those of *Bacillus stearothermophilus* (Walker et al., 1980a) and *Thermus aquaticus* (Walker et al., 1980b) the similarity is 50-55%. In parts of the GPD polypeptide known to be essential for enzymatic activity (Harris and Waters, 1976) similarity is almost 100%. Such relatively high percentages of similarity between different homologous polypeptides have also been found for other glycolytic enzymes (Pichersky et al., 1984; Tani et al., 1985; Forthergill-Gilmore, 1986).

Codon usage in the *A. nidulans gpdA* gene is clearly biased, with a preference for a pyrimidine in the third position. Of all codons 79% have a pyrimidine in that position (C:55%, T:24%, G:20%, A:2%), and when a choice between a purine or a pyrimidine is allowed, in 93% of the cases a pyrimidine is chosen. This bias is similar to that found for other highly expressed genes in filamentous fungi (Kinnaird and Fincham, 1983; Clements and Roberts, 1986; May et al., 1987) but clearly different from that in highly expressed genes of *S. cerevisiae* (Bennetzen and Hall, 1982).


-712
gagctctgtacagtgaccggtgactctttctggcatgcggagagacggacggtcgcagagaggagggctgagtaataagcgcactcatgt
-622
cagctctggcgctctgaggtgcagtggatgattattaatccgggaccggccgcccctccgccccgaagtggaaaggctggtgtgccccctc
-532
gttgaccaagaatctattgcatcatcggagaatatggagcttcatcgaatcaccggcagtaagcgaaggagaatgtgaagccaggggtgt
-442
atagccgtcggcgaaatagcatgccattaacctaggtacagaagtccaattgcttccgatctggtaaaagattcacgagatagtaccttc
-352
tccgaagtaggtagagcgagtacccggcgcgtaagctccctaattggcccatccggcatctgtagggcgtccaaatatcgtgcctctcct
-262
gctttgcccggtgtatgaaacccggaaaggccgctcaggagctggccagcggcgcagaccgggaacacaagctggcagtcgacccatccgg
-172
tgctctgcactcgacctgctgaggtccctcagtccctggtaggcagctttgccccgtctgtccgcccggtgtgtcggcggggttgacaag
-82
gtcgttgcgtcagtccaacatttgttgccatattttcctgctctcccccaccagctgtctttttcttttctcttttcttttcccatcttcag
9
tatattcatcttcccatccaagaacctttatttcccctaagtaagtactttgctacatccatactccatcctcccatccctttattcctt
intron I
99
tgaacctttcagttcgagctttcccacttcatcgcagcttgactaacagctaccccgcttgagcagacatcacaATGGCTCCCAAGgtac
M A P K
189
gacatttagatgaacgggcttttaccacaaggtcaattttactgaacttgttgcttccaggTCGGAATCAACGGCTTCGGTCGTATTGGA
intron II
V G I N G F G R I G
279
CGCATCgtgagttctgctttgcccccaattccggcctatctcctatctaacactgcctccaggTTTTCCGTAACGCgtgagtgtcaccca
R I
intron III
V F R N A
369
ccatattcaacagctatagcctatattgctgggtgtttagctcgttgaatcgcgaatttttctcgtgtgaatataggattctaacaatttt
intron IV
459
tctagCATCGAGGCGGGTACCGTCGATGTTGTTGCCGTCAACGACCCTTTCATCGAGACCCACTACGCTgtatgctacttcgaattctag
I E A G T V D V V A V N D P F I E T H Y A
549
ggagctaatttatgtttagctaacgtcgtgtgatgtagGCCTACATGCTCAAGTATGACTCACAGCACGGTCAGTTCAAGGGCACCATTGA
intron V
A Y M L K Y D S Q H G Q F K G T I E
639
GACCTACGACGAGGGTCTTATTGTCAACGGCAAGAAGATCCGCTTCCACACCGAGCGTGACCCCGCCAACATCCCCTGGGGCCAGGACGG
T Y D E G L I V N G K K I R F H T E R D P A N I P W G Q D G
729
TGCTGAATACATTGTCGAGTCCACCGGTGTCTTCACTACCCAGGAGAAGGCTAGCGCTCACCTGAAGGGTGGTGCCAAGAAGGTTGTCAT
A E Y I V E S T G V F T T Q E K A S A H L K G G A K K V V I
819
CTCTGCCCCATCTGCTGATGCCCCTATGTTCGTCATGGGTGTCAACAACGAGACCTACAAGAAGGACATTCAGGTCCTCTCCAACGCTTC
S A P S A D A P M F V M G V N N E T Y K K D I Q V L S N A S
909
TTGCACCACCAACTGCCTTGCCCCTCTCGCCAAGGTCATCAACGACAACTTCGGTATCATCGAGGGTCTGATGACCACCGTCCACTCCTA
C T T N C L A P L A K V I N D N F G I I E G L M T T V H S Y
999
CACTGCTACCCAGAAGGTCGTCGACGGCCCCTCGGCCAAGGACTGGCGTGGTGGCCGTACCGCTGCTACCAACATCATCCCCTCCTCCAC
T A T Q K V V D G P S A K D W R G G R T A A T N I I P S S T
1089
TGGTGCTGCCAAGGCTGTCGGCAAGGTCATTCCTTCGCTCAATGGCAAGCTCACCGGCATGGCGATGCGTGTTCCCACCTCCAACGTCTC
G A A K A V G K V I P S L N G K L T G M A M R V P T S N V S
1179
CGTTGTTGACCTGACCGTCCGCACCGAGAAGGCTGTTACCTACGACCAGATCAAGGATGCCGTCAAGAAGGCTTCTGAGAACGAGCTCAA
V V D L T V R T E K A V T Y D Q I K D A V K K A S E N E L K
1269
GGGtaatgtgaatgtgcttttgctgttggacatcttcgactaactagttgatttagGCATCCTTGGCTACACCGAGGACGACATCGTCTC
G[
intron VI
] I L G Y T E D D I V S
1359
TACCGACCTCAACGGTGACACCCGCTCTTCCATCTTCGATGCTAAGGCGGGTATTGCCCTCAACTCCAACTTCATCAAGCTCGTTTCCTG
T D L N G D T R S S I F D A K A G I A L N S N F I K L V S W
1449
GTACGACAACGAGTGGGGTTACTCCCGCCGTGTTGTTGACCTCATCAgtaagtcctcagttagatggaacattttgtgcgttgctaactc
Y D N E W G Y S R R V V D L I S[
intron VII
1539
gcacccagCCTACATCTCCAAGGTTGATGCCCAATAGgaaacaggtcggaagccaatggccaggagctccttgtaaaaaaatactccttg
] Y I S K V D A Q *
1629
gtctattaagttgcccattctttagcaggagtgtgcagactatgtccgtatcccacatgccgcaactgcagattcataggagctgttgggg
1719
atattggcataggatcccattgttacgtactatttaatgacaaatacacgatcaatttcaccactattgttcacttctactggtagctta
1809
gacgtactatttctcgtggaatagccagtacttgctcttatattggccgtcgcgaatttcggcgtcgacaacgagctaccacatttgttc
1899
atgccaggcaggctgaggaccttgaaagccttgaaatgccgaaggtagtatatcccgcgttcctttatcagattagaacaaatgccgttc
1989
tatcatctgggtatacttagtccttttgaccggggaaatatgtcacgtgcaaggcgct

(c) Introns

Of the seven introns detected by sequence analysis (Fig. 1) introns II to VII interrupt the coding region of the gene, intron I is part of the 5'-noncoding region of the gene. The nucleotide sequence of the exon-intron boundaries of all introns of the *gpdA* gene fit (with one or two mismatches) to the consensus sequence for fungal introns (Ballance, 1986). The size of the introns varies from 50 to 120 bp.

Comparison of the position of introns in different GPD-coding genes (Fig. 4) shows that intron V of the *A. nidulans* gene coincides with intron III of the chicken GPD-coding gene (Stone et al., 1985a). Furthermore, in the chicken as well as in a *D. melanogaster* GPD-coding gene the 5' part of the gene corresponding to the untranslated region of the mRNA contains an intron at 10-20 nt upstream from the start codon (ATG)(Stone et al., 1985a; Tso et al., 1985b). None of the other introns in the *A. nidulans* gene coincides with an intron found in one of the other GPD-coding genes of which the mRNA and/or genomic nucleotide sequence in known.

These results do not support the hypothesis that introns originally mediated exon assembly and thus are presented in homologous genes at corresponding positions at boundaries of regions encoding structural domains. Such a hypothesis was proposed on the basis of analysis of different triosephosphate isomerase genes (Straus and Gilbert, 1985; Gilbert et al., 1986) and the chicken GPD-coding gene (Stone et al., 1985b). The presence of very small exons in the *A. nidulans gpdA* gene (too small to be structural domains) is not consistent either with this hypothesis, although introns very close to each other may have resulted from

---

Fig. 2. Nucleotide sequence of the *gpdA* gene of *A. nidulans* and the predicted amino acid sequence. Coding regions are indicated in upper-case letters, all other sequences in lower-case letters. Below the coding regions, the predicted amino acid sequence is given using the standard one-letter code. Nucleotides are numbered with reference to the transcription start point (+1). The introns, numbered by roman numbers (see Fig. 1), are indicated by dotted underlining. The major transcription start point (+1) and the polyadenylation sites (+1778, +1781, +1784) are indicated by asterisks. The putative TATA box (-52 to -47) is overlined. The C + T-rich region (-47 to -1) is overlined with a dashed line. The putative polydenylation signal (+1760 to +1766) is underlined. The inverted repeats in the 3'-noncoding region (between +1640 and + 1730) are indicated by pairs of convergent arrows. The sequence upstream of the TATA box (-616 to -593) showing a clear similarity to a sequence upstream from the TATA box of the *A. nidulans pgk* gene is indicated by underlining.

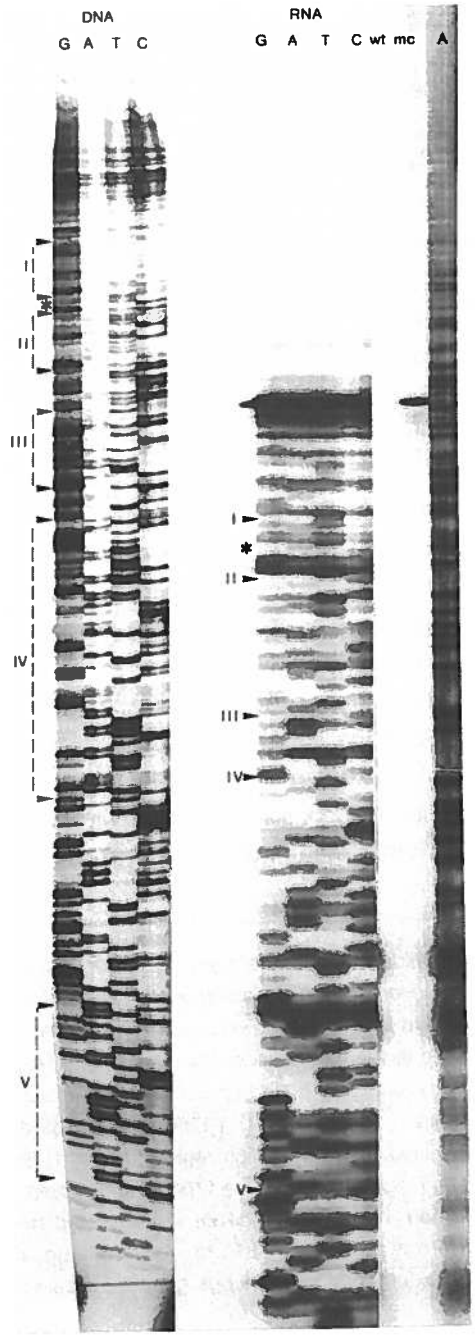


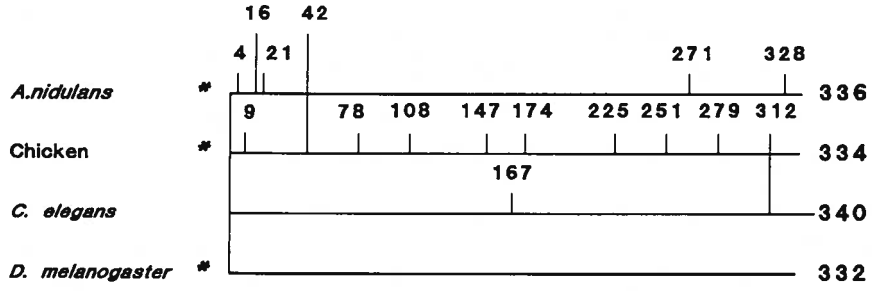
Fig. 3. Sequence analysis [illegible]

analysis to locate the transcription start point and the position of introns. Lanes G, A, T, C show the products of dideoxy chain-termination sequencing reactions with pAN5-22 DNA (indicated by 'DNA') and poly(A)$^+$RNA from *A. nidulans* MH1277 [pAN45-1A]1 (indicated by 'RNA') primed with an oligodeoxynucleotide primer (5'-CTCAATGGTGCCCTTGAACTGACCGTGC-3') complementary to a part of the coding region 3' of intron V. In the sequence ladder the exon-intron boundaries of the first five introns are indicated by arrowheads and roman numbers. In the sequence ladder obtained with poly(A)$^+$RNA the position of these introns is indicated by numbered arrowheads. The position of the translation start codon (ATG) is indicated by an asterisk (note that the sequence, read from bottom to top, is complementary to the RNA sequence and has the opposite polarity, thus ATG is read as CAT in the figure). In the RNA sequence several bands across all four lanes appear, probably due to stops caused by secondary structure of the RNA. To determine the transcription start point(s) (which were obscured in the RNA sequence as a consequence of aspecific bands) primer extension experiments were carries out with the same oligo-deoxynucleotide primer used for sequencing. Poly(A)$^+$RNA from *A. nidulans* FGSC4 (wt) and MH1277[pAN45-1A]1, containing multiple copies (mc) of the *gpdA* gene, was used for primer extension reactions. In the 'wt' and 'mc' lanes extension products from these RNA preparations are shown. In both cases one major band was observed. Minor wt bands (identical to additional mc bands) can only be seen after longer exposure. The rightmost lane A shows sequence reaction products used as size markers for the primer extension products.

lication and movement of early introns (Gilbert et al., 1986).
he presence of an intron outside of the coding region has been reported for one
ar fungal gene (Saloheimo et al., 1988). In several genes of higher eukaryotes,
uding the chicken (Stone et al., 1985a) and *D. melanogaster* (Tso et al., 1985b)
)-coding genes, introns outside the coding region have been observed.

3'-Noncoding sequences

ne sequence of the 3' end of the *gpdA* mRNA was determined by analysis of
A cDNA. In five out of eight cDNA clones analyzed, the poly(A) track started at
osition +207 downstream from the stop codon (UAG), in two clones at nt
ition +204 and in one clone at +201 (Fig. 2). Thus, some heterogeneity was
erved at the site of polyadenylation. The length of the poly(A) track in not known,
is at least 70 nt since this length was found in one of the cDNA clones. A
itive polyadenylation signal (AAUACA) was found 11-17 bp upstream from the
'(A) track. This sequence is related to a sequence (AAUAAA) found at
iparable sites in many other eukaryotic messengers (Proudfoot and Brownlee,
6). The 3'-noncoding sequence of the messenger shows stretches of dyad
metry, as indicated in Fig. 2. Comparable results were obtained for other *A.*
lans genes (Clements and Roberts, 1986; Ward and Turner, 1986). These
uences might be functional in 3' processing and polyadenylation of the precursor
ne messenger RNA (Platt, 1986).



g. 4. Position of the introns in the GPD-coding genes of *A. nidulans* (this report), chicken
tone et al., 1985a), *C. elegans* (Yarbrough et al., 1987) and *D. melanogster* (Tso et al.,
85b). The location of the introns in the genes is indicated by the codon numbers of the *A.*
fulans gene. The total number of codons in each gene is also given. Asterisks indicate the
sition of an intron in the 5'-noncoding region of the GPD-coding genes.

(e) 5'-Noncoding sequences

The transcription start point(s) of the *gpdA* gene were localized by sequer analysis of the *gpdA* messenger and by primer extension analysis. The m transcription start point is localized 172 bp (corresponding to a leader region of nt in the mRNA) upstream from the translation start codon (ATG) (Fig. 3). Sc minor sites were found between 140 and 170 bp in front of the ATG codon. In sequence immediately upstream from the major site a prominent C+T-rich reg of about 50 bp was observed. This sequence is preceded by a putative T/ element (TATTTT). Similar sequences have been observed upstream from ot *Aspergillus* and *N. crassa* genes (Ballance, 1986). In the highly expressed *nidulans oliC* (Ward and Turner, 1986) and *N. crassa am* (Kinnaird and Finch 1983) genes the C+T-rich region is particulary long. It has been suggested the length of the C+T-rich region between the TATA box and the major transcrip site in A. nidulans is related to the level of transcription (Ballance, 1986).

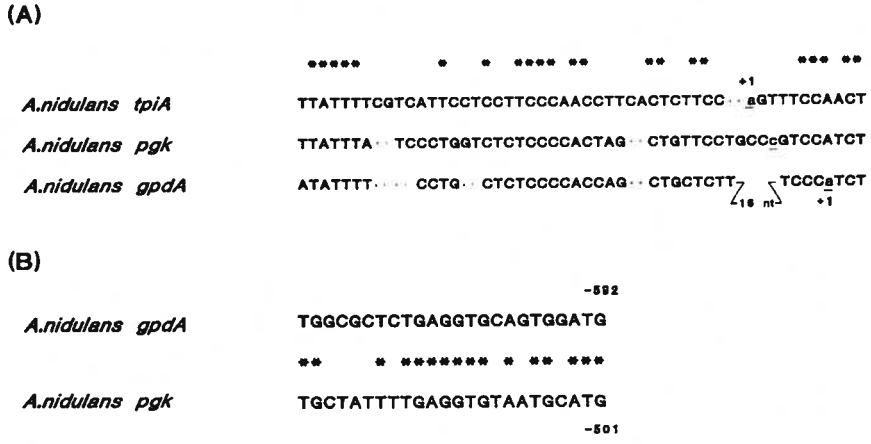


Fig. 5. Comparison of sequences of the 5' flanking region of glycolytic genes of *A. nidulan* (Part A) Comparison of the region around the transcription start point. The sequences of t *gpdA* (this paper), *pgk* (Clements and Roberts, 1986) and *tpiA* (McKnight et al., 1986) gen are aligned for maximal similarity by introducing gaps (indicated by dots). Nucleotides identic for all three genes are indicated by asterisks. The major transcription start point is indicate by an underlined lower-case letter. (Part B) Comparison of 5' upstream sequences of the *gpc* and *pgk* genes. The distance (in nt) from the transcription start point is given. Identic nucleotides are indicated by asterisks.

Analysis of the 5'-noncoding sequence of the *gpdA* messenger showed the presence of a repetitive oligomer (CCAUCU), one to several copies of which have also been found in many other *A. nidulans* genes (Ward and Turner, 1986).

The sequence around the AUG codon (ACAAUGG) which is thought to play an important role in the efficiency of translation initiation in eukaryotes (Kozak, 1986) closely resembles the consensus sequence ACCAUGG in higher eukaryotes (Kozak, 1986) and the consensus sequence AAAAUGG in glycolytic genes of *S. cerevisiae* (Cigan and Donahue, 1987).

Comparison of the sequences between the putative TATA box and the major transcription start point of three *A. nidulans* glycolytic genes (*gpdA*, this paper; *pgk*, Clements and Roberts, 1986; and *tpiA*, McKnight et al., 1986) shows a clear similarity (Fig. 5A). Comparison of sequences upstream from the TATA box of the *gpdA* gene and the *pgk* gene shows a region of similar sequence around 500-600 nt upstream from the transcription start point (Fig. 5B).

(f) Conclusions

The determination of the nucleotide sequence of the 5' flanking region of the *A. nidulans gpdA* gene and the comparison of this sequence with that of other genes suggests the presence of several regulatory regions. Detailed functional analysis using *lacZ* gene fusions (Van Gorcom et al., 1986) is in progress to correlate functional and structural features of the 5' flanking region.

Heterologous gene expression controlles by expression signals has been demonstrated in a number of *Aspergillus* species (van Gorcom et al., 1986; Punt et al., 1987; Mattern et al., 1987; Mullaney et al., 1988), *Penicillium chrysogenum* (Kolar et al., 1988), *Trichoderma reesei* (Penttilä et al., 1987) and in other filamentous fungi, showing that the expression signals of the *A. nidulans gpdA* gene are a useful tool in heterologous gene expression in filamentous fungi.

ACKNOWLEDGEMENTS

We gratefully acknowledge the valuable advice and assistance with regard to the construction of the cDNA clones by J. K. C. Knowles and T. T. Teeri. The authors wish to thank R. F. M. van Gorcom, F. Bleichrodt and P. Crowley for critical reading of the manuscript.

Ballance, D. J.: Sequences important for gene expression in filamentous fungi. Yeast 2 (1986) 229-236.

Bennetzen, J. L. and Hall, B. D.: Codon selection in yeast. J. Biol. Chem. 257 (1982) 3026-3031.

Branlant, G. and Branlant, C.: Nucleotide sequence of the *Escherichia coli gap* gene. Different evolutionary behaviour of the NAD^+-binding domain and of the catalytic domain of D-glyceraldehyde-3-phosphate dehydrogenase. Eur. J. Biochem. 150 (1985) 61-66.

Chen, E. Y. and Seeburg, P. H.: Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA 4 (1985) 165-170.

Cigan, A. M. and Donahue, T. F.: Sequence and structural features associated with translational initiator regions in yeast - a review. Gene 59 (1987) 1-18.

Clements, J. M. and Roberts, C. F.: Transcription and processing signals in the 3-phosphoglycerate kinase (PGK) gene from *Aspergillus nidulans*. Gene 44 (1986) 97-105.

Clutterbuck, A. J.: Glasgow stock list of *Aspergillus nidulans* strains 1986. Fungal Genet. Newslett. 33 (1986) 59-69.

Devereux, J., Haeberli, P. and Smithies, O.: A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12 (1984) 387-395.

Edens, ,L. E., Bom, J., Ledeboer, A. M., Maat, J., Toonen, M. Y., Visser, C. and Verrips, C. T.: Synthesis and processing of the plant thaumatin in yeast. Cell 37 (1984) 629-633.

Fort, P., Marty, L., Piechaczyk, M., El Sabouty, S., Dani, C., Jeanteur, P. and Blanchard, J. M.: Various rat adult tissues express only one major mRNA species from the glyceraldehyde-3-phosphate-dehydrogenase multigenic family. Nucleic Acids Res. 13 (1985) 1431-1442.

Fothergill-Gillmore, L. A.: The evolution of the glycolytic pathways. Trends Biochem. Sci. 11 (1986) 47-51.

Gilbert, W., Marchionni, M. and McKnight, G.: On the antiquity of introns. Cell 46 (1986) 151-154.

Hanauer, A. and Mandel, J. L.: The glyceraldehyde 3 phosphate dehydrogenase gene family: structure of the human cDNA and of an X-chromosome linked pseudogene; amazing complexity of the gene family in mouse. EMBO J. 3 (1984) 2627-2633.

Harris, J. I. and Waters, M.: Glyceraldehyde 3 phosphate dehydrogenase. In Boyer, P. D. (Ed.), The Enzymes. Vol. 13C. Academic Press, New York, 1976, pp. 1-49.

Holland, J. P. and Holland, M. J.: Structural comparison of two nontandemly repeated yeast glyceraldehyde-3-phosphate dehydrogenase genes. J. Biol. Chem. 255 (1980) 2596-2605.

Hynes, M. J., Corrick, C. M. and King, J. A.: Isolation of genomic clones containing the *amdS* gene of *Aspergillus nidulans* and their use in the analysis of the structural and regulatory mutations. Mol. Cell. Biol. 3 (1983) 1430-1439.

Johanningmeier, U., Bodner, U. and Wildner, G. F. : A new mutation in the gene coding for the herbicide binding protein in *Chlamydomonas*. FEBS Lett. 211 (1987) 221-224.

Kinnaird, J. H. and Fincham, J. R. S.: The complete nucleotide sequence of the *Neurospora crassa am* (NADP-specific glutamate dehydrogenase) gene. Gene 26 (1983) 253-260.

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

Kozak, M.: Point mutations define a sequence flanking the AUG initiation codon that modulates translation by eukaryotic ribosomes. Cell 44 (1986) 283-292.

Krebs, E. G., Rafter, G. W. and Junge, J. M.: Yeast glyceraldehyde 3 phosphate dehydrogenase. Yeast protein, II. J. Biol. Chem. 200 (1953) 479-492.

Maniatis, T., Fritsch, E. F. and Sambrook, J.: Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.

Mattern, I. E., Unkles, S., Kinghorn, J. R., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Transformation of *Aspergillus oryzae* using the *A. niger pyrG* gene. Mol. Gen. Genet. 210 (1987) 460-461.

May, G. S., Tsang, M. L.-S., Smith, H., Fidel, S. and Morris, N. R.: *Aspergillus nidulans* β-tubulin genes are unusually divergent. Gene 55 (1987) 231-243.

McAlister, L. and Holland, M. J.: Differential expression of the three yeast glyceraldehyde-3-phosphate dehydrogense genes. J. Bacteriol. 260 (1985) 15019-15027.

McKnight, G. L., O'Hara, P. J. and Parker, M. L.: Nucleotide sequence of the triosephosphate isomerase gene from *Aspergillus nidulans*. Implications for a differential loss of introns. Cell 46 (1986) 143-147.

Michels, P. A. M., Poliszczak, A., Osinga, K. A., Misset, O., Van Beeumen, J., Wierenga, R. K., Borst, P. and Opperdoes, F. R.: Two tandemly linked identical genes for the glycosomal glyceraldehyde-phosphate dehydrogenase in *Trypanosoma brucei*. EMBO J. 5 (1986) 1049-1056.

Mullaney, E. J., Punt, P. J. and Van den Hondel, C. A. M. J. J.: DNA-mediated transformation of *Aspergillus ficuum*. Appl. Microbiol. Biotechnol. (1988) in press.

Orr, W. C. and Timberlake, W. E.: Clustering of spore specific genes in *Aspergillus nidulans*. Proc. Natl. Acad. Sci. USA 79 (1982) 5976-5980.

Penttilä, M., Nevalainen, H., Rättö, M., Salminer, E. and Knowles, J.: A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. Gene 61 (1987) 155-164.

Pichersky, E., Gottlieb, L. D. and Hess, J. F.: Nucleotide sequence of the triosephosphate isomerase gene of *Escherichia coli*. Mol. Gen. Genet. 195 (1984) 314-320.

Piechaczyk, M., Blanchard, J. M., Marty, L., Dani, C., Panabieres, F., El Sabouty, S., Fort, P., and Jeanteur, P.: Posttranscriptional regulation of glyceraldehyde-3-phosphate dehydrogenase gene expression in rat tissues. Nucleic Acids Res. 12 (1984) 6951-6963.

Platt, T.: Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. 55 (1986) 339-372.

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

Proudfoot, N. J. and Brownlee, G. G.: 3' Non-coding region sequences in eukaryotic messenger RNA. Nature 263 (1976) 211-214.

Punt, P. J., Oliver, R. P., Dingemanse, M. A., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. Gene 56 (1987) 117-124.

Saloheimo, M., Lehtovaara, P., Penttilä, M., Teeri, T. T., Ståhlberg, J., Johansson, G., Petterson, G., Claeyssens, M., Tomme, P. and Knowles, J. K. C.: EGIII, a new endoglucanase from *Trichoderma reesei*: the characterization of both gene and enzyme. Gene 63 (1988) 11-21.

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

Shih, M.-C., Lazar, G. and Goodman, H. M.: Evidende in favor of the symbiotic origin of chloroplasts: primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. Cell 47 (1986) 73-80.

Skarżynski, T., Moody, P. C. E. and Wonacott, A. J.: Structure of hologlyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 Å resolution. J. Mol. Biol 193 (1987) 171-187.

Stone, E. M., Rothblum, K. N., Alevy, M. C., Kuo, T. M. and Schwartz, R. J.: Complete sequence of the chicken glyceraldehyde-3-phosphate dehydrogenase gene. Proc. Natl. Acad. Sci. USA 82 (1985a) 1628-1632.

Stone, E. M., Rothblum, K. N., Schwartz, R. J.: Intron dependent evolution of chicken glyceraldehyde phosphate dehydrogenase gene. Nature 313 (1985b) 498-500.

Straus, D. and Gilbert, W.: Genetic engineering in the precambrian: structure of the chicken triosephosphate isomerase gene. Mol. Cell. Biol. 5 (1985) 3497-3506.

Tani, K., Singer-Sam, J., Munns, M. and Yoshida, A.: Molecular cloning and structure of an autosomal processed gene for human phosphoglycerate kinase. Gene 35 (1985) 11-18.

Teeri, T. T., Kumar, V., Lehtovaara, P. and Knowles, J.: Construction of cDNA libraries by blunt-end ligation: high-frequency cloning of long cDNA's from filamentous fungi. Anal. Biochem. 164 (1987) 60-67.

Tso, J. Y., Sun, X.-H., Kao, T.-H., Reece, K. S. and Wu, R.: Isolation and characterization of rat and human glyceraldehyde-3-phosphate dehydrogenase cDNA's: genomic complexity and molecular evolution of the gene. Nucleic Acids Res. 13 (1985a) 2485-2502.

Tso, J. Y., Sun, X.-H., and Wu, R.: Structure of two unlinked *Drosophila melanogaster* glyceraldehyde-3-phosphate dehydrogenade genes. J. Biol. Chem. 260 (1985b) 8220-8228.

Van Gorcom, R. F. M., Punt, P. J., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: A system for the analysis of expression signals in *Aspergillus*. Gene 48 (1986) 211-217.

Van Hartingsveldt, W., Mattern, I. E., Van Zeyl, C. M. J., Pouwels, P. H. and Van de Hondel, C. A. M. J. J.: Development of an homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. Mol. Gen. Genet. 206 (1987) 71-75.

Walker, J. E.., Carne, A. F. Runswick, M. J. Bridgen, J. and Harris, J. I.: D-Glyceraldehyde-3 phosphate dehydrogenase: complete amino acid sequence of the enzyme from *Bacillus stearothermophilus*. Eur. J. Biochem. 108 (1980a) 549-565.

Walker, J. E., Wonacott, A. J. and Harris, J. L.: D-Glyceraldehyde-3-phosphate dehydrogenase complete aminoacid sequence of the enzyme from *Thermus aquaticus*. Eur. J. Biochem. 10 (1980b) 581-586.

Ward, M. and Turner, G.: The ATP synthase subunit 9 gene of *Aspergillus nidulans*: sequence and transcription. Mol. Gen. Genet. 205 (1986) 331-338.

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

brough, P. O., Hayden, M. A., Dunn, L. A., Vermersch, P. S., Klass, M. R. and Hecht, R. M.: The glyceraldehyde-3-phosphate dehydrogenase gene family in the nematode *Caenorhabditis elegans*: solation and characterization of one of the genes. Biochim. Biophys. Acta 908 (1987) 21-33.

Gene, 56 (1987) 117-124

CHAPTER 3

Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*

(*A. niger; A. nidulans;* dominant selection; heterologous gene expression; *hph* and *gpd* genes; recombinant DNA)

Peter J. Punt[a], Richard P. Oliver[b], Maria A. Dingemanse[a], Peter H. Pouwels[a] and Cees A. M. J. J. van den Hondel[a]

[a]TNO Medical Biological Laboratory, 2280 AA Rijswijk (The Netherlands) and [b]School of Biological Sciences, University of East Anglia, Norwich NR47TJ (U. K.).

SUMMARY

A new, heterologous, dominant marker for selection of *Aspergillus* transformants is described. This marker is based on the *Escherichia coli* hygromycin B (HmB) phosphotransferase gene (*hph*). Expression of the *hph* gene is controlled by *A. nidulans gpd* and *trpC* signals. An *Aspergillus* transformation vector was constructed which contains this marker and confers HmB resistance to *Aspergillus* species.

With both *A. niger* and *A. nidulans*, transformation frequencies of 5-20 transformants per μg vector DNA were obtained. Cotransformation with other vectors was shown to be very efficient in both species, when selection for HmB resistance was applied.

INTRODUCTION

A number of selectable markers has been described for the transformation of the filamentous fungi *A. niger* and *A. nidulans*. In most cases selection of the transformants is based on complementation of mutations (for a review see Rambosek and Leach, 1987).

Only two dominant selectable markers (which do not require special mutant strains) are in use. One of these allows growth on acetamide or acrylamide as the sole nitrogen source (*amdS*; Kelly and Hynes, 1985), whereas the other confers

resistance to oligomycin (*oliC*; Ward et al., 1986).

The use of these two selection markers has certain limitations. For instance, the *amdS* marker can be used only in strains that have no requirements for nitrogen-containing compounds interfering with the *AmdS* selection. The *oliC* marker, based on an oligomycin-resistant ATP synthase subunit 9 gene of *A. nidulans*, is probably species-specific since it requires the formation of an functional oligomycin-resistance ATP synthase complex. To overcome these limitations, a dominant selectable marker for *Aspergillus* transformation was developed using the HmB-resistance gene (hph) of *E. coli.*

Hygromycin is an aminoglycosidic antibiotic that inhibits protein synthesis in prokaryotes and eukaryotes by interfering with translocation and causing misreading (Gonzalez et al., 1987; Singh et al., 1979), HmB-resistance genes have been isolated and characterized from *Streptomyces hygroscopicus* (Zalacain et al., 1986) and *E. coli* (Gritz and Davies, 1983, Kaster et al., 1983). Both genes code for an HmB phosphotransferase which inactivates the antibiotic by phosphorylation (Malpartida et al., 1983; Gritz and Davies, 1983).

Based on the *E. coli hph* gene, cloning vectors conferring HmB resistance have been contructed for *E. coli* (Gritz and Davies, 1983; Kaster et al., 1983), *Saccharomyces cerevisae* (Gritz and Davies, 1983; Kaster et al., 1984), filamentous fungi (Queener et al., 1985; Yoder et al., 1986), plant cells (Van den Elzen et al., 1985; Waldron et al., 1985) and animal cells (Bernard et al., 1985).

In this paper the construction of a vector conferring HmB resistance to *A. niger* and *A. nidulans* is described. This vector contains the *E. coli hph* gene fused to the 5' expression signals of the *A. nidulans gpd* gene (P. J. P. and C. A. M. J. J. v. d. H., to be published elsewhere; Van Gorcom et al., 1986) and the transcription-termination region of the *A. nidulans trpC* gene (Mullaney et al., 1985).

MATERIALS AND METHODS

(a) Strains and plasmids

E. coli K-12 JM109 (Yanisch-Perron et al., 1985) was used for propagation of vectors molecules. *A. niger* N402 (*cspA*1 derivative of ATCC9029; Bos, 1986) and *A. nidulans* FGSC4 (Glasgow wildtype; Clutterbuck, 1986) were used as recipient strains for transformation experiments. Cotransformation experiments were carried

ısing *A. niger* AB4.1 (*cspA*1, *pyrG*; Van Hartingsveldt et al., 1987) or *A. nidulans* 277 (*biA*1, *amdS*320, *amdI*18, *amdA*7, *niiA*4; Hynes et al., 1983) as recipient ns.

JC18 has been described by Yanisch-Perron et al. (1985); M13mp10 and mp11 have been described by Messing (1983). Plasmid pVU1005 containing the *li hph* gene was kindly provided by Dr. P. van den Elzen. Plasmid pHY101 was ined from Dr. W. E. Timberlake and has been described by Van Gorcom et al. 5). Plasmid p3SR2 was obtained from Dr. M. Hynes (Hynes et al., 1983). mid pAB4-1 has been described by Van Hartingsveldt et al. (1987). Vector l5-3 was obtained by subcloning of a 2.5-kb *Eco*RI fragment from a λ Charon4A e containing the *A. nidulans gpd* gene (P. J. P. and C. A. M. J. J. v. d. H., to ublished).

Materials

Glucuronidase was obtained from Sigma, Novozym 234 from NOVO Industries HmB from Calbiochem.

Transformation

coli transformations were carried out according to Hanahan (1983) and *rgillus* transformations according to the method of Yelton et al. (1984). DNA-ed protoplasts were plated on osmotically stabilized agar plates containing 20-nl minimal growth medium (Pontecorvo et al., 1953). After 16-20 h of incubation 0°C (*A. niger*) or 37°C (*A. nidulans*), the plates were overlaid with an equal me of minimal medium agar containing HmB. Transformants appeared after 2-ys of incubation at 30 and 37°C, respectively.

DNA manipulations

ligodeoxynucleotide-directed deletions were made following the method of Zoller Smith (1984). *Aspergillus* chromosomal DNA was isolated according to Yelton l. (1984). The dideoxy chaintermination method (Sanger et al., 1977) was used sequencing the deletion constructions. All other methods were carried out as

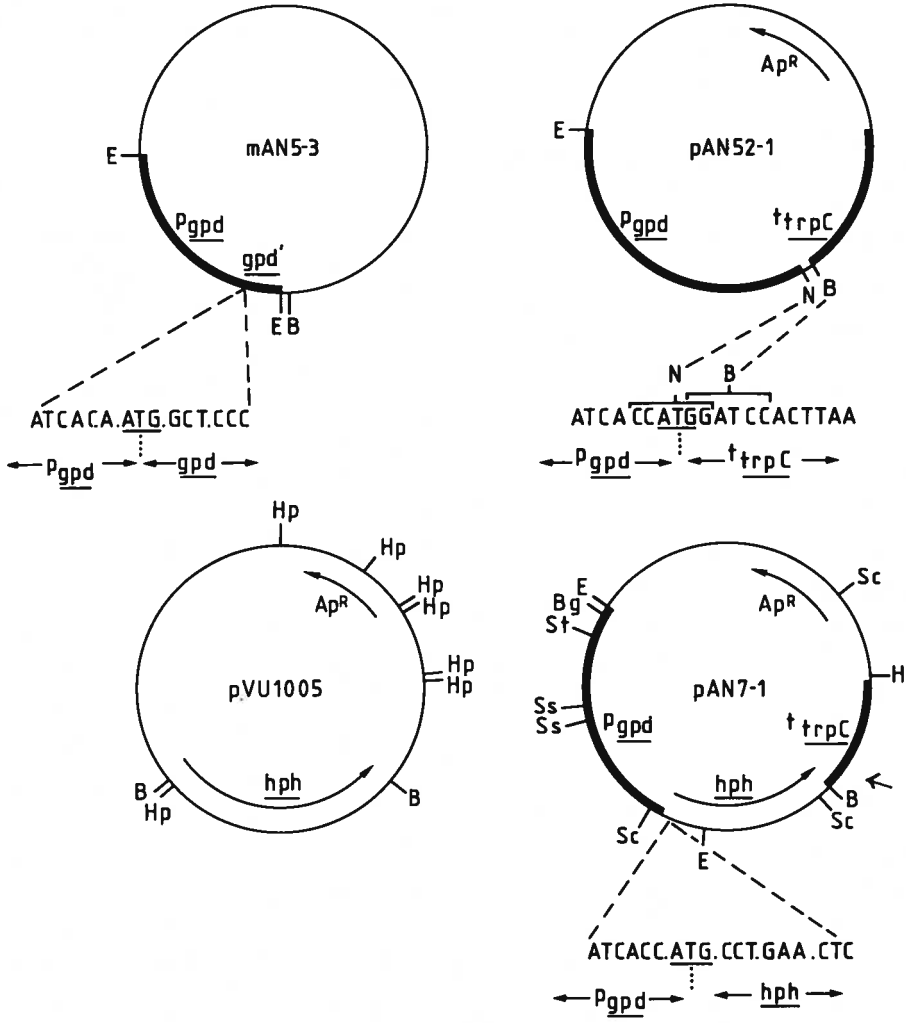
(a) Sensitivity to HmB of *Aspergillus niger* and *Aspergillus nidulans*

A prerequisite for the use of HmB resistance as a selection marker in *Asperg* transformation is the sensitivity of host strains to this drug. To test this, spores c *niger* and *A. nidulans* were plated on agar plates containing minimal medium different amounts of HmB. As shown in Table I growth of *A. niger* N402 was observed on plates containing 100 μg HmB/ml (or more). The sensitivity of th *nidulans* strains used was much lower. Even on plates containing 1000 μg Hml growth was not completely prevented although it was clearly reduced (not sho

(b) Construction of the *Aspergillus* vector pAN7-1, containing the *Escheri coli* HmB-resistance gene

The plasmids relevant for the construction of vector pAN7-1, that can be use transform *Aspergillus* to HmB resistance, are given in Fig. 1. The coding regio the *E. coli hph* gene, lacking the first three codons which are not necessary enzyme function (Kaster et al., 1983), has been fused at its N terminus to th expression signals, including the translation start codon (ATG) of the *A. nidulans* gene. Downstream from the coding region of the *hph* gene the termination re of the *A. nidulans trpC* gene (Mullaney et al., 1985) has been introduced. Detail the construction strategy are given in the legend to Fig. 1.

Fig. 1. Maps of relevant plasmids. Vector mAN5-3 was obtained by subcloning a 2.5-kb *E* fragment from a phage λCharon4A clone containing the *A. nidulans gpd* gene (P. J. P. and M. J. J. v. d. H., to be published). Sequence analysis of this subclone revealed that the transl start codon (ATG) is about 350 bp from the *Bam*HI site in mAN5-3. For the construction c *Aspergillus* expression vector all sequences between the ATG and the unique *Bam*HI in mA were deleted by oligodeoxynucleotide-directed mutagenesis. For convenience, at the same tim *Nco*I site was created around the ATG by changing an A at -1 into a C. The sequence o oligodeoxynucleotide used is CGATCTAGAGGATCCATGGTGATGTCTGC. From the resulting v a 1.8-kb *Eco*RI-*Bam*HI fragment containing the *gpd* expression signals was cloned into pL together with a 0.75-kb *Bam*HI-*Xho*I fragment containing the terminator region of the *A. nidulans*



gene (Mullaney et al., 1985), resulting in plasmid pAN52-1. Plasmid pVU1005 was cut with *Hph*I, treated with T4 polymerase and cut with *Bam*HI. A 1.0-kb *Hph*I-*Bam*HI fragment containing the *E. coli hph* gene (without the first three codons) was isolated and cloned into pAN52-1 (previously cut with *Nco*I, treated with T4 polymerase and cut with *Bam*HI) resulting in plasmid pAN7-1. Plasmid pAN7-1 does not contain *Eco*RV, *Kpn*I and *Sma*I sites. The nucleotide sequence around the ATG codon (underlined) in mAN5-3, pAN52-1 and pAN7-1 was confirmed by nucleotide sequence analysis. Thick lines represent *A. nidulans* DNA; thin lines represent *E. coli* DNA. B, *Bam*HI; Bg; *Bgl*II; E; *Eco*RI; H; *Hind*III; Hp, *Hph*I; N, *Nco*I; Sc, *Sca*I; Ss, *Sst*I; St, *Stu*I. p_{gpd} 5' expression signals of the *A. nidulans gpd* gene; t_{trpC}, termination region of the *A. nidulans trpC* gene; *gpd'*, 5' part of

TABLE I.
Analysis of *Aspergillus niger* transformants.

Transformant [plasmid]	Copy number[a]	Growth on agar plates with different concentrations of HmB (μg/ml)[b]				
		0	100	200	1000	2000
A. niger N402	0	+	-	-	-	-
[pAN7-1]1	1	+	+	+	+	(+)
[pAN7-1]2	1	+	+	+	-	-
[pAN7-1]3	2-3	+	+	+	+	(+)
[pAN7-1]4	1-2	+	+	+	+	(+)
[pAN7-1]5	5	+	+	+	+	(+)
[pAN7-1]6	3-5	+	+	+	+	(+)
[pAN7-1]7	1-2	+	+	+	+	(+)
[pAN7-1]8	2	+	+	+	+	(+)
[pAN7-1]9	1	+	+	+	(+)	-
[pAN7-1]10	1-2	+	+	+	(+)	-
[pVU1005]1	1	+	(+)	(+)	-	-
[pVU1005]3	1	+	+	+	-	-
[pVU1005]4	1	+	+	+	(+)	-

[a] Spotblots containing dilutions of chromosomal DNA of the transformants (ranging from 0.2 to 2μg) were probed with a ^{32}P-labelled *Bam*HI fragment from pVU1005 containing the *hph* gene or a ^{32}P-labelled *Xba*I fragment from plasmid pAB4-1 (Van Hartingsveldt et al., 1987) containing the *A. niger pyrG* gene. Signals obtained with the latter probe were used to determine the exact amount of transformant DNA spotted. From a comparison of the hybridization signals of the *hph* probe obtained with the spotted amount of transformant DNA and fixed amounts of pAN7-1 DNA the copy number was calculated.

[b] *A. niger* transformants were stab-inoculated on minimal-medium plates containing different concentrations of HmB. Growth was scored after three to four days of incubation at 30°C; +, normal growth; (+), clearly reduced growth; -, no growth.

(c) Transformation of *Aspergillus*

Plasmid pAN7-1 was used to transform *A. niger* and *A. nidulans* protoplasts. For selection of HmB^R transformants of *A. niger* an agar overlay containing 200 μg HmB/ml was sufficient to prevent growth of untransformed colonies. They could also be selected for by plating DNA-treated protoplasts directly on agar plates containing 100 μg HmB/ml. In the transformation experiments with *A. nidulans* protoplasts an overlay with at least 1000 μg HmB/ml was required to select HmB^R transformants. As shown in Table II, both fungi could be transformed to HmB resistance with pAN7-1 at a frequency of 5-20 transformants per μg DNA. HmB^R colonies were not observed on plates with pUC18-treated protoplasts. pVU1005 containing the *E. coli hph* gene without *A. nidulans* expression signals flanking the *hph* gene was also used to transform *A. niger*. Very few HmB^R colonies were found in this case (Table II).

(d) HmB-resistance levels of the *Aspergillus* transformants

To determine the level of HmB resistance in a number of transformants, they were stab-inoculated on minimal medium agar containing different concentrations of HmB. Most of the *A. niger* transformants obtained with pAN7-1 grew readily on plates containing 1000 μg HmB/ml, wheras a small number (three out of ten) only grew well on plates with 200 μg HmB/ml (Table I). Growth on plates with more than 1000 μg HmB/ml was clearly reduced for all *A. niger* transformants tested. The observed differences in resistance might be due to differences in the number of copies of pAN7-1 in the transformants. However, since in all transformants the vector DNA was integrated into the chromosomal DNA (see next paragraph) differences in resistance may also be due to differences in chromosomal environment of the integrated pAN7-1 sequences.

The *A. niger* transformants obtained with pVU1005 did not grow normally on plates containing more than 200 μg HmB/ml (Table I). The difference is resistance level between pVU1005 transformants and pAN7-1 transformants probably reflects a difference in efficiency of expression of the *hph* gene. In pAN7-1 transformants the expression of the *hph* gene is controlled by (strong) *A. nidulans* expression signals, which were shown to be functional in *A. niger* (Van Gorcom et al., 1986). Since

transformants obtained with this vector is probably dependent on integration of the plasmid into the genome in such a way that the *hph* gene is fused to *A. niger* expression signals. Integration of the *hph* gene close to strong expression signals is likely to be very infrequent.

For *A. nidulans* pAN7-1 transformants differences in growth rate were observed on plates containing 1000 μg HmB/ml (not shown). This might also be due to differences in number of copies or in chromosomal environment of the vector in the transformants.

(e) Spotblot and Southern analysis of HmB-resistant transformants

To obtain information about the number of plasmid copies present in the transformants and as to whether the plasmids are integrated into the chromosomal

TABLE II.
Transformation of *Aspergillus niger* and *Aspergillus nidulans*[a].

Strains	Selected marker (vector)	Cotransformed marker (vector)	Transformation frequency[b]	Cotransformation frequency[c]
A. niger N402	*hph* (pAN7-1)	-	10	-
A. niger N402	*hph* (pVU1005)	-	0.4	-
A. niger AB4.1	*hph* (pAN7-1)	-	15	-
A. niger AB4.1	*hph* (pAN7-1)	*pyrG* (pAB4-1)	20	15/18
A. nidulans FGSC4	*hph* (pAN7-1)	-	5	-
A. nidulans MH1277	*hph* (pAN7-1)	-	15	-
A. nidulans MH1277	*hph* (pAN7-1)	*amdS* (p3SR2)	10	9/16

[a] 5-10 μg DNA was used per transformation. In cotransformation experiments equimolar amounts of both plasmids were used. (For details see MATERIALS and METHODS, section c.)
[b] HmB^R transformants per μg DNA.
[c] A number of HmB^R transformants was stab-inoculated on plates selective for the second marker. The number of cotransformants divided by the total number analysed are given.

IA, spotblot and Southern analyses were carried out. As is shown in Table I, the *niger* transformants analysed contained different numbers of copies of pAN7-1. Southern blot analysis of chromosomal DNA of the transformants cut by *Bam*HI ealed that single and/or tandem copies of the plasmid were integrated into the nome at various sites (results not shown). Similar results were obtrained from uthern analyses of *A. nidulans* transformants (not shown).

All three pVU1005 transformants of *A. niger* analysed contained only one copy of plasmid sequences (Table I). DNA from non-transformed strains did not show y hybridization with the *hph* probe.

Cotransformation

To analyse the possibility of cotransformation with vector pAN7-1, auxotrophic itants of *A. niger* and *A. nidulans* were transformed with a mixture of pAN7-1 and vector that contains the gene that complements the mutation. *A. niger* AB4.1 (a G mutant) was transformed with pAN7-1 and pAB4-1, in a 1:1 molar ratio, *A. lulans* MH1277 (an *amdS* mutant) was transformed with pAN7-1 and p3SR2. HmB^R transformants were obtained without selection for the auxotrophic marker at quencies similar to those obtained for wild-type strains. These HmB^R transformants re subsequently stab-inoculated on media selective for the second marker (*pyrG amdS*). As shown in Table II, about 80% of the HmB^R transformants of *A. niger* 4.1 and 60% of the HmB^R of *A. nidulans* MH1277 were cotransformed with the cond marker.

Conclusions

A. niger and *A. nidulans* can be transformed with the vector pAN7-1 containing *hph* gene of *E. coli*. Expression of the *hph* gene confers resistance to HmB in th species. Since the sensitivity levels for the drug in both species were clearly erent, different concentrations of HmB were used for the selection of HmB^R nsformants.

The transformation frequencies obtained in *A. niger* and *A. nidulans* are mparable. Since expression of the *hph* gene on plasmid pAN7-1 is controlled by *A. nidulans gpd* and *trpC* expression signals, this indicates that these

several other *A. nidulans* and *A. niger* expression signals (Kos et al., 1985; Kelly ar Hynes, 1985; Buxton et al., 1985; Van Gorcom et al., 1986; Van Hartingsveldt et a 1987).

Analysis of HmB^R transformants showed that the vector molecules integrate in chromosomal DNA, as was found in virtually all other *Aspergillus* transformati systems described (Rambosek and Leach, 1987). In a high percentage of t transformants more than one copy of pAN7-1 is integrated into the genome various locations). In general, transformants containing more than one copy of pAN 1 in the genome show a higher resistance level than those containing only one co (Table I). This result suggests a gene-dose effect for *hph* expression. Howev some transformants with only one copy of pAN7-1 integrated into the genome sh a high level of resistance (Table I) as well, suggesting that the site of integration m also influence the expression of the *hph* gene. Therefore a more likely explanati for these observations could be that integration of one (or more) copy of pAN7 in the genome results in a high HmB resistance level but that in a few cases t chromosomal environment of the integrated copy causes a decrease in expressi of the *hph* gene.

Cotransformation of vector molecules in combination with pAN7-1 is very efficie for both *Aspergillus* species (Table II). Cotransformation frequencies of more th 50% are obtained by transformation with equimolar amounts of pAN7-1 and second vector. These efficiencies are comparable to those found for other selecti systems (Wernars, 1986; Kelly and Hynes, 1985) and can be very useful introducing non-selectable genes into host strains without the need to constr special vectors and strains.

As shown in Table II, *A. niger* transformants can be obtained with pVU1005. Sin there are no *Aspergillus* expression signals on this vector, the *hph* gene is probal integrated behind *Aspergillus* expression signals which cause *hph* expression. In t way pVU1005 can be used as a promoter probe vector. Similar vectors have be used in *E. coli* and *S. cerevisiae* to isolate (strong) expression signals (e. Casadaban et al., 1983).

Since many other ascomycetes and deuteromycetes are related to *A. nidulans A. niger*, the use of *A. nidulans* and *A. niger* vectors in the transformation of thc fungi seems promising and has already been proven fruitful in some cases (Turge et al., 1985; P. J. P. and C. A. M. J. J. v. d. H., unpublished observations). Sir HmB resistance conferred by vector pAN7-1 is based on inactivation of the dr

species specificity of drug resistance is not expected for this marker. This would mean that vector pAN7-1 could be used as a transformation vector for any HmB-sensitive filamentous fungus. Indeed, vector pAN7-1 has been used successfully in transforming *Fulvia fulvum*, a plant-pathogenic deuteromycete (R. P. O., unpublished results), and *Schizophyllum commune*, a basidiomycete (J. H. S., unpublished results).

In conclusion, transformation based on HmB resistance might be a useful tool in molecular genetic research in *Aspergillus* and other filamentous fungi.

ACKNOWLEDGEMENTS

The authors wish to thank R. F. M. van Gorcom and J. F. Bleichrodt for critical reading of the manuscript and J. H. Sietsma and R. P. van Weeghel for providing results from their experiments before publication. This work was made possible by a travel grant to R. P. O. from EMBO (ASTF 4888).

REFERENCES

Bernard, H.-U., Krammer, G. and Rowekamp, W. G.: Construction of a fusion gene that confers resistance against hygromycin B to mammalian cells in culture. Exp. Cell. Res. 158 (1985) 237-243.

Bos, C. J.: Induced Mutation and Somatic Recombination as Tools for Genetic Analysis and Breeding of Imperfect Fungi. Ph. D. Thesis Agricultural University, Wageningen, The Netherlands, 1985.

Buxton, F. P., Gwynne, D. I. and Davies, R. W.: Transformation of *Aspergillus niger* using the *arg*B gene of *Aspergillus nidulans*. Gene 37 (1985) 207-214.

Casadaban, M. J., Martinez-Arias, A., Shapira, S. K. and Chou, J.: β-Galactosidase gene fusions for analysing gene expression in *Escherichia coli* and yeast. Methods Enzymol. 100 (1983) 293-308.

Clutterbuck, A. J.: Glasgow stock list of *Aspergillus nidulans* strains 1986. Fungal Genet. Newslett. 33 (1986) 59-59.

Gonzalez, A., Jimenez, A., Vasquez, D., Davies, J. E. and Schindler, D.: Studies on the mode of actions of hygromycin B, an inhibitor of translocation in eukaryotes. Biochim. Biophys. Acta 521 (1987) 459-469.

Gritz, L. and Davies, J.: Plasmid encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. Gene 25 (1983) 557-580.

Hanahan, D.: Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol 166 (1983) 557-580.

Hynes, M. J., Corrick., C. M. and King, J. A.: Isolation of genomic clones containing the *amdS* gene of *Aspergillus nidulans* and their use in the analysis of the structural and regulatory mutations. Mol. Cell. Biol. 3 (1983) 1430-1439.

Kaster, K. R., Burgett, S. G., Rao, R. N. and Ingolia, T. D.: Analysis of a bacterial hygromycin B resistance gene by transcriptional and translational fusions and by DNA sequencing. Nucl. Acids Res. 11 (1983) 6895-6911.

Kaster, K. R., Burgett, S. G. and Ingolia, T. D.: Hygromycin B resistance as dominant selectable marker in yeast. Curr. Genet. 8 (1984) 353-358.

Kelly, J. M. and Hynes, M. J.: Transformation of *Aspergillus niger* by the *amdS* gene of *Aspergillus nidulans*. EMBO J. 4 (1985) 475-479.

Kos, A., Kuijvenhoven, J., Wernars, K., Bos, C. J., Van den Broek, H. W. J., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Isolation and characterization of the *Aspergillus niger trpC* gene. Gene 39 (1985) 231-238.

Maniatis, T., Fritsch, E. F. and Sambrook, J.: Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.

Malpartida, F., Zalacain, M., Jimenez, A. and Davies, J.: Molecular cloning and expression in *Streptomyces lividans* of a hygromycin B phosphotransferase gene from *Streptomyces hygroscopicus*. Biochem. Biophys. Res. Commun. 117 (1983) 6-12.

Messing, J.: New M13 vectors for cloning. Methods Enzymol. 101 (1983) 20-79.

Mullaney, E. J., Hamer, J. E., Yelton, M. M. and Timberlake, W. E.: Primary structure of the *trpC* gene from *Aspergillus nidulans*. Mol. Gen. Genet. 199 (1985) 37-45.

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

Queener, S. W., Ingolia, T. D., Skatrud, P. L., Chapman, J. L. and Kaster, K. R.: A system for genetic transformations of *Cephalosporium acremonium*. In: Schlessinger, D. (Ed.), Microbiology, Am. Soc. Microbiol. (1985) 468-472.

Rambosek, J. A. and Leach, J.: Recombinant DNA in filamentous fungi; progress and prospects. CRC Crit. Rev. Biotechnol. (1987) in press.

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

Singh, A., Ursic, D. and Davies, J.: Phenotypic suppression and misreading in *Saccharomyces cerevisiae*. Nature 277 (1979) 146-148.

Turgeon, B. G., Garber, R. C. and Yoder, O. C.: Transformation of the fungal maize pathogen *Cochliobolus heterostrophus* using the *Aspergillus nidulans amdS* gene. Mol. Gen. Genet. 201 (1985) 450-453.

Van den Elzen, P. J. M., Townsend, J., Lee, K. Y. and Bedbrook, J. R.: A chimaeric hygromycin resistance gene as a selectable marker in plant cells. Plant Mol. Biol. 5 (1985) 299-302.

Van Gorcom, R. F. M., Pouwels, P. H., Goosen, T., Visser, J., Van den Broek, H. W. J., Hamer, J. E., Timberlake, W. E. and Van den Hondel, C. A. M. J. J.: Expression of an *Escherichia coli* β-galactosidase fusion gene in *Aspergillus nidulans*. Gene 40 (1985) 99-106.

an Gorcom, R. F. M., Punt, P. J., Pouwels, P. H. en Van den Hondel, C. A. M. J. J.: A system for the analysis of expression signals in *Aspergillus*. Gene 48 (1986) 211-217.

an Hartingsveldt, W., Mattern, I. E., Van Zeijl, C. M. J., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Development of an homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. Mol. Gen. Genet. 206 (1987) 71-75.

'aldron, C., Murphy, E. B., Roberts, J. L., Gustafson, G. D., Armour, S. L. and Malcolm, S. K.: Resistance to hygromycin B. A new marker for plant transformation studies. Plant Mol. Biol. 5 (1985) 103-108.

'ard, M., Wilkinson, B. and Turner, G.: Transformation of *Aspergillus nidulans* with a cloned, oligomycin-resistant ATP synthase subunit 9 gene. Mol. Gen. Genet. 202 (1986) 265-270.

'ernars, K.: DNA-mediated Transformations of the Filamentous Fungus *Aspergillus nidulans*. Ph. D. Thesis, Agricultural University, Wageningen, The Netherlands, 1986.

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

elton, M. M., Hamer, J. E. and Timberlake , W. E.: Transformation of *Aspergillus nidulans* using a *trpC* plasmid. Proc. Natl. Acad. Sci. USA 81 (1984) 1470-1474.

oder, O. C., Weltring, K., Turgeon, B. G., Garber, R. C. and Van Etten, H. D.: Technology for molecular cloning of fungal virulence genes. In: Bailey, J. A. (Ed.), Biology and Molecular Biology of Plant-Pathogen Interactions. NATO ASI Series H, Cell Biology, Vol. 1, 1986, pp. 371-384.

alacain, M., Gonzalez, A., Guerrero, M. C., Mattaliano, R. J., Malpartida, F. and Jimenez, A.: Nucelotide sequences of hygromycin B phosphotransferase gene of *Streptomyces hygroscopicus*. Nucl. Acids. Res. 14 (1986) 1565-1581.

oller, M. J. and Smith, M.: Oligonucleotide-directed mutagenesis: A simple method using two oligonucleotide primers and a single stranded template. DNA 3 (1984) 479-488.

Gene, 93 (1990) 101-109

CHAPTER 4

Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene encoding glyceraldehyde-3-phosphate dehydrogenase

(Recombinant DNA; *lacZ* fusion gene; site-directed mutagenesis; targeted single-copy integration; *Aspergillus niger*; transcription activation; transcription start point)

Peter J. Punt, Maria A. Dingemanse, Anneke Kuyvenhoven, Ronald D. M. Soede, Peter H. Pouwels and Cees A. M. J. J. Van den Hondel

TNO Medical Biological Laboratory, Rijswijk (The Netherlands)

SUMMARY

Analysis of the promoter region of the highly expressed *Aspergillus nidulans gpdA* gene is described. The nucleotide (nt) sequence of a 1.3-kb region upstream of the ATG was determined. Comparison with promoter regions of other *Aspergillus* and *Neurospora* genes revealed several regions of similar sequence. Both random and site-specific mutations were introduced into the promoter region of the *gpdA* gene, and the resulting mutant promoters were fused to the *Escherichia coli lacZ* gene. The constructed fusions were introduced into *A. nidulans* and transformants that contained one copy of these fusions at the *argB* locus were analysed. β-Galactosidase assays and primer extension experiments were used to identify sequence elements involved in transcription activation and transcription initiation. Two elements, located around 650 and 250 nt upstream of the major transcription start point (*tsp*), were identified as transcription activation elements. These elements coincide with regions of putative secondary structure (direct or inverted-repeats). A third element, a C+T-rich region directly upstream from the major *tsp*, was shown to be involved in correct initiation of transcription.

Filamentous fungi are attractive organisms for studying structure-function relationships of genetic (chromosomal and mitochondrial) elements, because, being lower eukaryotes, they have a complex life cycle analogous to that of higher eukaryotes. Yet they can be cultivated in simple, defined growth media. Moreover, they can be analysed by simple biochemical and (molecular and classical) genetic means.

Despite the extensive data that have been accumulated about the primary structure of the 5'- and 3'-flanking regions of a number of fungal genes and the many sequence elements that have been indicated (for reviews see Ballance, 1986; Gurr et al., 1988), very little is known about the functional organisation of the expression signals of filamentous fungi. Only for a few genes an (even limited) functional analysis of the promoter region has been carried out (Hamer and Timberlake, 1987; Davis et al., 1988; Goc and Weglenski, 1988; Soliday et al., 1989).

Much more is known about the organisation of the expression signals of *Saccharomyces cerevisiae* and higher eukaryotes. A detailed analysis of elements involved in both transcription activation and transcription initiation has been carried out for a number of genes. Several elements, such as UAS, TATA box etc., have been functionally identified (for reviews see Struhl, 1986, 1987; Wasylik, 1986; Guarente, 1988).

To identify and functionally characterize regulatory elements for filamentous fungi, a detailed functional analysis is needed. Since in various organisms, including *S. cerevisiae*, the use of *lacZ* fusions has been very fruitful for the functional analysis of expression signals (e. g., Struhl, 1986), a similar strategy was developed for filamentous fungi (Van Gorcom et al., 1986). This strategy was used to analyse the expression signals of the GPD-encoding gene of *A. nidulans*.

The GPD-encoding gene of *A. nidulans* (*gpdA*) including its expression signals has been isolated and sequenced recently (Punt et al., 1988). GPD is a key enzyme in glycolysis and gluconeogenesis. In several organisms, such as *S. cerevisiae* (Krebs et al., 1953) and in higher eukaryotes (Piechaczyk et al., 1984) up to 5% of the soluble cellular proteins consists of GPD. Similar results were also obtained for *A. nidulans* (P. P. F. Hanegraaff, in preparation). This indicates that the expression signals of the *gpdA* gene(s) are very strong. Therefore, results obtained from the functional analysis of the *gpdA* expression signals will probably encourage the use

these expression signals for expression of homologous and heterologous genes fungal host strains.

ATERIALS AND METHODS

Strains and plasmids

The *A. nidulans* ArgB⁻ strain (*argB*2, *biA*1, *methG*2) used as a recipient in nsformation experiments was obtained from Dr. W. E. Timberlake (Univ. of orgia, Athens, GA). Vectors were constructed and propagated in *E. coli* K-12 109 (Yanisch-Perron et al., 1985). Uracil-containing DNA templates were prepared *E. coli* K-12 BW313 (Kunkel, 1985). M13mp18 has been described by Yanisch-rron et al. (1985), plasmids pAN923-41B/42B and pAN5-41B by Van Gorcom et (1986) and pAN52-1 by Punt et al. (1987). M13mAN5-7B was obtained by ocloning of an 0.34-kb *Sst*I fragment containing the 3' end of the coding region the *gpdA* gene (Punt et al., 1988) into M13mp11 (Messing, 1983).

Transformations

E. coli transformations were carried out according to Hanahan (1983) and M13 nsfections according to Kunkel (1985). *A. nidulans* transformations were carried t according to Yelton et al. (1984) or Goosen et al. (1987).

Molecular methods

Fungal DNA and RNA isolations were performed as described by Kolar et al. 988) and Teeri et al. (1987), respectively. If not indicated, all other methods were sentially as described in Maniatis et al. (1982).

SULTS AND DISCUSSION

Sequence analysis of the *gpdA* promoter

Combination of the data of the sequence analysis of different subclones of the omoter region of the *A. nidulans gpdA* gene resulted in the complete nt sequence

found compared to the sequence previously published (Punt et al., 1988). Tl changes are indicated in Fig. 1.

Comparison of the 5' flanking sequences of the *A. nidulans gpdA* gene with th of 5' flanking regions of other glycolytic *Aspergillus* genes (*pgkA*; Clements a Roberts, 1986; and *tpiA*; McKnight et al., 1986) indicated two regions of simi sequence (Punt et al., 1988), namely, a region around 600 nt upstream from the *ts* further referred to as *pgk* box and the C+T-rich region immediately upstream of t major *tsp*, referred to as *ct* box (Fig. 1).

Recently, the complete sequence of the *A. niger gpdA* gene was also establish in our laboratory (R. F. M. Van Gorcom, unpublished results). Although the ent promoter regions of the two *gpdA* genes show about 70% similarity, a region arou 250 nt upstream of the major *tsp* in the *A. nidulans gpdA* promoter shows as mu as 96% similarity with a region of the *A. niger gpdA* gene located at a simi distance from the *tsp* (Fig. 2). This region is further referred to as the *gpd* box (F 1). Both the *gpd* box and the *pgk* box are found in regions with a considerat number of inverted-repeat sequences (Figs. 1, 4D). Since inverted-repeat sequenc are considered to be regulatory sites in promoter sequences through binding transcription factors and/or regulatory proteins (Johnson and McKnight, 1989), th result may indicate that the *pgk* and *gpd* boxes are a part of such regulatory site

Elements similar to the ones identified in the *A. nidulans alcA* promoter (Gwyn et al., 1987) and in the promoter regions of genes in the *qa* gene cluster *Neurospora crassa* (Baum et al., 1987; Geever et al., 1989) and the *qut* gene clust of *A. nidulans* (Hawkins et al., 1989) were also found in the *gpdA* promoter (Fig. These elements are indicated as *alc*, *qa*, and *qut* boxes, respectively, in Fig. 1. T significance of these boxes for transcription activation and/or regulation in the *gp* promoter remains to be determined. However, it is interesting to note that most these boxes coincide with regions of direct- or inverted-repeat sequence suggesting that they are also involved in binding of regulatory proteins and/ transcription factors. One area in the promoter region around 600 nt upstream fro the major *tsp* draws special attention, since *pgk*, *qa* and *qut* boxes partially overl each other there.

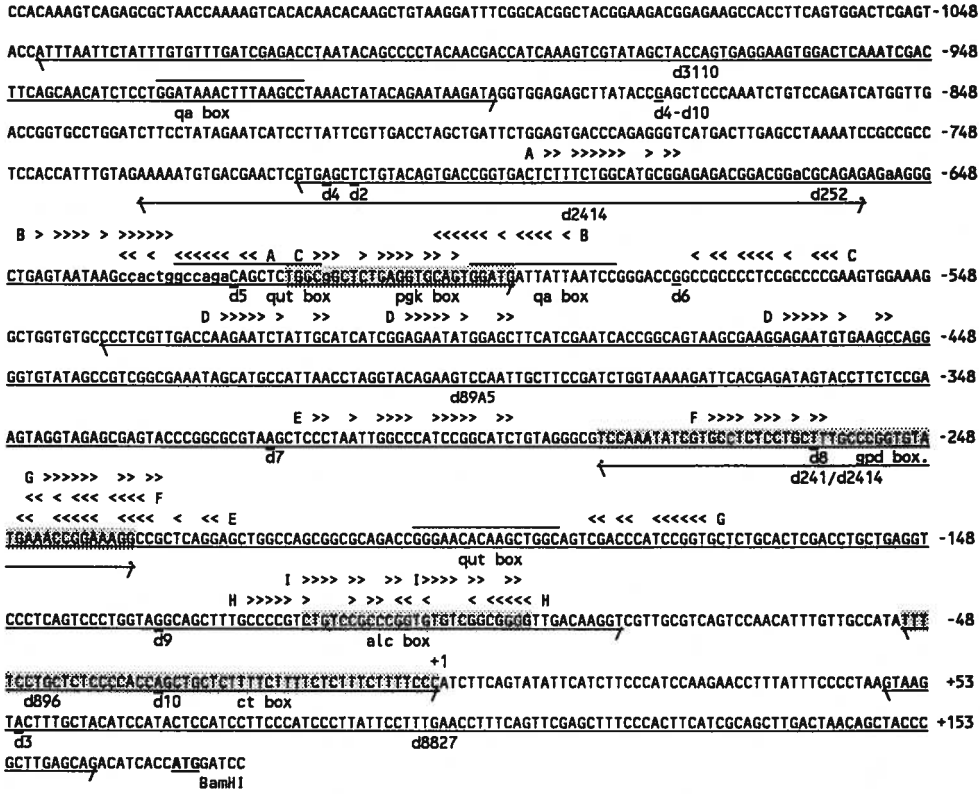


Fig. 1. Nucleotide sequence of the 5' flanking region of the *A. nidulans gpdA* gene in pAN52-1 (Punt et al., 1987). Sequence analysis was carried out using chain termination methods (Chen and Seeburg, 1985; Sanger et al., 1977). Nucleotides are numbered with reference to the major *tsp* (+1). Corrections of previously published sequence data (Punt et al., 1988) are indicated in lower-case letters. The start codon (ATG) is given in bold underlined letters. Specific deletions are underlined with lines flanked by arrowheads. The deletion endpoints of unilateral deletions are underlined. All deletions are indicated with the appropriate deletion (d) numbers. Elements identified by comparison of the 5'-flanking sequences of the *gpdA* gene with the 5'-flanking sequences of other fungal genes are indicated by shaded letters (*pgk*-, *gpd*-, *alc*- and *ct* box, see RESULTS AND DISCUSSION section **a**) or overlining (*qa* and *qut* box). Regions with inverted- or direct-repeat sequences are indicated by > or < symbols. Corresponding regions are indicated with pairs of identical upper-

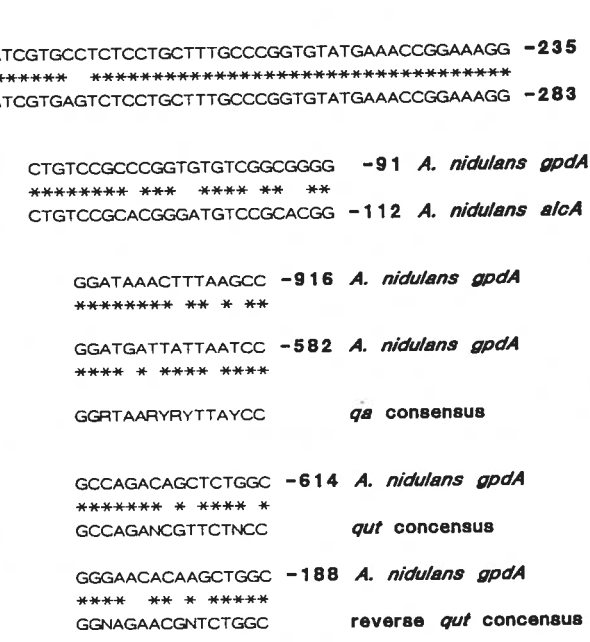


Fig. 2. Sequence elements identified by comparison of 5'-flanking regions of the *A. nidulans gpdA* gene and other fungal genes. **(A)** Nucleotide sequence of a region of similar sequence in the 5' flanking regions of the *gpdA* genes of *A. nidulans* and *A. niger*. **(B)** Sequence element similar to a direct repeat sequence identified in the *alcA* control region (Gwynne et al., 1987). **(C)** Sequence elements similar to the binding site of the regulatory protein QA-1F of *N. crassa* (Geever et al., 1989). **(D)** Sequence elements similar to the 16-nt motif in the *qut* gene cluster (Hawkins et al., 1988). Identical nt are indicated by asterisks. The distance (in nt) from the major *tsp* is given (not for *qa* and *qut* consensus elements).

(b) Construction of vectors containing promoter deletion mutations

To identify sequences in the 5'-flanking regions of the *gpdA* gene important for efficient gene expression we have constructed a set of promoter deletion mutants and analysed the effect of the deletions on transcription efficiency by determining the expression of the *E. coli lacZ* gene, which was fused to the mutant promoters. Initially, vectors pAN923-41B or pAN923-42B (Van Gorcom et al., 1986; Fig. 3) were used for the construction of fusion genes. To select for integration of the fusion gene at the *argB* gene locus (Hamer and Timberlake, 1987) these vectors were modified

filling in the unique *Bgl*II site in the *argB* gene with PolIK, creating a mutant *argB* ele. The resulting vectors pAN923-41B$_{Bgl\,II}$ and pAN923-42B$_{Bgl\,II}$ yield a reduced nsformation frequency (less than one transformant per μg of DNA), but a large (at least 50%) of the transformants obtained with this vector contain a single copy the plasmid integrated at the *argB* locus.

The location and extent of the different promoter deletions described below is licated in Fig. 4.

At first, a number of deletion mutants was constructed to localize promoter quences in the 2.1-kb *Stu*I fragment of pAN5-d0 (=pAN5-41B) (Van Gorcom et al., 86), resulting in pAN5-d2 and pAN5-d3 (Fig. 4A). In each of these vectors, gether with the promoter, the 5' part of the coding region of the *gpdA* gene was ed to the *E. coli lacZ* gene. Subsequently, a number of promoter mutations was nstructed whereby only the promoter plus 5'-noncoding sequences of the *gpdA* ne were fused to the *lacZ* gene, resulting in pAN5-d1 and pAN5-d4 to pAN5-0 (Fig. 4B).

Several of the elements identified by sequence comparison (RESULTS AND SCUSSION, section **a**; Fig. 1) were removed by site-directed mutagenesis using gos spanning the region to be deleted. Thus pAN5-d241, pAN5-d252, pAN5-110, pAN5-d8827 and pAN5-d896 were obtained (Fig. 4C).

n oligo-directed mutagenesis experiments as described above, aberrant mutants re obtained at a low frequency. Vectors pAN5-d2414 and pAN5-d89A5 were tained by fusing such aberrant mutant promoters to *lacZ* in pAN923-42B$_{Bgl\,II}$ (Fig.).

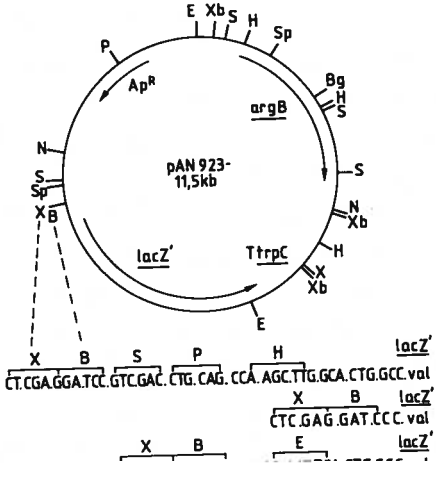


Fig. 3. Restriction map of the vectors used for the analysis of expression signals. Promoter fragments (including ATG) can be cloned in three different reading frames in front of *lacZ*. The construction of these vectors is described by Van Gorcom et al. (1986). Vectors pAN923-41B$_{Bgl\,II}$ and pAN923-42B$_{Bgl\,II}$ were obtained from pAN923-41B and -42B by filling-in (with PolIK) the unique *Bgl*II site. Thick lines represent *Aspergillus* DNA, thin lines represent *E. coli* DNA. B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; N, *Nru*I; P, *Pst*I; S, *Sal*I; Sp, *Sph*I; X, *Xho*I; Xb, *Xba*I; t_{trpC}, transcription termination

The deletion endpoints of mutant promoters obtained by site-directed mutagene were determined by sequence analysis and found to be as was expected from t sequence of the oligos used for the mutagenesis. Sequence analysis of one of t aberrant deletions (pAN5-d89A5) revealed the presence of inverted-repeat sequenc at the endpoints of the deletion. This suggests that secondary structure in t template used for mutagenesis may be a source of aberrant mutations.

(c) Construction of the *Aspergillus nidulans* strains

Vectors with different promoter deletions were used to transform *A. nidula* ArgB$^-$. Southern analysis was used to identify *A. nidulans* transformants containi a single copy of the fusion gene at the *argB* locus (results not shown). One or t independently isolated, single-copy transformants of each mutant promoter we used for further analysis.

Upon visual inspection of transformants on XGal plates we consistently observ both blue and white colonies. Even with the use of the mutant *argB* allele fr pAN923-42B$_{BglII}$ white colonies were obtained. Southern analysis showed that the white transformants had arisen from a gene conversion or double-crossover eve This type of integration event has been found in other systems as well (Fincha 1989).

(d) βGal activities

βGal activities were measured for all deletion mutants in protein extracts prepar from cells grown in minimal medium. The results described in section **e**, below, sh that identical transcripts are produced in the different promoter mutants (except and d896). Therefore, βGal activity levels may be used as a measure for transcripti efficiency of all mutants, except d3 and d896.

From the relative activity levels presented in Fig. 4A, it can be concluded that t major expression signals are located within an 0.8-kb *Sst*I-*Sca*I fragment (compa d0 with d2 in Fig. 4A). In d3 only 5% of the activity of d0 is found, indicating th only a very low promoter activity is left, which probably originates from vec sequences upstream from the promoter region. Further analysis was focussed the region located between the upstream *Sst*I site and the start codon (AT Deletion of a region upstream of the 5' *Sst*I site (d3110; Fig. 4C) has no or only v limited effect on expression levels, in agreement with the previous conclusion t

the signals for efficient transcription are present within a region of about 0.9 kb upstream of the ATG codon. The effect of the different Bal31 deletions on βGal activity is shown in Fig. 4B.

From the unilateral deletions d4 to d10 it is clear that sequences within 260 bp immediately upstream from the major *tsp* (see section **e**, below) result in wt promoter activity (compare d1 and d8; Fig. 4B). Deletions extending into this 260-bp fragment (i. e., d9 and d10) result in a considerable decrease in promoter activity indicating the presence of positive promoter elements in this region. Compared to d1 and d8, d5 shows a 20-30% decrease in activity, suggesting the presence of one or more negative promoter elements between (and/or around) the 3'-endpoints of the deletions in pAN5-d5 and pAN5-d8, and the presence of one or more positive promoter elements between (and/or around) the upstream *Sst*I site and the 3'-endpoint of the deletion in pAN5-d5.

It cannot be excluded, however, that the effects observed are not only the result of deletion of sequences but also of joining of endpoints of deletions, creating new (artificial) elements. Therefore, not only unilateral deletions were analysed but also a number of specific deletions, removing conserved sequence elements identified by comparison of the promoter sequence of the *A. nidulans gpdA* gene with promoter sequences of other *Aspergillus* genes (see section **a**, above; Fig. 4C).

Deletion of the *gpd* box (d241) results in a 50% decrease of activity. From this result it can be concluded that the *gpd* box contains at least a part of a positively acting, functional promoter element. Deletion of an 124-bp fragment including the *pgk* box (d252) results in a 30% decrease of activity. Deletion of a 79-bp fragment upstream of the *pgk* box (which partially overlaps d252), in combination with a deletion of the *gpd* box (d2414), results in an 80% decrease of activity. From these results we can conclude that sequences directly upstream of the *pgk* box, and presumably not the *pgk* box itself, contain at least a part of a functional promoter element. Comparison of the expression levels of d5 and d6, the deletion endpoints of which map upstream and downstream, respectively, from the *pgk* box, also suggests that removal of the *pgk* box has no significant effect on gene expression. Deletion of the *ct* box (d896) results in an 80% decrease of βGal activity. As described in section **e**, below, deletion of the *ct* box results in aberrant transcription initiation. The decreased gene expression observed for d896 cannot unambiguously be assigned to decreased promoter activity, since also the stability of the altered transcripts could be different. However, it is tempting to speculate that the *ct* box

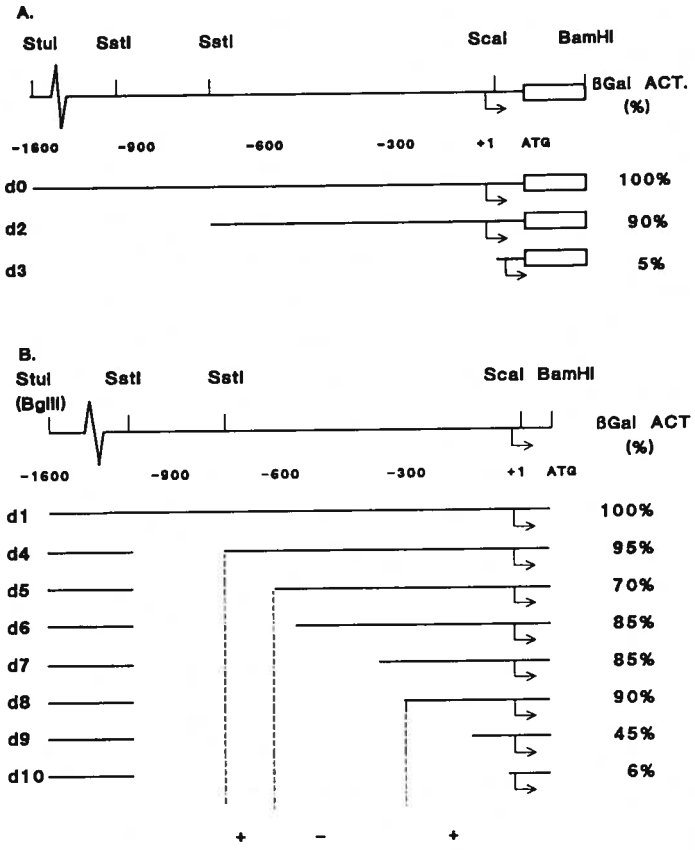
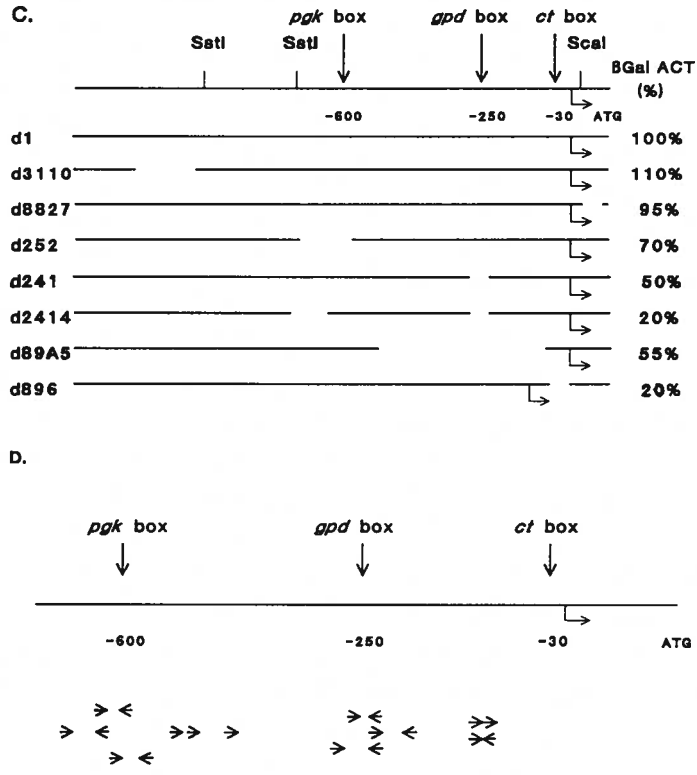


Fig. 4. Relative levels of βGal activity directed by *gpdA* promoter mutants. The extent of the deletion for the promoter mutants is indicated. Deletions which remove from the promoter segments of increasing length were made with Bal31 and/or different restriction enzymes. The methods of Zoller and Smith (1984), and Schold et al. (1984), combined with the method of Kunkel (1985), were used for site-directed mutagenesis. With either ss or ds templates mutant frequencies of 70% and 5-10%, respectively, were obtained. For each deletion the major *tsp* is indicated by a rightward arrow. **(A)** Promoter mutations obtained by restriction enzyme digestion of pAN5-41B (Van Gorcom et al., 1986). A 1.3-kb *Sst*I-*Bam*HI fragment and an 0.5-kb *Sca*I-*Bam*HI fragment from pAN5-41B (=pAN5-d0) were fused to *lacZ* in pAN923-41B to give pAN5-d2 and pAN5-d3, respectively. In these constructions the 5' end of the coding region of the *gpdA* gene is included in the fusion gene. **(B)** Unilateral deletions obtained from a derivative of pAN52-1 (Punt et al., 1987). A 1.8-kb *Stu*I-*Bam*HI fragment from pAN52-1 (Punt et al., 1987) was fused in phase to the *lacZ* gene in pAN923-42B$_{BglII}$ to give pAN5-d1. Furthermore, the *Stu*I site of this fragment was converted into a *Bgl*II site and the resulting *Bgl*II-*Bam*HI fragment was subcloned in a pUC derivative with unique *Bgl*II and *Bam*HI cloning sites. From the resulting vector the 0.34-kb *Sst*I fragment was deleted, resulting in



pAN5-23. Deletions were initiated at the unique *Sst*I site of pAN5-23 using Bal31. After digestion with *Bam*HI, different promoter fragments containing sequential deletions were isolated and cloned in (the large fragment of) pAN5-23 digested with *Bam*HI and *Sst*I. The *Bgl*II-*Bam*HI fragments from the resulting vectors were fused to *lacZ* in pAN923-42B (or pAN923-42B$_{Bglll}$) giving pAN5-d4 to pAN5-d10. In these vectors the start codon for the *gpdA* gene is fused in phase to *lacZ*. For convenience, in pAN52-1, an *NcoI* site and a *BamHI* site were created around the ATG, using site-directed mutagenesis. The location of putative positive (+) and negative (-) promoter elements is indicated. **(C)** Site-specific and aberrant deletions obtained with site-directed mutagenesis using pAN5-d1 or mAN5-10, an M13mp18 vector containing the 1.8-kb *Bam*HI fragment from pAN5-d1, as templates for mutagenesis. The resulting deletions were fused as *Bam*HI fragments to *lacZ* in pAN923-42B or pAN923-42B$_{Bglll}$ resulting in pAN5-d241, pAN5-d2414, pAN5-d252, pAN5-d3110, pAN5-d8827, pAN5-d896 and pAN5-d89A5. **(D)** Location of direct (->->) and inverted (-><-) repeat sequences in the *gpdA* promoter region. The βGal activity assays of protein extracts were carried out as described by Van Gorcom et al. (1985) with *ortho*-nitrophenyl-β-*D*-galactoside as a substrate (ϵ= 0.0045 ml/nMol/cm at 37°C). Cells were cultivated at 35°C for 16-20 h in minimal medium (Pontecorvo et

Fig.4 (continued).-Protein concentrations in extracts were determined using the BioRad protein assay described by Bradford (1976). βGal activities were qualitatively determined using plate assays with XGal as a substrate, as described by Van Gorcom et al. (1985). The enzyme activities were determined in cell extracts from two independently isolated single-copy transformants or duplicate cultures of one transformant. Standard errors between replicate assays within one experiment were always lower than 10%. Although, between experiments some differences in absolute activities were observed, relative activities consistently varied less than 10%. The specific enzyme activities were related to the activity of d0 (panel **A**) or d1 (panels **B** and **C**). In a representative experiment for these mutants a specific activity of respectively 9100 and 11300 units/mg protein was found (1 unit is defined as 1nMol *ortho*-nitrophenol produced per min at 37°C).

Vectors containing promoter mutants with specific deletions of *alc*, *qa* and *qut* boxes were not analysed in this study. Therefore, conclusions about the specific role of these elements in transcription activation/regulation can not be drawn. Since the *alc*, *qa* and *qut* boxes are thought to play a role in transcription activation through binding of regulatory proteins in either ethanol- or quinic acid-induced growth conditions (Gwynne et al., 1987; Hawkins et al., 1988; Baum et al., 1987; Geever et al., 1989), the possible role of these elements in regulation of the *gpdA* promoter should be studied under these growth conditions. Deletion analysis of these elements may provide further information about the role of these boxes.

The expression of many genes involved in carbon metabolism in *A. nidulans*, including the *alcA* and *qut* genes, is carbon-catabolite repressed via the product of the regulatory gene *creA* (Bailey and Arst, 1975). Preliminary results of experiments in which *A. nidulans* was cultivated under catabolite derepressed growth conditions (0.1% fructose; Bailey and Arst, 1975) or in which a mutant *creA* allele was introduced into a strain containing the wt *gpdA* promoter fused to *lacZ*, did not result in a significant change in *lacZ* expression (P. J. P., unpublished results). These results suggest that carbon-catabolite repression does not play a significant role in expression of the *gpdA* gene. Thus, the presence of an *alc* box, suggested to be a possible CreA-binding site in the *alcA* gene (Gwynne et al., 1987), does not confer carbon-catabolite repression to the *gpdA* promoter.

Deletion of the intron in the 5'-untranslated region of the *gpdA* gene (d8827) has no effect on expression levels, which indicates that this intron does not play a role in determining gene expression.

) Transcription initiation analysis

To be able to draw valid conclusions about transcription efficiency from βGal roduction, it is imperative that the fusion transcripts derived from the different romoter mutants are identical. Only then, differences in translation efficiency and/or anscript stability, which could have an effect on βGal levels, can be excluded. The *p* of all deletion mutants was determined by primer extension analysis. Fig. 5 ıows the results of such an analysis for a number of deletion mutants. In all ıutants except d896 (Fig. 5) and d3 (data not shown) identical *tsp* were used. The ıajor *tsp* of the fusion transcript maps at the same position as that found for the ətural *gpdA* transcript. Also the minor *tsp* map at identical positions (Punt et al., 988). One additional minor *tsp* upstream of the major *tsp* was observed for the ısion transcripts at nt position -117. Initiation at the corresponding site, although not ɔted previously, was found for *gpdA* (Punt et al., 1988).

It should be noted that the extension products of d0 and d2 differ from those of 1 and the other deletion mutants (data not shown). However, this is not due to fferent *tsp* but is caused by the presence in d0 and d2 (and d3), but not in the ther deletion mutants, of part of the coding region of the *gpdA* gene (see Fig. 4A).

The major *tsp* is not used in d896 (deletion of the *ct* box). Transcription initiation ɔcurs in this case only at the upstream *tsp* at nt position -117.

Primer extension analysis of d3 revealed one *tsp* at +510 and a number of sites ətween +70 and +130 (results not shown). Since the 3' deletion endpoint of pAN5-3 (Fig. 1) is located in the first intron of the *gpdA* gene (Intron I; Punt et al., 1988) ese intron sequences will probably not be spliced out in this mutant. If that is the ase the upstream *tsp* in d3 map within a C+T-rich region between +70 and +130 ig. 1) such that almost every purine nt is used as a *tsp*. The *tsp* at +510 maps rectly downstream of a small CT region in the *gpdA* coding region of the fusion ene (not shown in Fig. 1) and will presumably not result in a functional transcript. om the results obtained with d896 and d3 we conclude that the presence of a +T-rich region plays an important role in determining the *tsp*. Deletion of the ɔmplete *ct* box - with (d3) or without (d896) upstream sequences - results in tered transcription initiation, whereas in d10 where only part of the *ct* box is eleted, normal transcription initiation is found. Thus, at least that part of the *ct* box, ıich is present in d10 is necessary and sufficient for normal transcription initiation. ırthermore, the altered *tsp* observed with d3 and d896 map downstream from

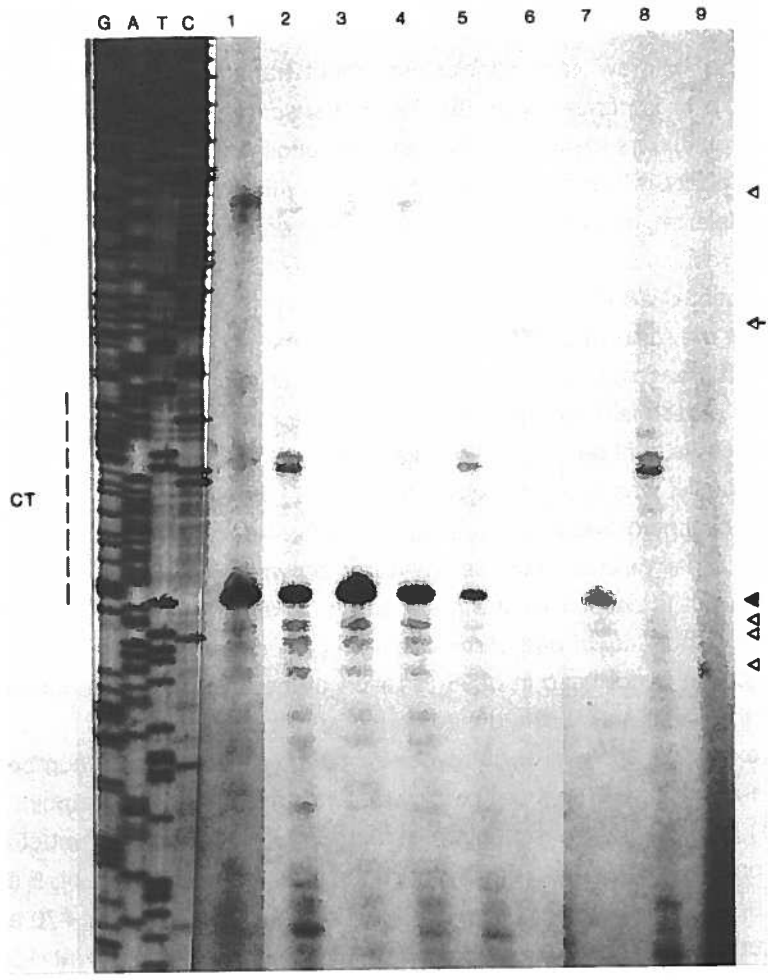


Fig. 5. Primer extension analysis of *gpdA* promoter mutants. Primer extension experiments were performed as previously described (Punt et al., 1988). RNA isolated from different promoter mutants was primed with an oligo (bro48; 5'-AGGCGATTAAGTTGGGTAAC-3') complementary to the 5' part of the coding region of the *lacZ* gene. Lanes G, A, T, C show the products of dideoxy chain-termination sequencing reactions with pAN5-d8827 DNA (in this vector the intron present between the major *tsp* and the ATG was removed, thus the position of the primer extension products corresponds directly to that of the sequencing products). The location of the (sequence complementary to the) *ct* box is indicated with CT and a dashed line. Lanes 1 to 9 give the primer extension products of RNA of (1); d1, (2); d252, (3); d3110, (4); d241, (5); d2414, (6); d896, (7); d8827, (8); $ArgB^{-}$ and (9); without *Aspergillus* RNA. In lane 8 a

determining *tsp*.

C+T-rich regions have been identified directly upstream from most *tsp* of fungal genes (Ballance, 1986; Gurr et al., 1988), most noticebly in the *A. nidulans* and *A. niger oliC* genes (Ward and Turner, 1986; Ward et al., 1988). Deletion of these regions from the *A. nidulans oliC* promoter results in aberrant transcription initiation (J. Brown, in preparation), again suggesting a role for C+T-rich regions in determining the *tsp*. Similar results were also obtained in the *A. nidulans trpC* promoter, although in this case no exact deletion of the C+T-rich region was constructed (Hamer and Timberlake, 1987).

(f) Nuclease S1 analysis

To verify the correlation between transcription efficiency and βGal levels, nuclease S1 analysis (Fig. 6) was carried out with RNA from a number of deletion mutants. The results were quantified by laserscan densitometry. In general, a good correlation was found between the amount of RNA identified in the nuclease S1 analysis and the βGal levels (although standard deviations for the nuclease S1 analysis are much higher than those obtained for βGal assays). Thus, βGal activity measurement is an accurate tool for determining transcription efficiency.

The results of the nuclease S1 analysis also indicated that the amount of RNA from the intact fusion gene is about three fold lower than that of the *gpdA* gene, suggesting a lower transcript stability of the fusion gene. Possibly the use of the *trpC* terminator region or the presence of the *lacZ* gene itself are the cause of this.

---

number of extension products is observed which are not *lacZ*-specific, probably due to hybridization of bro48 with endogenous *Aspergillus* RNA. In lane 1 (besides the aspecific signals) a number of specific extension products is observed, indicated by closed (major extension product) or open (minor products) arrowheads. The major extension product corresponds to transcription initiation at the first purine nt (A) downstream from the *ct* box. Minor products map at a purine nt (G) downstream from a minor C+T-rich region (upstream from the *ct* box around -120; 5'-CTTTGCCCCG) and at purine nt downstream from the major *tsp*. For all other deletions except d896 identical results were observed. For d896 (lane 6) a single extension product is observed upstream from the *ct* box (indicated by an arrow). This extension product maps (after correction for the deleted sequences in d896) at the minor

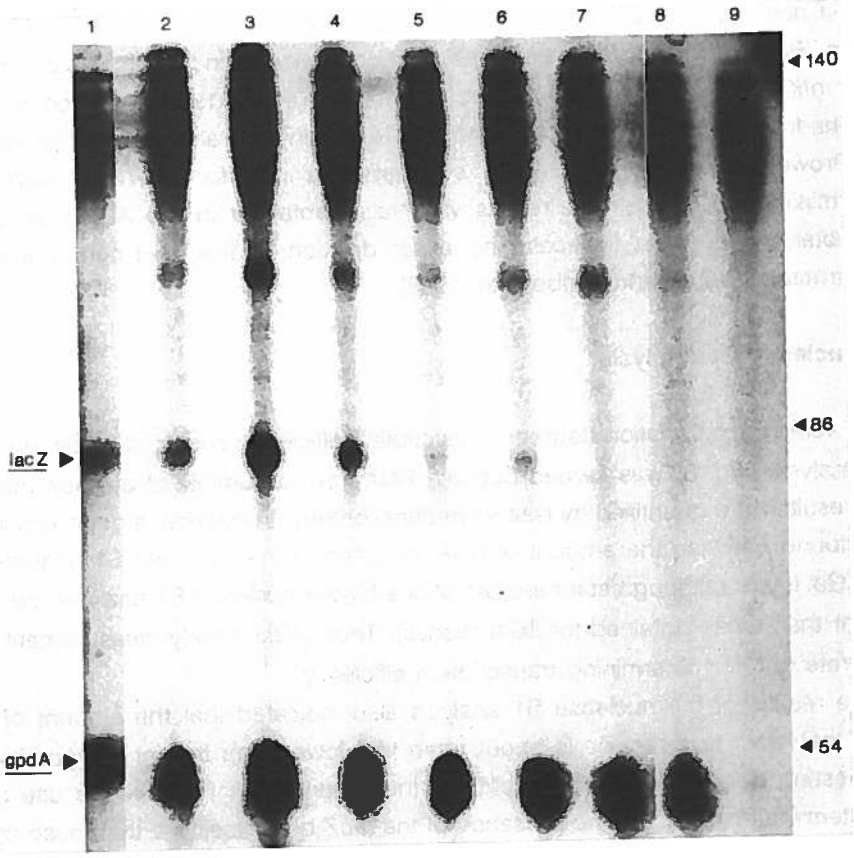


Fig. 6. Nuclease S1 protection analysis of *gpdA* promoter mutants. Nuclease S1 protection experiments were performed as described by Hamer and Timberlake (1987). RNA from a number of promoter mutants was hybridised to a radioactively labeled ss DNA fragment complementary to the coding region of the *gpdA* gene and the coding region of the *lacZ* gene. The ss probe was prepared as described by Hamer and Timberlake (1987). To M13mAN5-7B ss DNA a synthetic oligo (MBL91; 5'–GGCGAAAGGGGGATGTGCTGC-3') complementary to a part of the coding region of the *lacZ* gene, was annealed and extended using PolIK and [$^{\alpha-32}$P] dNTP's . The reaction product was digested with *Sfa*NI and the 140-nt fragment was isolated using denaturing PAGE. Hybridisation and nuclease S1 digestions were performed as described by Hamer and Timberlake (1987). *S. cerevisiae* tRNA was used as carrier in the reactions. The nuclease-resistant DNA fragments were analysed on 6% sequencing gels. For *gpdA* and *lacZ*, protected fragments of 49 and 80 nt, respectively, were

(g) Conclusions

In this paper several functional elements are identified in the *A. nidulans gpdA* promoter by using a *lacZ* fusion strategy. Two of these elements (around nt position -650 and -250) were identified as transcription activation elements. We suggest that the *gpd* box (at nt -250), identified by sequence comparison with the *A. niger gpdA* gene, is a binding site for transcription factors and/or regulatory proteins. Alternatively, this element could also be involved in nucleosome folding and unfolding as was suggested for promoter elements identified in several yeast genes (Struhl, 1986) and the *N. crassa qa* gene cluster (Geever et al., 1989).

Detailed protein-DNA binding assays as performed for *S. cerevisiae* (Struhl, 1986; Johnston, 1987 and references therein) and the *N. crassa qa* gene cluster (Baum et al., 1987) need to be carried out to confirm either of these possibilities.
Further analysis of the *gpdA* promoter will focus on the characterization of the *gpd*-box and on the analysis of proteins binding to this and other functional elements.

ACKNOWLEDGEMENTS

We gratefully acknowledge M. Kolar and M. Busscher for assistance in some of the experiments described in this paper. The GenBank accession number for the sequence published in this paper is M33539.

---

expected (based on the sequence of the probe). Lanes 1 to 9 show the nuclease S1-resistant DNA fragments obtained with RNA from (1); d1, (2); d252, (3); d3110, (4); d241, (5); d2414, (6); d896, (7); a gene conversion, $ArgB^+$ strain, (8); $ArgB^-$ and (9); without *Aspergillus* RNA. As markers, the 140-nt fragment and an *Eco*RI digest of this fragment resulting in a 86-nt and a 54-nt fragment, complementary to *lacZ* and *gpdA*, respectively, were used. The nuclease S1-resistant DNA fragments observed in lane 9 are probably a result of fortuitous hybridization of the probe with carrier tRNA. In lanes 7 and 8 a *gpdA* specific 49-nt fragment is observed (furthermore around 100 nt a protected fragment of unknown origin is observed). In lanes 1 to 6 also a 80-nt *lacZ*-specific fragment is observed. The relative intensities of the *lacZ* signals in the different promoter deletions (using the *gpdA* signals as an internal standard) were calculated using laser-scan densitometry.

Bailey, C. and Arst Jr., H. N.: Carbon catabolite repression in *Aspergillus nidulans*. Eur. J. Biochem. 51 (1975) 573-577.

Ballance, D. J.: Sequences important for gene expression in filamentous fungi. Yeast 2 (1986) 229-236.

Baum, J. A., Geever, R. F. and Giles, N. H.: Expression of qa-1F activator protein: identification of upstream binding sites in the *qa* gene cluster and localization of the DNA-binding domain. Mol. Cell. Biol. 7 (1987) 1256-1266.

Bradford, M. M.: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72 (1976) 248-254.

Chen, E. Y. and Seeburg, P. H.: Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA 4 (1985) 165-170.

Clements, J. M. and Roberts, C. F.: Transcription and processing signals in the 3-phosphoglycerate kinase (PGK) gene from *Aspergillus nidulans*. Gene 44 (1986) 97-105.

Davis, M. A., Cobbett, C. S. and Hynes, M. J.: An *amdS-lacZ* fusion for studying gene regulation in *Aspergillus*. Gene 63 (1988) 199-212.

Fincham, J. R. S.: Transformation in fungi. Microbiol. Rev. 53 (1989) 148-170.

Geever, R. F., Huiet, L., Baum, J. A., Tyler, B. M., Patel, V. B., Rutledge, B. J., Case, M. E. and Giles, N. H.: DNA sequence, organization and regulation of the *qa* gene cluster of *Neurospora crassa*. J. Mol. Biol. 207 (1989) 15-34.

Goc, A. and Weglenski, P.: Regulatory region of the *Aspergillus nidulans argB* gene. Curr. Genet. 14 (1988) 425-429.

Goosen, T., Bloemheuvel, G., Gysler, C., De Bie, D. A., Van den Broek, H. W. J. and Swart, K.: Transformation of *Aspergillus niger* using the homologous orotidine-5'-phosphate-decarboxylase gene. Curr. Genet. 11 (1987) 499-503.

Guarente, L.: UAS's and enhancers: common mechanism of transcriptional activation in yeast and mammals. Cell 52 (1988) 303-305.

Gurr, S. J., Unkles, S. E. and Kinghorn, J. R.: The structure and organisation of nuclear genes of filamentous fungi. In Kinghorn J. R. (Ed.), Gene Structure in Eukaryotic Microbes. SGM Special Publication, Vol. 23. IRL Press, Oxford, 1988, pp. 93-139.

Gwynne, D. I., Buxton, F. P., Sibley, S., Davies, R. W., Lockington, R. A., Scazzochio, C. and Sealy-Lewis, H. M.: Comparison of the *cis*-acting control regions of two coordinately controlled genes involved in ethanol utilization in *Aspergillus nidulans*. Gene 51 (1987) 205-216.

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

Hanahan, D.: Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166 (1983) 557-580.

Hawkins, A. R., Lamb, H. K., Smith, M., Keyte, J. W. and Roberts, C. F.: Molecular organisation of the quinic acid utilization (*QUT*) gene cluster of *Aspergillus nidulans*. Mol. Gen. Genet. 214 (1988) 224-231.

Johnson, P. F. and McKnight, S. L.: Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. 58 (1989) 799-839.

nston, M.: A model fungal gene regulatory mechanism: the *GAL* genes of *Saccharomyces erevisiae*. Microbiol. Rev. 51 (1987) 458-476.

ar, M., Punt, P. J., Van den Hondel, C. A. M. J. J. and Schwab, H.: Transformation of *Penicillium hrysogenum* using dominant selection markers and expression of an *Escherichia coli lacZ* fusion ene. Gene 62 (1988) 127-134.

bs, E. G., Rafter, G. W. and Junge, J. M.: Yeast glyceraldehyde-3-phosphate dehydrogenase, . Yeast protein. J. Biol. Chem. 200 (1953) 479-492.

ıkel, T. A.: Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. \cad. Sci. USA 82 (1985) 488-492.

ıiatis, T., Fritsch, E. F. and Sambrook, J.: Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.

Knight, G. L., O'Hara, P. J. and Parker, M. L.: Nucleotide sequence of the triosephosphate somerase gene from *Aspergillus nidulans*. Implications for a differential loss of introns. Cell 46 1986) 143-147.

ssing, J.: New M13 vectors for cloning. Methods Enzymol. 102 (1983) 20-78.

chaczyk, M., Blanchard, J. M., Marty, L., Dani, C., Panabieres, F., El Sabouty, S., Fort, P. and leanteur, P.: Posttranscriptional regulation of glyceraldehyde-3-phosphate dehydrogenase gene expression in rat tissues. Nucleic Acids Res. 12 (1984) 6951-6963.

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

ıt, P. J., Oliver, R. P., Dingemanse, M. A., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. Gene 56 (1987) 117-124.

ıt, P. J., Dingemanse, M. A., Jacobs-Meijsing, B. J. M., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Isolation and characterization of the glyceraldehyde-3-phosphate dehydrogenase gene of *Aspergillus nidulans*. Gene 69 (1988) 49-57.

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

ıold, M., Colombero, A., Reyes, A. A. and Wallace, R. B.: Oligonucleotide-directed mutagenesis using plasmid DNA templates and two primers. DNA 3 (1984) 469-477.

iday, C. L., Dickman, M. B. and Kolattukudy, P. E.: Structure of the cutinase gene and detection of promoter activity in the 5'-flanking region by fungal transformation. J. Bacteriol. 171 (1989) 1942-1951.

uhl, K.: Yeast promoters. In Reznikoff, W. and Gold, L. (Eds.), Maximizing Gene Expression. Biotechnol. Series 9. Open University Press, Milton Keynes, 1986, pp. 35-78.

uhl, K.: Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell 49 (1987) 295-297.

eri, T. T., Kumar, V., Lehtovaara, P. and Knowles, J.: Construction of cDNA libraries by blunt end ligation: high frequency cloning of long cDNA's from filamentous fungi. Anal. Biochem. 164 (1987) 60-67.

n Gorcom, R. F. M., Pouwels, P. H., Goosen, T., Visser, J., Van den Broek, H. W. J., Hamer, J. E., Timberlake, W. E. and Van den Hondel, C. A. M. J. J.: Expression of an *Escherichia coli* ß-

Van Gorcom, R. F. M., Punt, P. J., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: A syst for the analysis of expression signals in *Aspergillus*. Gene 48 (1986) 211-217.

Ward, M. and Turner G.: The ATP synthase subunit 9 gene of *Aspergillus nidulans*: sequence a transcription. Mol. Gen. Genet. 205 (1986) 331-338.

Ward, M., Wilson, L. J., Carmona, C. L. and Turner G.: The *oliC3* gene of *Aspergillus nig* isolation, sequence and use as a selectable marker for transformation. Curr. Genet. 14 (19 37-42.

Wasylik, B.: Protein coding genes of higher eukaryotes: promoter elements and *trans*-acting factc In Reznikoff, W. and Gold, L. (Eds.), Maximizing Gene Expression. Biotechnol. Series 9, Op University Press, Milton Keynes, 1986, pp. 79-99.

Yanisch-Perron, C., Vieira, J. and Messing, J.: Improved M13 phage cloning vectors and h strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33 (1985) 103-11

Yelton, M. M., Hamer, J. E. and Timberlake, W. E.: Transformation of *Aspergillus nidulans* usin *trpC* plasmid. Proc. Natl. Acad. Sci. USA 81 (1984) 1470-1474.

Zoller, M. J. and Smith, M.: Oligonucleotide-directed mutagenesis: a simple method using t oligonucleotide primers and a single stranded template. DNA 3 (1984) 479-488.

Submitted

CHAPTER 5

upstream activating sequence from the *Aspergillus nidulans dA* gene.

combinant DNA, *amdS::lacZ* fusion gene; targeted single-copy integration; hybrid moters; carbon catabolite repression; nitrogen metabolite repression; ω-amino d induction)

ter J. Punt, Cas Kramer, Anneke Kuyvenhoven, Peter H. Pouwels and Cees M. J. J. van den Hondel

partment of Molecular Genetics and Gene-technology, Medical Biological poratory TNO, PO box 45, 2280 AA Rijswijk, The Netherlands

---

MMARY

ntroduction of a previously identified promoter element of the *Aspergillus nidulans dA* gene, the so-called *gpd* box, into the upstream region of the highly regulated *nidulans amdS* gene, significantly increased (up to 30-fold) the expression of the *Z* reporter gene fused to these expression signals. This increase was dependent the orientation of the *gpd* box and on the site of introduction into the *amdS* stream region. The presence of additional *gpdA* sequences which flank the *gpd* x reduced or even extinguished positive effects of the *gpd* box. ω-Amino acid-, bon catabolite- and ammonium-regulation of the *amdS* promoter were retained er introduction of the *gpd* box, indicating that the *gpd* box does not abolish eractions of the regulatory proteins, AmdR, AreA and CreA with the *amdS* nscription control sequences. Based on the results it is suggested that the *gpd* x comprises at least two separate activities; one being orientation dependent, but atively independent of position of the *gpd* box in the upstream region; the other only functional near other sites of transcription control. Most likely, both activities not involved in regulation of the *amdS* promoter.

Aspergillus nidulans is an attractive organism to study the regulation of eukaryc gene expression for a variety of reasons. Firstly, extensive genetical and biochemi research concerning the regulation of various metabolic pathways has been carr out (e. g., Cove, 1979; Arst and Scazzocchio, 1985; Davis and Hynes, 19 Secondly, molecular-genetical approaches such as gene cloning, transformati gene-replacement, are well developed in this organism (for recent reviews s Timberlake, 1991; Van den Hondel and Punt, 1991).

The expression of several *A. nidulans* genes has been studied. In most cases *vivo* analysis was carried out using a strategy in which various mutant promot were fused to a reporter gene. The expression of this reporter gene is monitorec *A. nidulans* strains containing these fusion genes integrated into the fungal genor Using this type of approach the presence of several positively and negatively act promoter elements has been indicated (e. g., Hamer and Timberlake, 1987; Ada and Timberlake, 1990; Turner et al., 1990).

Initial analysis of the upstream region of one of the *A. nidulans* house-keep genes, *gpdA*, encoding glyceraldehyde-3-phosphate dehydrogenase, revealed presence of various upstream sequences involved in the expression of the ge (Punt et al., 1990). Deletion analysis identified upstream regions involved transcription initiation and transcription efficiency.

In this report a more detailed analysis of one of the identified elements, the g box, is described. This 50 bp element was identified by sequence comparison of upstream regions of the *A. nidulans* and *A. niger gpdA* genes (Punt et al., 19 Punt and Van den Hondel, 1991). Deletion of this element from the upstream reg resulted in a two- to three-fold reduction in gene expression in *A. nidulans* (Punt al., 1990) and *A. niger* (unpublished results). To analyze the effects of introduct of the *gpd* box in other fungal promoters, we introduced DNA fragments carry the *gpd* box into the upstream region of the highly regulated *amdS* gene a determined the level of gene expression obtained with the hybrid promoters. T upstream region of the *amdS* gene is chosen for this study, as much is kno about the regulatory circuits of this gene (for a review see Hynes et al., 1989; Da and Hynes, 1991).

CHAPTER 5

MATERIALS AND METHODS

(a) Strains, plasmids and transformation procedures

A. nidulans ArgB$^-$ (*argB*2, *biA*1, *methG*2) was used as a recipient in all transformation experiments. Vectors were constructed and propagated in *E. coli* K-12 JM109 (Yanisch-Perron et al., 1985). Plasmids pAN5-23 and pAN5-d7 were previously described by Punt et al. (1990), pLIT14 by Richardson et al. (1989), M13mp18 by Yanish-Perron et al. (1985).

E. coli transformations were carried out according to Hanahan (1983) and M13 transfections according to Kunkel (1985). *A. nidulans* transformations were carried out as described previously (Punt and Van den Hondel, 1992).

(b) β-Galactosidase assays

Qualitative βGal activities were determined using plate assays with XGal as a substrate as described by Van Gorcom et al. (1985). For the detection of low levels of activity the method of Kolar et al. (1991) was used.

Quantitative βGal activity assays of mycelial extracts were carried out as described previously (Punt et al., 1990). Mycelial extracts were prepared from cells that were incubated at 35°C for 18h in Aspergillus minimal growth medium (Bennett and Lasure, 1991) supplemented with methionine (100 μg/ml) and biotin (200 ng/ml). Unless otherwise indicated 1% glucose and 10mM nitrate were used as carbon- and nitrogen-source.

(c) Molecular methods

Fungal DNA isolations were performed as described by Kolar et al. (1988). All other methods were essentially as described by Sambrook et al. (1989).

RESULTS AND DISCUSSION

(a) Construction of *amdS::lacZ* vectors carrying *gpd* box variants

To analyze the effects of the *gpd* box on transcription activation DNA fragments

amdS gene fused to *E. coli lacZ* (Fig. 1). A fragment containing the 50 bp *gpd* box with flanking *Sau*IIIA cloning-sites was chemically synthesized. Two additional DNA fragments containing the *gpd* box with flanking sequences were isolated as restriction fragments. These fragments have been included in this study, since they contain a number of inverted repeat sequences which overlap with the *gpd* box (Fig. 1). A 115 bp *Bgl*I-*Sal*I fragment was isolated from pAN5-23 (Punt et al., 1990) and a 170 bp *Sst*I-*Ava*II fragment was isolated from pAN5-d7 (Punt et al., 1990). These fragments were blunted with T4 DNA polymerase and cloned into the *Sma*I site of M13mp18, resulting in vectors mAN5-11 and mAN5-21, respectively. From these vectors *Sst*I-*Bam*HI fragments were isolated and blunted with T4 DNA polymerase for subsequent cloning. (These fragments will be referred to as 115 bp and 170 bp *gpd* box fragments; Fig. 1).

The three *gpd* box fragments were introduced into pAN49-1 (Fig. 2A). This plasmid is obtained by introducing the 3.4-kb *Xba*I fragment from pAN923-41B_{BglII}, carrying a mutant *argB* allele (Punt et al., 1990), into the unique *Eco*RI site of pLIT14 (Richardson et al., 1989). In pAN49-1 both *Eco*RI and *Xba*I sites were regenerated.

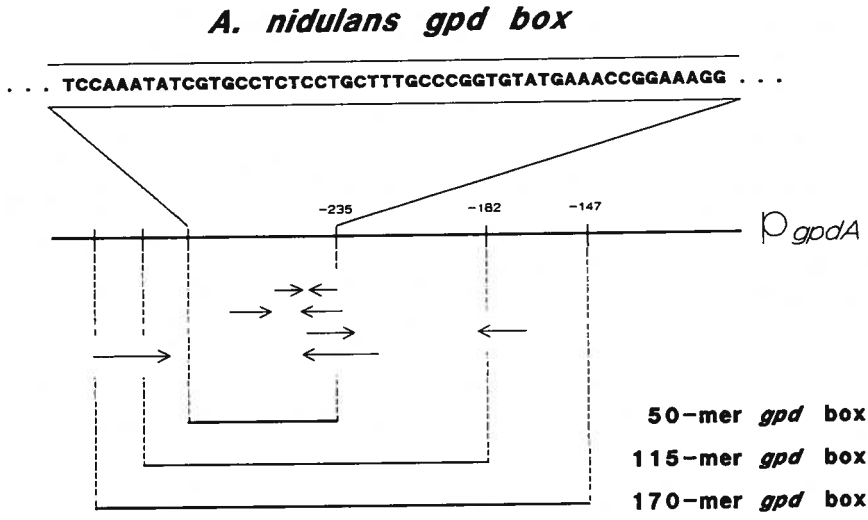


Fig. 1. Schematic representation of the location of the *gpd* box containing fragments in the *gpdA* upstream region. The distance from the major *tsp* is given in nt. The location and extent of inverted repeat sequences is given by converging arrows.

ıe *gpd* box fragments were introduced at different positions in the upstream region the *amdS* gene; **1)** at the *Bam*HI site (at -81 nt from the major *tsp*), within the *ndS* transcription control region as defined by *in vivo* titration analysis (Hynes and avis, 1986; Hynes et al., 1988; Hynes et al., 1989), (pAN49-3✱ vectors; Fig. 2B), at the *Bgl*II site (at -616 nt), far upstream of the *amdS* transcription control region, AN49-2✱ vectors; Fig. 2B). Furthermore, vectors were constructed in which *gpd* ox fragments replaced the *Bgl*II-*Bam*HI fragment (pAN49-4✱ vectors; Fig. 2B). In ost constructions the *gpd* box fragment was cloned in two orientations. onstructions carrying the *gpd* box in the same orientation as in the upstream gion of the *gpdA* gene are labelled "A". The opposite orientation is labelled "B".

) Construction of *A. nidulans* strains

The constructed vectors were introduced into *A. nidulans* ArgB$^-$. *A. nidulans* ınsformants were plated onto Xgal plates to identify those transformants that have corporated the *lacZ*-fusion gene. Only in a few cases LacZ$^+$ transformants could identified in this way. In previous experiments with vectors carrying the mutant *gB* allele about 50% of the transformants contained the vectors integrated at the gB locus. Therefore, it was expected that the expression level of the *amdS::lacZ* sion was too low to be detected in this plate assay. Accordingly, a more sensitive ate assay (Kolar et al., 1991) was adopted, which resulted in the identification of any more LacZ$^+$ transformants. Southern analysis of LacZ$^+$ transformants was used identify transformants containing a single copy of the vector at the *argB* locus esults not shown). Two independently isolated single copy transformants of each asmid were used for further analysis.

) Expression analysis of the *gpd* box containing *amdS::lacZ* single copy ınsformants

Quantitative βGal assays were carried out in mycelial extracts from single copy ınsformants grown in minimal medium. Introduction of the 50 bp *gpd* box in either entation in the *Bam*HI site within the *amdS* transcription control region led to a nsiderable increase in βGal activity (Fig. 2B; pAN49-31 transformants). Also at a ısition far upstream of the transcription control region the 50 bp *gpd* box was own to be active. However, in this case only the "A" orientation led to an increase

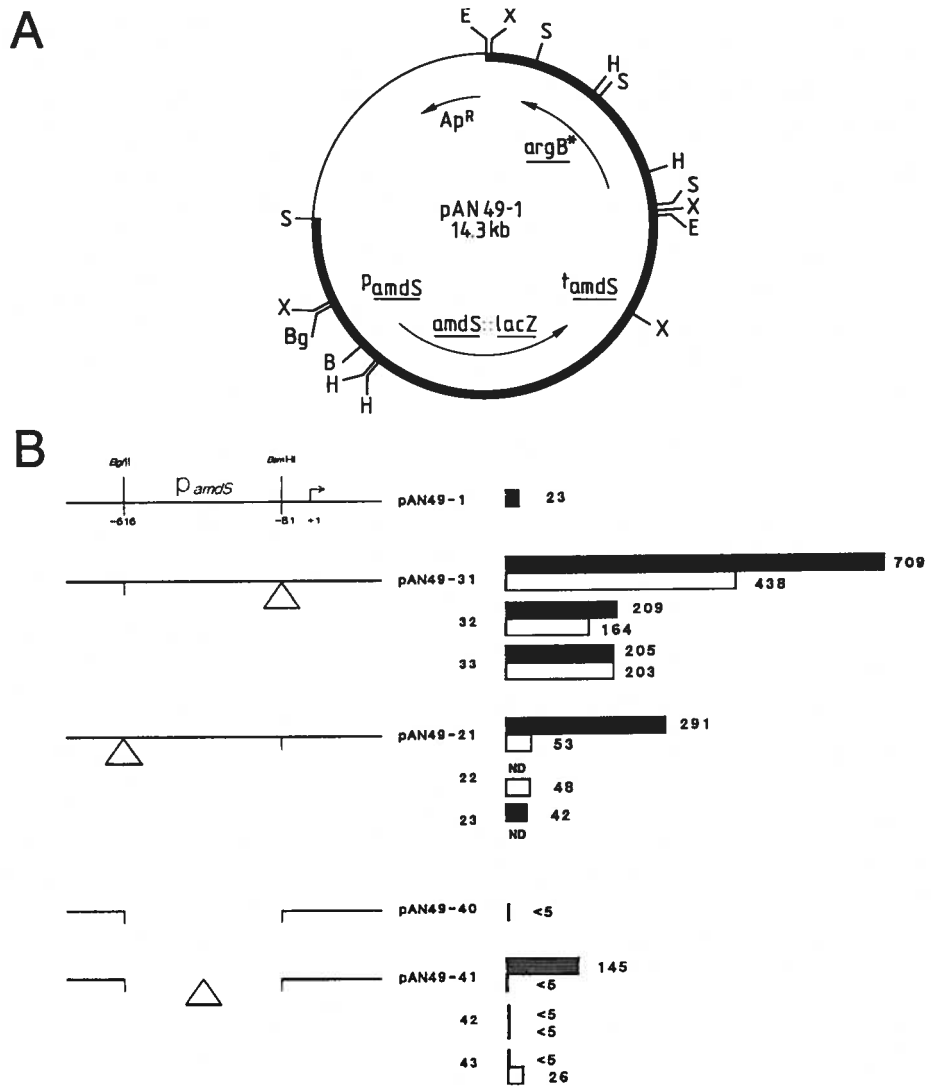


Fig. 2 **(A)**. Restriction map of vector pAN49-1. Fragments containing the *gpd* box were cloned in the unique *Bam*HI (B) or *Bgl*II (Bg) sites in the *amdS* promoter region (p_{amdS}). E, *Eco*RI; H, *Hind*III; S, *Sal*I; X, *Xba*I; *argB**, mutant *argB* allele (to obtain site specific integration); p_{amdS}, promoter region of the *amdS* gene; t_{amdS}, terminator region of the *amdS* gene. **(B)** Graphical representation of the level of βGal activity obtained with transformants carrying a single copy of one of the various *gpd* box plasmids cultivated in medium with 1% glucose and 10 mM nitrate as carbon and nitrogen source. Vectors pAN49-21, -31 and -41 carry the 50 bp *gpd* box fragment; -22, -32 and -42 the 115 bp *gpd* box fragment; -23, -33 and -43 the 170-bp

in gene expression (Fig. 2B; pAN49-21 transformants). A similar orientation dependent effect was observed when the 50 bp *gpd* box replaced the *Bgl*II-*Bam*HI fragment (Fig. 2B; pAN49-41 transformants).

These results indicate that the 50 bp *gpd* box is capable of activating gene expression when positioned at various sites in the *amdS* upstream region. Furthermore, our results indicate that the level of activity of the *gpd* box is depending on; 1) its position in the *amdS* upstream region and 2) the orientation of the *gpd* box in the *amdS* upstream region. Preliminary deletion analysis of the 50 bp *gpd* box indicates the presence of separate functional elements which contribute to the total activity of the *gpd* box (P. J. P, A. K., unpublished results). These results, in combination with the results described above, suggest that the *gpd* box consists of more than one functional element. It is feasible that each of the elements engage activation pathways with different properties.

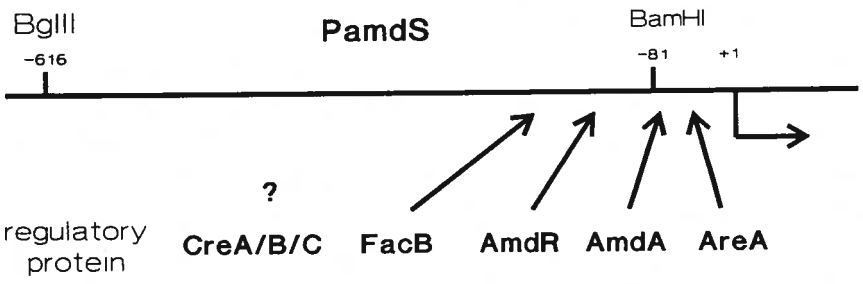
In addition to the 50 bp *gpd* box also *gpd* box containing fragments with flanking *gpdA* sequences were tested for their transcription activation activity. The results presented in Fig. 2B indicate that, in all cases tested, these flanking sequences reduce (in the case of pAN49-32/33 transformants) or even almost completely extinguish the activiting properties of the 50 bp *gpd* box (pAN49-22/23 and pAN49-42/43 transformants). One of the explanations of these results could be the presence of negative promoter elements in the flanking sequences of the 50 bp *gpd* box. The presence of such negative promoter elements in the *gpdA* upstream region was already indicated by results obtained with deletion analysis (Punt et al., 1990). However, the position of these elements does not coincide with the *gpd* box flanking sequences present in the 115 bp and 170 bp *gpd* box containing fragments.

The introduction of the *gpd* box into the upstream region of a *amdS*::*lacZ* fusion gene results in a 6 to 30-fold increase in gene expression (Fig. 2B), which is far

---

gpd box fragment. For the different plasmids the position of the *gpd* box containing fragment in the *amdS* promoter region is indicated by a triangle. Solid bars indicate the results obtained with plasmids without *gpd* box sequences (pAN49-1, -40) and plasmids carrying the *gpd* box fragments in the A orientation. Open bars indicate the results obtained with the *gpd* box fragments in the B orientation. The level of βGal activity indicated is the average of at least two independent experiments using two independently isolated single copy transformants. Standard errors between replicate assays were always lower than 15%. The distance of the *Bam*HI and *Bgl*II sites to the major transcription initiation site as determined by Corrick et al.

more than its effect in the *gpdA* upstream region (2 to 3-fold increase; Punt et al., 1990). The maximal level of gene expression (ca. 700 u βGal/mg; Fig. 2B), however, is still considerably less than the *gpd* box dependent level of gene expression in the *gpdA* upstream region (6000 u βGal/mg; Punt et al., 1990). This result indicates that introduction of the *gpd* box does not simply have an independent effect on the level of gene expression. Apparently, other elements of the upstream region, such as transcription initiation sequences, are of importance for the effect of the *gpd* box on gene expression. Adams and Timberlake (1990) have shown that the effect of introduction of upstream sequences of the *abaA* gene into the *trpC* upstream region is influenced by the presence of *trpC* transcription initiation sequences, which also indicates direct or indirect communication between different elements in fungal



CreA/B/C	carbon catabolite repression
FacB/AmdA	acetate induction
AmdR	ω-aminoacid induction
AreA	nitrogen metabolite repression

Fig. 3. Schematic representation of the approximate location of target sites (indicated with arrows) of the products of various regulatory genes involved in the expression of the *amdS* gene (Hynes et al., 1988). The position (and distance from the major *tsp* in nt) of the *Bam*HI and *Bgl*II cloning sites used in this study is indicated. The regulatory circuits of the different regulatory genes are indicated. The location of the target site for the regulatory proteins involved in carbon catabolite repression has not been determined experimentally.

ıstream regions. Similar relationships between transcription control sequences and ınscription initiation sequences have been observed for other eukaryotic genes /en et al., 1991; Mellor et al., 1991).

As shown in Fig. 2B, the activity of the *gpd* box is position dependent, showing ə strongest effect when a 50 bp *gpd* box is positioned in the region of the *amdS* ıstream sequences which was shown to be involved in transcription control AN49-3✱ transformants; Hynes et al., 1988; Hynes et al., 1989). To analyze in ore detail whether the activity of the *gpd* box is interfering with the normal ;tivation circuits of the *amdS* gene (Fig. 3; Hynes et al., 1989; Davis and Hynes, ı91), βGal expression was analyzed in mycelial extracts from single copy ınsformants with 50 bp *gpd* box plasmids, cultivated under various ıducing/repressing) growth conditions.

) Effect of introduction of the *gpd* box on ω-amino acid regulation of the *ndS* promoter

Expression of the *amdS* gene and the *amdS::lacZ* fusion gene is induced by ω-nino acids, such as β-alanine and γ-amino butyric acid (Arst, 1976; Hynes, 1978a, ıvis et al., 1988). This regulation is mediated by the action of the gene product of ə regulatory gene *amdR*, which is indicated to be a DNA-binding (Zn-finger like) otein (Andrianopoulos and Hynes, 1988; 1990). The βGal level of pAN49-1 ınsformants cultivated in medium containing nitrate + β-alanine (induced) is clearly ɔreased compared to the level obtained in the absence of β-alanine (non-induced; g. 4). Also in pAN49-31 and -21 transformants significant ω-amino acid induction ıs observed (Fig. 4), indicating that transcription activation by the *amdR* gene-oduct is not abolished by the introduction of the *gpd* box. In pAN49-40 and \N49-41 transformants no induction in the presence of β-alanine was observed ig. 4). These results are in agreement with previous data, obtained by sequence ıalysis of mutant *amdS* alleles (Corrick et al., 1987; Hynes et al., 1988), *in vivo* ɾation experiments (Hynes et al., 1988) and mobility shift assays (Van Heeswijck ıd Hynes, 1991), which localized the target-site for the AmdR protein to a region ɾectly upstream of the *Bam*HI site in pAN49-1.

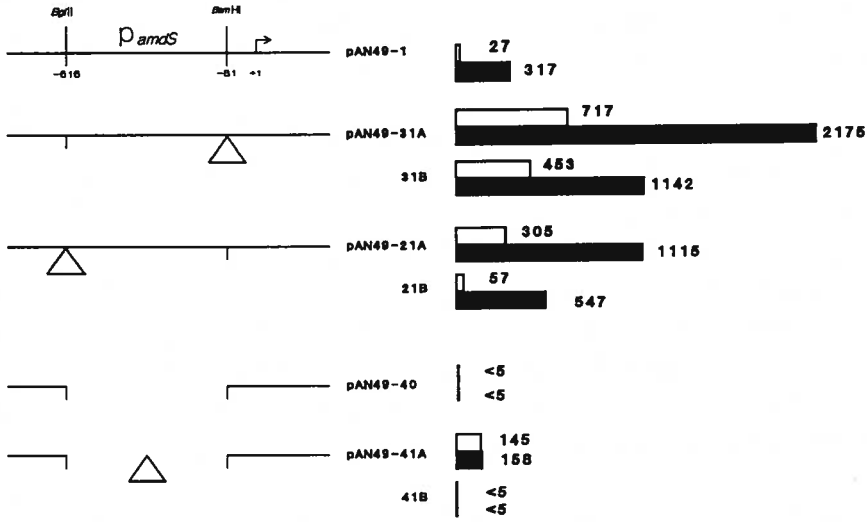


Fig. 4. Graphical representation of the level of βGal activity obtained with transformants carrying a single copy of one of the 50 bp *gpd* box plasmids cultivated in medium with 1% glucose and either 10 mM nitrate (non-induced; open bars) or 10 mM nitrate + 10 mM β-alanine (induced; solid bars). The level of βGal activity indicated is the average of at least two independent experiments. For further details see the legend to Fig. 2.

(e) Effect of introduction of the *gpd* box on carbon catabolite repression of t *amdS* promoter

As is true for many *A. nidulans* genes involved in carbon catabolism, the *am* gene is controlled by carbon catabolite repression. In the presence of sufficie amounts of favourable carbon-sources (e. g., 1% glucose) expression of the *am* gene is repressed, while under conditions of (limited) carbon starvation (0.1 fructose) increased *amdS* expression is observed (Bailey and Arst, 1975; Davis al., 1988). Untill now not much is known about the mechanism by which carb catabolite repression is affecting transcription activation. Mutant analysis h identified at least three genes involved in the process (*creA/B/C*; Bailey and Ar 1975; Hynes and Kelly, 1977). Of these three genes only *creA* was cloned (Dowz and Kelly, 1989). This gene is suggested to encode a (negatively-acting) DN binding regulatory protein, as its predicted aminoacid sequence has a number c

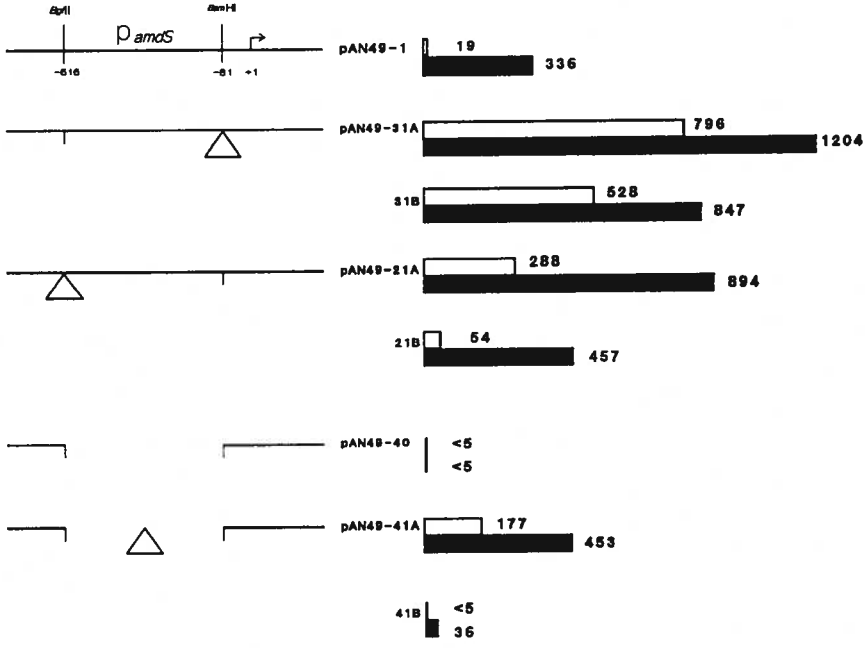


Fig. 5. Graphical representation of the level of βGal activity obtained with transformants carrying a single copy of one of the 50 bp *gpd* box plasmids cultivated in medium with 10 mM nitrate and either 1% glucose (catabolite repressed; open bars) or 0.1% fructose (derepressed; solid bars). For further details see the legend to Fig. 4.

features in common with established DNA-binding proteins (Dowzer and Kelly, 1991). At present no experimental evidence concerning the target-site(s) for carbon catabolite control are available.

As shown in Fig. 5, the expression of the *amdS::lacZ* gene in pAN49-1 transformants was increased under conditions of catabolite derepression (0.1% fructose). Also in the presence of the *gpd* box a considerable increase was observed (Fig. 5). In contrast to the results obtained for ω-amino acid induction (see RESULTS AND DISCUSSION, section d), pAN49-41A/B transformants respond to carbon catabolite derepression (Fig. 5). The most likely explanation for this result is that the target-site(s) for carbon catabolite repression are positioned downstream of the *Bam*HI site or upstream of the *Bgl*II site in pAN49-1. There is preliminary evidence that at least one of the promoter element(s) involved in carbon catabolite

repression is positioned downstream of the *Bam*HI site (Hynes et al., 1989).

Since effects of carbon catabolite repression were not observed for the expression of the p_{gpdA}::*lacZ* fusion gene (Punt et al., 1990) an alternative explanation, i. e. that the *gpd* box itself is involved in carbon control, is invalidated.

(f) Effect of introduction of the *gpd* box on ammonium repression of the *amdS* promoter

Expression of the *amdS* gene is also affected by nitrogen metabolite repression. In the absence of favourable nitrogen sources such as ammonium the expression of the *amdS* gene is derepressed, through the action of the gene product of the regulatory gene *areA* (Arst and Cove, 1973; Hynes, 1978a). Based on its deduced aminoacid sequence, it is suggested that the *areA* gene encodes a specific DNA-binding regulatory protein (Kudla et al., 1990). The exact site of interaction in the *amdS* upstream region is not known for the *areA* gene product. It is suggested that the target site for the *areA* gene product is identical to that of the homologous *nit-2* gene product from *Neurospora crassa* (Davis and Hynes, 1987; Fu and Marzluf, 1990). The upstream sequences of many *A. nidulans* genes regulated by the action of *areA* protein, including *amdS*, contain sequences similar to this target site (TATCTA; Fu and Marzluf,1990). However, no experiments have been reported to demonstrate the functionality of these sequences.

As shown in Fig. 6, for pAN49-1 transformants no decrease of expression (even an increase) was observed in media containing both nitrate+ammonium (non-induced, repressed) compared to media containing only nitrate (non-induced). Even in the presence of ammonium only (repressed) expression levels were not lower than in nitrate medium (results not shown). In contrast to the results obtained for pAN49-1 transformants, considerable ammonium repression was observed in all transformants containing a *gpd* box plasmid, except in pAN49-21B transformants in which only a low level of expression of the *lacZ* gene was found (Fig. 6). One explanation for the absence of ammonium repression in pAN49-1 (and pAN49-21B) transformants could be the introduction of the *Bam*HI cloning site at the position of the *Sma*I site in the wt *amdS* promoter region in pLIT14 (Richardson et al., 1989) and, thus, in all pAN49-plasmids except pAN49-40 and its derivatives. Sequences near this *Sma*I site are involved in *amdS* expression, as was inferred from sequence analysis of an *amdS* up-promoter mutation (*amdI*18; Hynes, 1978b; Corrick et al., 1987) which was selected as a revertant of an *areA* mutant (Hynes, 1978b).

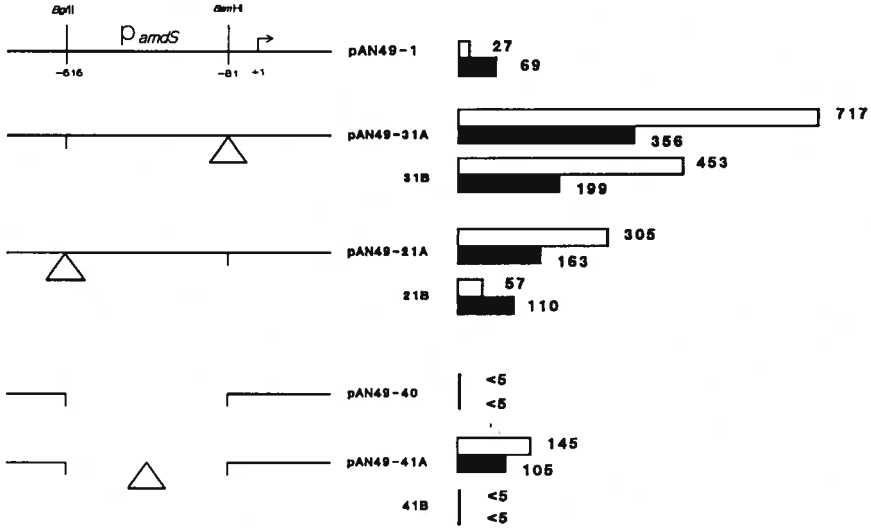


Fig. 6. Graphical representation of the level of βGal activity obtained with transformants carrying a single copy of one of the 50 bp *gpd* box plasmids cultivated in medium with 1% glucose and either 10 mM nitraat (non-induced; open bars) or 10 mM nitrate + 10 mM ammonium-tartrate (repressed; solid bars). For further details see the legend to Fig. 4.

Therefore, it is conceivable that the sequence alteration as a consequence of the introduction of the *Bam*HI site causes an elevated level of expression in the presence of nitrate + ammonium as nitrogen sources, which may obscure effects of nitrogen metabolite repression in pAN49-1, pAN49-21B and, to some extent, pAN49-41A transformants (Fig. 6). However, the results obtained with transformants containing an active *gpd* box upstream of the *amdS::lacZ* fusion gene demonstrated that transcription activation by AreA is not abolished by the presence of the *gpd* box (Fig. 6). The requirement for a functional AreA protein for the expression of the *amdS::lacZ* fusion gene was also demonstrated under conditions of carbon catabolite derepression (0.1% fructose, see RESULTS AND DISCUSSION section (e)). Under these conditions ammonium repression could be seen also in pAN49-1 and pAN49-21B transformants (Fig. 7).

In accordance with previous data (Hynes et al., 1988), the results obtained for pAN49-41 transformants (Fig. 6, 7) indicate that ammonium repression is still observed in the absence of *amdS* sequences from the region upstream of the *Bam*HI site in pAN49-1.

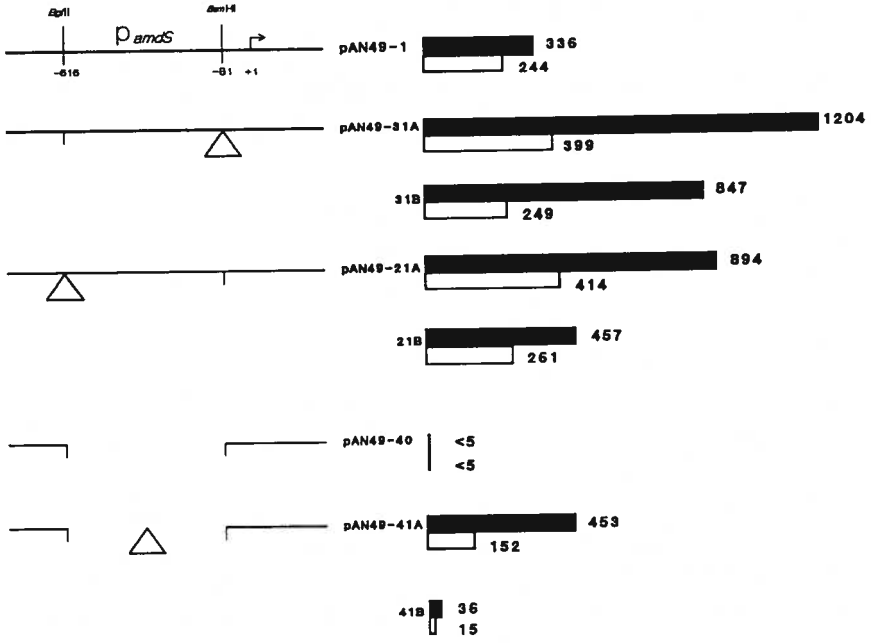


Fig. 7. Graphical representation of the level of βGal activity obtained with transformants carrying a single copy of one of the 50 bp *gpd* box plasmids cultivated in medium with 0.1% fructose and either 10 mM nitraat (non-induced; solid bars) or 10 mM nitrate + 10 mM ammonium-tartrate (repressed; open bars). For further details see the legend to Fig. 4.

(g) Conclusions

Based on sequence comparison and *in vivo* deletion analysis of the upstream region the *A. nidulans gpdA* gene a putative 50 bp transcription control element, the *gpd* box, was identified (Punt et al., 1990; Punt and Van den Hondel, 1991). In this report we show that this 50 bp element, when introduced into another fungal promoter is capable of increasing gene expression. We show that introduction of the *gpd* box in the *amdS* upstream region results in orientation- and position-dependent effects on the level of gene expression, which suggests that the *gpd* box does not consist of one single functional element but of two or more (maybe overlapping) elements. Furthermore, we show that the presence of sequences from the *gpdA*

ɔstream region, flanking the 50 bp *gpd* box negatively affects the activity of the ɔd box.

The results also show that regulation of the *amdS::lacZ* fusion gene is still subject ω-amino acid induction, carbon catabolite repression and nitrogen metabolite pression after introduction of the *gpd* box, which indicates that the trans-acting ctors involved in these regulation circuits, AmdR, CreA and AreA proteins, can still teract with (at least some of) their target sequences in the *amdS* upstream region. ıe effect of the *gpd* box and those of the various regulatory circuits are more or ss additive. A similar additivity was also observed for the different regulatory rcuits involved in *amdS* expression (Hynes, 1978a; Hynes and Davis, 1986; Hynes al., 1989; Davis and Hynes, 1991). This additivity indicates that the transcription ɔtivating effects of the *gpd* box do not interfere with the action of the regulatory ·oteins.

Interestingly, our results may indicate that ammonium repression and carbon ıtabolite repression of the *amdS* expression are mediated by *amdS* sequences cated in a region between the *Bam*HI site and the startcodon in pAN49-1. Based ı sequence analysis of *amdS* upstream sequences isolated from strains carrying s-acting regulatory mutations (Hynes et al., 1988; Katz et al., 1990), this region was so shown to be involved in the *amdA* dependent acetate induction. This result ɔuld imply that the target-sites for transacting regulatory proteins involved in these ree regulatory circuits may consist of (partially) overlapping sequences. This triguing hypothesis may offer a way to explain several untill now puzzling results dicating an interrelationship between carbon and nitrogen control circuits (Arst and ove, 1973; Arst and Scazzocchio, 1985; Davis and Hynes, 1991).

CKNOWLEDGEMENTS

Dr. Michael J. Hynes is acknowledged for fruitful discussions and critical reading the manuscript. This paper is dedicated with great appreciation to Dr. Frits erends on the occasion of his retirement as head of the department of iochemistry of the TNO Medical Biological Laboratory.

EFERENCES

dams, T. H., and Timberlake, W. E.: Upstream elements repress premature expression of an

Andrianopoulos, A., and Hynes, M. J.: Cloning and analysis of the positively acting regulatory ger *amdR* from *Aspergillus nidulans*. Mol. Cell. Biol. 8 (1988) 3532-3541.

Andrianopoulos, A., and Hynes, M. J.: Sequence and functional analysis of the positively actir regulatory gene *amdR* from *Aspergillus nidulans*. Mol. Cell. Biol. 10 (1990) 3194-3204.

Arst, H. N.: Integrator gene in *Aspergillus nidulans*. Nature 262 (1976) 231-234.

Arst, H. N., and Cove, D. J.: Nitrogen metabolite repression in *Aspergillus nidulans*. Mol. Ge Genet.126 (1973) 111-141.

Arst, H. N. and Scazzochio, C.: Formal genetics and molecular biology of the control of ger expression in *Aspergillus nidulans*. In Bennett, J. W. and Lasure, L. L. (Eds.), Gene Manipulatio in Fungi, Academic Press, New York, 1985, pp. 309-343.

Bailey, C. and Arst Jr., H. N.: Carbon catabolite repression in *Aspergillus nidulans*. Eur. J. Biochei 51 (1975) 573-577.

Bennett, J. W. and Lasure, L. L.: Growth media. In Bennett, J. W. and Lasure, L. L. (Eds.), Mo Gene Manipulations in Fungi, Academic Press, New York, 1991, pp. 441-458.

Corrick, C. M., Twomey, A. P., and Hynes, M. J.: The nucleotide sequence of the *amdS* gene *Aspergillus nidulans* and the molecular characterization of 5' mutations. Gene 53 (1987) 63-7

Cove, D. J.: Genetic studies of nitrate assimilation in *Aspergillus nidulans*. Biol. Rev. 54 (1979) 29 327.

Davis, M. A., Cobbett, C. S. and Hynes, M. J.: An *amdS-lacZ* fusion for studying gene regulati in *Aspergillus*. Gene 63 (1988) 199-212.

Davis, M. A. and Hynes, M. J.: Complementation of *areA*$^{-}$ regulatory gene mutations of *Aspergill nidulans* by the heterologous regulatory gene *nit-2* of *Neurospora crassa*. Proc. Natl. Acad. S USA 84 (1987) 3753-3757.

Davis, M. A. and Hynes M. J.: Regulatory circuits in *Aspergillus nidulans*. In Bennett, J. W. a Lasure, L. L. (Eds.), More Gene Manipulations in Fungi, Academic Press, New York, 1991, p 151-189.

Dowzer, C. E. A. and Kelly, J. M.: Cloning of the *creA* gene from *Aspergillus nidulans*: a ger involved in carbon catabolite repression. Curr. Genet. 15 (1989) 457-459.

Dowzer, C. E. A. and Kelly, J. M.: Analysis of the *creA* gene, a regulator of carbon catabol repression in *Aspergillus nidulans*. Mol. Cell. Biol. 11 (1991) 5701-5709.

Fu, Y.-H. and Marzluf, G. A.: *Nit-2*, the major positive-acting nitrogen regulatory gene of *Neurospo crassa*, encodes a sequence-specific DNA-binding protein. Proc. Natl. Acad. Sci. USA 87 (199 5331-5335.

Hamer, J. E. and Timberlake, W. E.: Functional organization of the *Aspergillus nidulans trp* promoter. Mol. Cell. Biol. 7 (1987) 2352-2359.

Hanahan, D.: Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166 (198 557-580.

Hynes, M. J.: A *cis*-dominant regulatory mutation affecting enzyme induction in the eukaryo *Aspergillus nidulans*. Nature 253 (1975) 210-211.

Hynes, M. J.: Multiple independent control mechanisms affecting the acetamidase of *Aspergill nidulans*. Mol. Gen. Genet. 161 (1978a) 59-65.

Hynes, M. J.: An "up-promoter" mutation affecting the acetamidase of *Aspergillus nidulans*. Mol. Ge Genet. 166 (1978b) 31-36.

Hynes, M. J. and Davis, M. A.: The *amdS* gene of *Aspergillus nidulans*: control by multiple regulatory signals. BioEssays 5 (1986) 123-128.

Hynes, M. J. and Kelly, J. M.: Pleiotropic mutants of *Aspergillus nidulans* altered in carbon metabolism. Mol. Gen. Genet. 150 (1977) 193-204.

Hynes, M. J., Andrianopoulos, A., Davis, M. A., Van Heeswijck, R., Katz, M. E., Littlejohn, T. G., Richardson, I. B. and Saleeba, J. A.: Multiple circuits regulating the *amdS* gene of *Aspergillus nidulans*. In Nevalainen, H. and Penttilä, M. (Eds.) Proceedings of the EMBO-Alko workshop on Molecular Biology of Filamentous Fungi, Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, 1989, pp. 53-62.

Hynes, M. J., Corrick, C. M., Kelly, J. M. and Littlejohn, T. G.: Identification of the sites of action for regulatory genes controlling the *amdS* gene of *Aspergillus nidulans*. Mol. Cell. Biol. 8 (1988) 2589-2596.

Katz, M. E., Saleeba, J. A., Sapats, S. I. and Hynes, M. J.: A mutation affecting *amdS* expression in *Aspergillus nidulans* contains a triplication of *cis*-acting regulatory sequence. Mol. Gen. Genet. 220 (1990) 373-376.

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

Kolar, M., Holzmann, K., Weber, G., Leitner, E. and Schwab, H.: Molecular characterization and functional analysis in *Aspergillus nidulans* of the 5'-region of the *Pennicillium chrysogenum* isopenicillin N synthetase gene. J. Biotechn. 17 (1991) 67-80.

Kudla, B., Caddick. M. X., Langdon, T., Martinez-Rossi, N. M., Bennett, C. F., Sibley, S., Davies, R. W. and Arst, H. N.: The regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus nidulans*. Mutations affecting specificity of gene activation alter a loop residue of a putative zinc finger. EMBO J. 9 (1990) 1355-1364.

Kunkel, T. A.: Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA 82 (1985) 488-492.

Mellor, J., Midgely, C., Kingsman, A. J., Kingsman, S. M. and Kim, S.: Transcriptional activation by upstream activation sequences requires distinct interactions with downstream elements in the yeast *TRP1* promoter. Mol. Gen. Genet. 225 (1991) 217-224.

Punt, P. J., and Van den Hondel, C. A. M. J. J.: Analysis of transcription control sequences in filamentous fungi. In Tudzynski, P. and Stahl, U. (Eds.), Proceedings of the EMBO workshop on Molecular Biology of Filamentous Fungi, VCH Verlag, Berlin, in press.

Punt, P. J. and Van den Hondel, C. A. M. J. J.: Transformation of filamentous fungi based on hygromycin B and phleomycin resistance markers. Meth. Enzymol., in press.

Punt, P. J., Dingemanse, M. A., Kuyvenhoven, J., Soede, R. D. M., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene encoding glyceraldehyde-3-phosphate dehydrogenase. Gene 93 (1990) 101-109.

Richardson, I. B., Hurley, S. K. and Hynes, M. J.: Cloning and molecular characterisation of the *amdR* controlled *gatA* gene of *Aspergillus nidulans*. Mol. Gen. Genet. 217 (1989) 118-125.

Sambrook, J., Fritsch, E. F. and Maniatis, T.: Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.

Timberlake, W. E.: Cloning and analysis of fungal genes. In Bennett, J. W. and Lasure, L. L. (Eds.) More gene manipulations in fungi, Academic Press, New York, 1991, pp. 51-85.

Turner, G., Brown, J., Kerry-Williams, S., Bailey, A. M., Ward, M., Punt. P. J. and Van den Hondel, C. A. M. J. J.: Analysis of the *oliC* promoter of *Aspergillus nidulans*. In Nevalainen, H. and Penttilä, M. (Eds.), Proceedings of the EMBO-Alko workshop on Molecular Biology of Filamentous Fungi, Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, 1989, pp. 101-109.

Van den Hondel, C. A. M. J. J. and Punt, P. J.: Gene transfer systems and vector development for filamentous fungi. In Peberdy, J. F., Caten, C. E., Ogden, J. E. and Bennett, J. W. (Eds.), Applied Molecular Genetics of Fungi, Cambridge University Press, Cambridge, 1991, pp. 1-28.

Van Gorcom, R. F. M., Pouwels, P. H., Goosen, T., Visser, J., Van den Broek, H. W. J., Hamer, J. E., Timberlake, W. E. and Van den Hondel, C. A. M. J. J.: Expression of an *Escherichia coli* ß-galactosidase fusion gene in *Aspergillus nidulans*. Gene 40 (1985) 99-106.

Van Heeswijck, R. and Hynes, M. J.: The *amdR* product and a CCAAT-binding factor bind to adjacent, possibly overlapping DNA sequences in the promoter region of the *Aspergillus nidulans amdS* gene. Nucl. Acid. Res. 19 (1991) 2655-2660.

Wen, P., Groupp, E. R., Buzard, G., Crawford, N. and Locker, J.: Enhancer, repressor and promoter specificities combine to regulate the rat α-fetoprotein gene. DNA Cell Biol. 10 (1991) 525-536.

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

SUMMARY

Summary

The study of the regulation of fungal gene expression is aimed primarily at an understanding of the molecular mechanisms involved in the process of transcription of genetic information in these organisms. Research is focused on two aspects of this process: 1) Analysis of DNA sequences located upstream of the genes of interest (promoter sequences). Based on data obtained from research in other organisms, these sequences are thought to be involved in particular in determining both the sites of transcription initiation and the efficiency of transcription of genes downstream of these sequences. 2) Characterization and analysis of (regulatory) proteins interacting with these sequences.

Several strategies have been used to identify and characterize fungal promoter-(or transcription control-)sequences (reviewed in Chapter 1). Besides rather simple methods such as the comparison of upstream sequences of different genes also more sophisticated *in vitro* or *in vivo* techniques have been used. The *in vitro* approach, which is focused on the investigation of binding of regulatory protein(s) to promoter sequences, is primarily used in those cases where regulatory mutants and genes encoding regulatory proteins have been characterized.

Both sequence analysis and protein-DNA binding analysis may indicate sequence elements involved in the regulation of gene expression. However, the conclusions inferred from such an analysis should be verified by additional *in vivo* approaches.

These *in vivo* approaches include the analysis of titration of regulatory proteins and the analysis of the effects of mutations generated in promoter sequences on gene expression. Titration analysis is used to identify the sequences involved in regulation by regulatory proteins. This approach is based on the concept that introduction of multiple copies of the target sequence of a specific regulatory protein may lead to sequestering of this protein. This may result in aberrant expression of the gene of interest, which may affect the growth characteristics of such multi-copy strains.

In the research described in this thesis analysis of sequences involved in fungal gene expression was carried out by a mutation analysis. In this type of approach the effect of various mutations generated in the promoter sequences of the gene under study, are monitored by the expression of a convenient reporter gene fused to the mutated promoter sequences. Previous experiments had already shown that the

expression in *Aspergillus nidulans*, and other filamentous fungi. Therefore, the *lacZ* gene fusion strategy was adopted for our research.

As a model gene for the study of fungal gene expression by *in vivo* mutation analysis, the efficiently expressed glyceraldehyde-3-phosphate dehydrogenase-encoding gene from *A. nidulans* (*gpdA*) was chosen. Besides being a suitable model gene, the promoter sequences of this gene were expected to be very useful for the development of fungal expression vectors. Development of such vectors was and still is an important topic for the (over)expression of fungal and non-fungal genes in fungi.

The *gpdA* gene from *A. nidulans* was isolated by heterologous hybridization with one of the glyceraldehyde-3-phosphate dehydrogenase-encoding genes of *Saccharomyces cerevisiae* as a probe. The structure of the *gpdA* gene was further elucidated based on comparison of the nucleotide- and deduced aminoacid-sequences of this gene with those of glyceraldehyde-3-phosphate dehydrogenase-encoding genes from other organisms (Chapter 2). The transcribed region of the gene, which comprises 7 introns, as well as its upstream (promoter) and downstream (terminator) flanking regions were analyzed. Sequence comparison of the promoter regions with those of other glycolytic genes from *A. nidulans* revealed 2 putative transcription control sequences. An extensive C+T-rich sequence was observed in the region directly upstream of the major transcription start point. Compilation of sequence data from fungal genes indicate that similar sequences are present in most fungal genes analyzed to date (see Chapter 1 and references therein). Furthermore, comparison of the upstream sequences of *gpdA* and *pgkA* (encoding phospho-glycerate kinase) revealed an element of similar sequence around 500-600 nt upstream of the major transcription start point.

The promoter sequences of the *gpdA* gene were used for the construction of expression vectors, which may be used for efficient protein production in filamentous fungi. To demonstrate the ability of the *gpdA* upstream region to drive gene expression, this region was fused to the coding region of the *E. coli* hygromycin B phosphotransferase gene, *hph* (Chapter 3). Introduction of the resulting plasmid, pAN7-1, into various *Aspergillus* species, resulted in the expression of the *hph* gene, as was judged from the increased hygromycin B resistance level of the transformants obtained. Plasmid pAN7-1 was shown to be very useful in the development of gene-transfer systems based on hygromycin B resistance for a great variety of fungi.

To identify and characterize the sequences in the upstream region of the *gpdA*

gene involved in the regulation of gene expression, the above-mentioned strategy based on *in vivo* mutation analysis was used (Chapter 4). An important aspect of his strategy is that in all cases transformants carrying a single copy of the P_{gpdA}::*lacZ* fusion at a specific site in the fungal genome (*argB* locus) were used for urther analysis. In this way effects of copy number and the site of integration on the evel of expression can be excluded. Initially, deletions were generated in the *gpdA* promoter region by removing fragments of extending length from the upstream sequences. Subsequently, the effects of these deletions on the level of gene expression were analyzed. Based on the data obtained with this analysis it was concluded that probably all sequences involved in expression of the *gpdA* gene were located in a region of about 700 nt upstream of the major transcription start site. Subsequently, the effects of more specific deletions, removing sequences dentified on the basis of a comparison with upstream sequences of other *Aspergillus* genes, were analyzed. Deletion of a 50 bp sequence, identified by comparison of the upstream sequences of the *gpdA* genes of *A. nidulans* and *A. niger* (*gpd* box), resulted in a decrease of the level of gene expression of about 50%. Deletion of the C+T-rich region directly upstream of the major transcription start point resulted in a marked (>90%) decrease of the level of gene expression. n addition, this deletion also affected transcription initiation. The original transcription start points downstream of the deletion were no longer used. Transcription initiation primarily occurred from a minor upstream transcription start point. These data ndicate that C+T-rich sequences identified in the upstream sequences of many fungal genes, may be involved in determining the site of transcription initiation.

Further analysis was focused on the activity of the *gpd* box. Introduction of this element into the upstream region of the acetamidase-encoding gene (*amdS*) of *A. nidulans* fused to the *lacZ* gene, resulted in a 6-30 fold increase of the level of gene expression. The level of increase was shown to depend on the site of ntegration and on the orientation of the *gpd* box. Integration in the region involved n regulation of the *amdS* gene (at about 80 nt upstream of the major transcription start point) in the same orientation as in the intact *gpdA* upstream region, resulted n the highest level of gene expression. Introduction of the *gpd* box with flanking *gpdA* upstream sequences resulted in a much lower increase of the level gene expression. From the available data it is inferred that the *gpd* box comprises at least two activities, one being orientation dependent, but relatively independent of position of the *gpd* box in the upstream region; the other being functional in either

As the *amdS* gene is affected by complex nitrogen and carbon control circuits, also the effects of introduction of the *gpd* box on several of these control circuits were analyzed. From the data obtained in this analysis it is concluded that the introduction of the *gpd* box does not interfere with the interaction of the *amdS* promoter sequences with regulatory proteins involved in ω-amino acid induction, carbon catabolite repression and nitrogen metabolite repression.

SAMENVATTING

Samenvatting

Het onderzoek van de regulatie van expressie van genetische informatie in schimmels is voornamelijk gericht op de moleculaire processen die ten grondslag liggen aan de regulatie van gen-expressie op het niveau van transcriptie. De aandacht is hierbij in het bijzonder gericht op twee aspecten van de transcriptie regulatie: 1) Analyse van genetische informatie die "stroomopwaarts" van de betrokken genen is gelegen (promotor sequenties). Op basis van onderzoek van andere organismen wordt verondersteld dat deze promotor sequenties betrokken zijn bij het bepalen van de transcriptie-initiatie en de efficiëntie van transcriptie van de "stroomafwaarts" gelegen genen. 2) Karakterisering en analyse van (regulatie-) eiwitten die interacties aangaan met deze sequenties.

Voor de identificatie en verdere karakterisering van promotor sequenties zijn verschillende benaderingswijzen beschreven (zoals samengevat in Hoofdstuk 1). Naast het eenvoudigweg vergelijken van de nucleotidevolgorde van promotor gebieden behorend bij verschillende genen, zijn ook een aantal meer verfijnde benaderingen uitgewerkt. Een van deze benaderingen omvat onderzoek naar *in vitro* binding van (regulatie-)eiwitten aan promotor sequenties. Deze methode kan echter alleen goed worden toegepast als regulatie mutanten en genen die coderen voor de regulatie eiwitten, beschikbaar en gekarakteriseerd zijn.

Beide bovengenoemde methoden zijn zeer geschikt om aanwijzingen te verkrijgen welke promotor sequenties een rol spelen bij de regulatie van gen-expressie. Het is echter vrijwel altijd noodzakelijk deze aanwijzingen met een *in vivo* benaderingswijze te bevestigen.

Voor het onderzoek naar de expressie van schimmelgenen is een tweetal *in vivo* methoden beschreven, te weten, titratie-analyse en mutatie-analyse. Met behulp van titratie-analyse kunnen promotor sequenties die betrokken zijn bij de interactie met regulatie eiwitten worden geïdentificeerd. Deze analyse-methode is gebaseerd op de veronderstelling dat introductie van een groot aantal van dergelijke specifieke sequenties in een schimmelstam zal leiden tot het wegvangen van het bijbehorende regulatie eiwit. Hierdoor zal voor de expressie van genen die onder invloed staan van datzelfde regulatie eiwit een te kort aan dit eiwit ontstaan. Dit zal veelal leiden tot een veranderde expressie van deze genen, hetgeen kan resulteren in veranderde groei-eigenschappen van de betrokken schimmelstam.

gebruik gemaakt van mutatie-analyse. Hierbij worden de effecten van veranderingen aangebracht in de te onderzoeken promotor sequenties, onderzocht. In de meeste gevallen wordt hierbij van de expressie van een geschikt "reporter" gen gebruik gemaakt. Op basis van voorgaand onderzoek is gekozen voor het *Escherichia coli lacZ* gen als reporter gen.

Als model gen voor het onderzoek van de expressie van schimmelgenen is gekozen voor het glyceraldehyde-3-fosfaat dehydrogenase gen van *A. nidulans* (*gpdA*). Gelet op de efficiënte wijze van expressie van dit gen, was het de verwachting dat de promotor sequenties van dit gen ook geschikt zouden zijn voor het op efficiënte wijze tot expressie brengen van andere (niet-schimmel) genen hetgeen een belangrijk onderwerp van onderzoek is in schimmels. Het *gpdA* gen is geïsoleerd met behulp van heterologe hybridisatie met een van de glyceraldehyde-3-fosfaat dehydrogenase genen van *Saccharomyces cerevisiae*. De structuur van het *gpdA* gen is opgehelderd aan de hand van vergelijking van de nucleotiden- (en daarvan af te leiden aminozuur-) volgorde van het *gpdA* gen met die van glyceraldehyde-3-fosfaat dehydrogenase genen van andere organismen (Hoofdstuk 2). Van zowel het coderende gedeelte van het gen als van de flankerende promotor en terminator gebieden werd de basepaarvolgorde bepaald. Vergelijking van de basepaarvolgorde van het promotor gebied met die van andere schimmelgenen leidde tot de identificatie van twee mogelijk functionele onderdelen van de promotor. Het gedeelte van de promotor direct voorafgaand aan de belangrijkste transcriptie startplaats, bestaat (in de coderende DNA-streng) vrijwel volledig uit de nucleotiden C en T. Dergelijke (C+T)-rijke gebieden worden bij een groot aantal schimmelgenen aangetroffen op vrijwel dezelfde positie (zie ook referenties bij Hoofdstuk 1). Verder is op ongeveer 500-600 nucleotiden vóór de transcriptie startplaats een nucleotidevolgorde aangetroffen die overeenkomst vertoont met een gedeelte van het promotor gebied van het fosfo-glyceraat kinase gen van *A. nidulans*.

De promotor van het *gpdA* gen is vervolgens gebruikt bij de constructie van expressie vectoren, bedoeld om (niet-schimmel) genen op efficiënte wijze in schimmels tot expressie te brengen. Om te bevestigen dat deze promotor in staat is een gen efficiënt tot expressie te brengen, is in een van de expressie vectoren het *E. coli* hygromycine B fosfo-transferase gen (*hph*) gefuseerd aan de *gpdA* promotor (Hoofdstuk 3). Introductie van de aldus verkregen vector in verschillende *Aspergillus* species, resulteerde in expressie van het *hph* gen, zoals werd afgeleid uit het verhoogde hygromycine B resistentie niveau van de verkregen transformanten. Verder bleek het mogelijk met dezelfde vector transformatie systemen op basis van

hygromycine B resistentie te ontwikkelen voor een groot aantal verschillende schimmels.

Om de functionele gedeelten van de *gpdA* promotor te identificeren en verder te karakteriseren is de bovengenoemde *in vivo* mutatie-analysemethode gebruikt (Hoofdstuk 4). Een belangrijk aspect van deze analysemethode is dat steeds gebruik wordt gemaakt van transformanten die één kopie van de fusie van een mutant *gpdA* promotor en het *lacZ* reporter gen bevatten. In alle gevallen is daarbij het fusie-gen ook geïntegreerd op een vaste plaats in het genoom. Hierdoor kunnen bij vergelijking van het niveau van *lacZ* expressie in de verschillende transformanten effecten van het kopie-aantal en de integratie-plaats buiten beschouwing worden gelaten. Aanvankelijk is, met behulp van deleties waarbij steeds meer van de promotor sequenties werden verwijderd, bepaald dat voor maximale expressie ongeveer 700 nucleotiden van de vóór het gen gelegen sequenties (gerekend vanaf de belangrijkste transcriptie startplaats) nodig zijn. Meer specifieke deleties in dit gebied, gekozen op basis van vergelijking met promotor gebieden van andere schimmelgenen, heeft geleid tot de identificatie van een aantal functionele onderdelen van de *gpdA* promotor. Deletie van een 50 nucleotiden groot element, aanwezig in de promotor van zowel het *A. nidulans* als *A. niger gpdA* gen (*gpd* box) resulteerde in een afname in het niveau van gen-expressie met 50%. Deletie van het (C+T)-rijke gedeelte van de promotor resulteerde in een afname van het niveau van gen-expressie met meer dan 90%. Bovendien resulteerde deze deletie in een effect op de transcriptie initiatie. De transcriptie startplaatsen stroomafwaarts van de deletie worden niet langer gebruikt. Deze laatste resultaten wijzen op de betrokkenheid van (C+T)-rijke gebieden, zoals aanwezig in vrijwel alle schimmelgenen, bij het proces van transcriptie initiatie.

Het onderzoek is vervolgens volledig gericht op de activiteit van de *gpd* box (Hoofdstuk 5). Introductie van de *gpd* box in de promotor van het aceetamidase gen *(amdS)* van *A. nidulans*, gefuseerd aan het *lacZ* gen, resulteerde in een 6 tot 30-voudige verhoging van het niveau van expressie van het fusie-gen. Zowel de oriëntatie als de plaats van introductie waren van invloed op het niveau van gen-expressie. Het hoogste niveau van expressie werd bereikt bij introductie van de *gpd* box, in dezelfde oriëntatie als in de *gpdA* promotor, in het gedeelte van de *amdS* promotor dat betrokken is bij de regulatie van het *amdS* gen. Wanneer behalve de *gpd* box ook flankerende gebieden (uit de *gpdA* promotor) werden geïntroduceerd in de *amdS* promotor leidde dit tot een veel geringere verhoging van

geconcludeerd worden dat de *gpd* box waarschijnlijk meer dan één functioneel domein bevat, waarbij de activiteit van één van deze domeinen afhankelijk is van de oriëntatie van de *gpd* box en min of meer onafhankelijk van de plaats van introductie. De andere domein is, onafhankelijk van de oriëntatie, alleen actief is in de nabijheid van andere functionele promotor gebieden. Naast het onderzoek naar het effect van de *gpd* box op de efficiëntie van gen-expressie is ook onderzoek verricht naar het effect van de *gpd* box op de regulatie van het *amdS* gen. Uit dit onderzoek is gebleken dat introductie van de *gpd* box de interactie met de bij de expressie van het *amdS* gen betrokken regulatie eiwitten niet verhindert. Zowel de ω-aminozuur inductie, als de koolstof cataboliet repressie en de stikstof metaboliet repressie van expressie van het *amdS* gen vinden ook in aanwezigheid van de *gpd* box nog steeds plaats.

ABBREVIATIONS

A.	*Aspergillus*
Ap	ampicillin
amdS	*A. nidulans* gene encoding acetamidase
βGal	β-galactosidase
bp	base pair(s)
d	deletion
ds	double strand(ed)
E.	*Escherichia*
GPD	glyceraldehyde-3-phosphate dehydrogenase
gpd/gpdA	*A. nidulans* gene encoding GPD
HmB	hygromycin B
hph	*E. coli* gene encoding HmB phosphotransferase
kb	1000 bp
mc	multiple copies
nt	nucleotide(s)
oligo	oligodeoxyribonucleotide
PAGE	polyacrylamide gelelectrophoresis
pgk	*A. nidulans* gene encoding phosphoglycerate kinase
PolIK	Klenow (large) fragment of *E. coli* DNA polymerase I
^{R}	resistant
SDS	sodium dodecyl sulfate
ss	single strand(ed)
SSC	0.15 M NaCl + 0.015 M Na_3.citrate pH 7.6
tdh	*S. cerevisiae* gene encoding GPD
tpiA	*A. nidulans* gene encoding triosephosphate isomerase
tsp	transcription start point(s)
wt	wild type
XGal	5-bromo-4-chloro-3-indolyl-ß-*D*-galactopyranoside
[ ]	designates plasmid-carrier state.

CURRICULUM VITAE

De schrijver van dit proefschrift is geboren op 3 maart 1961 te Rotterdam-Overschie.

Na het voorspoedig doorlopen van de Ds. G. A. v. d. Hooftschool voor lager onderwijs is hij in 1972 begonnen met de middelbare-schoolopleiding aan het Marnix Gymnasium te Rotterdam. In 1978 behaalde hij het diploma voor Voorbereidend Wetenschappelijk Onderwijs (Gymnasium β), waarna in dat zelfde jaar de studie Biologie aan de Rijksuniversiteit Leiden werd begonnen. In 1981 werd het kandidaatsexamen B4 (biologie met scheikunde als 2e hoofdvak) behaald. De hieropvolgende doctorale studiefase omvatte deelname aan het onderzoek binnen; 1) de vakgroep Botanische Morfogenese (Drs. A. C. Maan / Prof. Dr. K. R. Libbenga) van de Rijksuniversiteit Leiden, 2) de vakgroep Celbiologie en Genetica (Drs. A. T. J. Bianchi / Prof. Dr. R. Benner) van de Erasmus Universiteit te Rotterdam, en 3) de sectie Recombinant DNA (Ir. R. F. M. van Gorcom / Dr. C. A. M. J. J. van den Hondel / Dr. P. H. Pouwels) van het Medisch Biologisch Laboratorium / TNO te Rijswijk. Het doctoraalexamen werd in 1984 afgelegd.

Van 1984-1985 is als gedetacheerd wetenschappelijk medewerker van de Rijksuniversiteit Leiden onderzoek verricht aan filamenteuse schimmels in het Medisch Biologisch Laboratorium / TNO. Dit onderzoek is uitgevoerd onder leiding van Dr. C. A. M. J. J. van den Hondel. Van 1985 tot heden is dit onderzoek voortgezet in dienst van TNO en heeft geleid tot de resultaten beschreven in dit proefschrift.

LIST OF PUBLICATIONS

lanegraaf, P. P. F., Punt, P. J., Van den Hondel, C. A. M. J. J., Dekker, J., Yap, W., Van Verseveld, H. W. and Stouthamer, A. H.: Construction and physiological characterization of glyceraldehyde-3-phosphate dehydrogenase overproducing transformants of *Aspergillus nidulans*. Appl. Microbiol. Biotechnol. 34 (1991) 765-771.

:olar, M., Punt, P. J., Van den Hondel, C. A. M. J. J. and Schwab, H.: Transformation of *Penicillium chrysogenum* using dominant selection markers and expression of an *Escherichia coli lacZ* fusion gene. Gene 62 (1988) 127-134.

lattern, I. E., Punt, P. J. and Van den Hondel, C. A. M. J. J.: A vector of *Aspergillus* transformation conferring phleomycin resistance. Fung. Genet. Newslett. 35 (1988) 25.

looibroek, H., Kuipers, A. G. J., Sietsma, J. H., Punt, P. J. and Wessels, J. G. H.: Introduction of hygromycin B resistance into *Schizophyllum commune*: Preferential methylation of donor DNA. Mol. Gen. Genet. 222 (1990) 41-48.

lullaney, E. J., Punt, P. J. and Van den Hondel, C. A. M. J. J.: DNA mediated transformation of *Aspergillus ficuum*. Appl. Microbiol. Biotechnol. 28 (1988) 451-454.

)liver, R. P., Roberts, I. N., Harling, R., Kenyon, L., Punt, P. J., Dingemanse, M. A. and Van den Hondel, C. A. M. J. J.: Transformation of *Fulvia fulva*, a fungal pathogen of tomato, to hygromycin B resistance. Curr. Genet. 12 (1987) 231-233.

'unt, P. J., Oliver, R. P., Dingemanse, M. A., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. Gene 56 (1987) 117-124.

'unt, P. J., Dingemanse, M. A., Jacobs-Meijsing, B. J. M., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Isolation and characterization of the glyceraldehyde-3-phosphate dehydrogenase gene of *Aspergillus nidulans*. Gene 69 (1988) 49-57.

'unt, P. J., Greaves, P. A. and Van den Hondel, C. A. M. J. J.: Linkage mapping of the *gpdA* gene of *Aspergillus nidulans*. Fung. Genet. Newslett. 37 (1990) 33.

'unt, P. J., Zegers, N. D., Busscher, M., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: Intracellular and extracellular production of proteins in *Aspergillus* under the control of expression signals of the highly expressed *A. nidulans gpdA* gene. J. Biotechnol. 17 (1991) 19-34.

Punt, P. J., Dingemanse, M. A., Kuyvenhoven, A., Soede, R. D. M., Pouwels, P. H and Van den Hondel, C. A. M. J. J.: Functional elements in the promotor regio of the *Aspergillus nidulans gpdA* gene encoding glyceraldehyde-3-phosphat dehydrogenase. Gene 93 (1990) 101-109.

Punt, P. J., Greaves, P. A., Kuyvenhoven, A., Van Deutekom, J. C. T., Kinghorn, R., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: A twin-reporter vector fc simultaneous analysis of expression signals of divergently transcribed, contiguou genes in filamentous fungi. Gene 104 (1991) 119-122.

Punt, P. J. and Van den Hondel, C. A. M. J. J.: Analysis of transcription contr sequences of fungal genes, in: Molecular Signals in Plant-Microb Communications, D. P. S. Verma ed. CRC Press, Boca Raton, FL, 1991, pp. 2 48.

Roberts, I. N., Oliver R. P., Punt, P. J. and Van den Hondel, C. A. M. J. J Expression of the *Escherichia coli* ß-glucuronidase gene in industrial an phytopathogenic filamentous fungi. Curr. Genet. 15 (1989) 177-180.

Turner, G., Brown, J., Kerry-Williams, S., Bailey, A. M., Ward, M., Punt. P. J., and Va den Hondel, C. A. M. J. J.: Analysis of the *oliC* promoter of *Aspergillus nidulans* in, Proceedings of the EMBO-Alko workshop on Molecular Biology of Filamentou Fungi, Nevalainen, H., and Penttilä, M., Foundation for Biotechnical and Industri Fermentation Research, Helsinki, 1989, pp. 101-109.

Unkles, S. E., Campbell, E. I., Punt, P. J., Hawker, K. L., Contreras, R., Hawkins, R., Van den Hondel, C. A. M. J. J. and Kinghorn, J. R.: The *Aspergillus niger niaI* gene encoding nitrate reductase: upstream nucleotide and amino acid sequenc comparisons. Gene 111 (1992) 149-155.

Van den Hondel, C. A. M. J. J. and Punt, P. J.: Gene-transfer systems and vect development for filamentous fungi, BSM Symposium Series Volume 18, Applie Molecular Genetics of Fungi, Peberdy, J. F., Caten, C. E., Ogden, J. E. an Bennett, J. W. eds., British Mycological Society, Cambridge University Pres Cambridge, UK, 1991, p1-29.

Van den Hondel, C. A. M. J. J., Punt, P. J. and Van Gorcom, R. F. M.: Heterologou gene expression in filamentous fungi, in, *More gene manipulation in fungi*, Bennet J. W., and Lasure, L. L. Eds., Academic Press San Diego, CA, 1991, p396-428.

Van den Hondel, C. A. M. J. J., Punt, P. J. and Van Gorcom, R. F. M.: Productio of extracellular proteins by the filamentous fungus *Aspergillus*, Anthonie va Leeuwenhoek 61 (1992), 153-160.

PUBLICATIONS

Van Gorcom, R. F. M., Punt, P. J., Pouwels, P. H. and Van den Hondel, C. A. M. J. J.: A system for the analysis of expression signals in *Aspergillus*. Gene 48 (1986) 211-217.

DANKWOORD

Bij het schrijven van dit dankwoord kwamen onwillekeurig herinneringen op van itnodigingen voor verjaardagsfeestjes op de lagere school: wie zullen we wel, en ie zullen we niet vragen?

Het met naam en toenaam noemen van iedereen die aan de totstandkoming van it proefschrift heeft bijgedragen is vrijwel onmogelijk. Mijn dank is er desalniettemin iet minder om. Een aantal mensen wil ik toch speciaal bedanken.

In de eerste plaats mijn ouders, die hun kinderen altijd hebben gestimuleerd in het emen van verantwoordelijkheid. Mijn opvoeding in het licht van Gods woord is nog teeds het fundament van mijn leven. Ik vindt het erg verdrietig dat mijn vader deze ag niet meer heeft kunnen meemaken.

Lieve Eline, jou wil ik bedanken voor zoveel wat met de pen niet is te beschrijven. iet gevoel dat bij me opkwam bij het zien van Hilde's eerste rondje fietsen zonder jwieltjes of Koen's imitatie van een grasmaaier komen er misschien nog wel het ichtst bij in de buurt.

Natuurlijk geen proefschrift zonder (co-)promotores. Prof. Pouwels, beste Peter, ju wil ik bedanken dat je mijn promoter wilde zijn. Verder voor je zuiver 'etenschappelijke interesse in het onderzoek, ook als het soms niet zo oepassingsgericht" was. Je kritische beoordeling van manuscripten (met een scherp eslepen potlood, zodat het in een eerste oogopslag altijd lijkt mee te vallen) 'aardeer ik nog steeds. Dr. Van den Hondel, beste Kees, jou komt mijn dank toe oor je begeleiding in mijn ontwikkeling tot een zelfstandig onderzoeker. Vanaf mijn erste dag bij het MBL is jouw ideeënrijkdom een bron waaruit ik heb geput.

Ook geen proefschrift zonder analisten. Beste Brigit, hartelijk dank voor je opzetje an mijn *gpd* bestaan. Beste Marianne, als "mijn" eerste analist heb je een zeer grote ijdrage geleverd aan zo'n beetje de helft van dit proefschrift. Ik bewonder je oorzettingsvermogen, waardoor ik je binnenkort als Dr. Dingemanse kan begroeten. ieve Anneke, hartelijk dank voor jouw bijdrage aan dit proefschrift. Zonder je auwgezetheid en ervaring in een groot aantal technieken was er vast veel vaker prake geweest van "unreproducible results". Ik vind het erg fijn dat je mijn paranimf ilde zijn.

In mijn geval ook geen proefschrift zonder tijdelijke-, gast-, student-, en stagiair-nedewerkers, waarvan ik er drie in het bijzonder wil noemen. Beste Cas, door jouw ijdrage aan het onderzoek, zoals beschreven in het laatste hoofdstuk, is het nogelijk geworden dit proefschrift dan ook eindelijk te schrijven. Ik heb onze [illegible] altijd gewaardeerd, misschien vooral wel omdat je niet iemand bent

3. Plasmid pAN7-1 has become more than we anticipated. Ronald Soede wil il bedanken voor zijn uitstekende bijdrage aan het onderzoek, zoals beschreven ir hoofdstuk 4.

Verder wil ik alle "Aspen" en "Lacto's" bedanken voor hun bijdrage aan d plezierige sfeer in onze sectie, waardoor ik elke dag mijn fietstochtje van achter d kassen zonder tegenzin maak. Op het gevaar af voor te trekken -ik denk weer aa dat verjaardagsfeestje- wil ik nog een paar mensen apart noemen. Beste Robert, j was mijn eerste directe chef. Het feit dat jij wel een student kon gebruiken was mi of meer het begin van dit proefschrift. Beste Ineke, al jaren mijn kamergenoot, jouv interesse ook voor dingen buiten het werk, leverde altijd genoeg stof tot gesprek op We zijn nog niet uitgepraat. Beste Wim en Cora, ook jullie wil ik een plaatsje gevei in dit dankwoord. Onze koppels gingen een tijdlang gelijk op. Met dit verschil; julli werden gesponsord. Dr. Frits Berends wil ik bedanken voor de zorg die hij altij besteed heeft en besteedt aan het reilen en zeilen van onze afdeling Biochemie.

Tenslotte, Jeane, Henny, Michel, Ronald, Jan, Frans, Ria, Rinus, Ferry, Nel, Floo en alle andere medewerkers van de ondersteunende diensten wil ik hartelij bedanken voor hun ondersteuning.

STELLINGEN

1

De aanwezigheid van TATA box-achtige sequenties in de promoter regio van schimmel genen is geenszins het bewijs dat deze sequenties ook een functioneel onderdeel van deze promoter zijn.

Dit proefschrift.

2

De conclusie van Dunn-Coleman et al. (1991) dat, in de gebruikte chymosine over-producerende transformanten geen sprake is van een limitatie op transcriptie-niveau, wordt door de resultaten van de door de auteurs beschreven experimenten eerder tegengesproken dan ondersteund.

Dunn-Coleman et al., Bio/Technology 9 (1991) 976-981.

3

Bij de analyse van de Taka-amylase (*Taa*) gen-familie van *Aspergillus oryzae* door Tsukagoshi et al. (1989) wordt in een van de *Taa* genen (*Taa-G1*) ten onrechte een afwijkende positie van een van de introns verondersteld. De beschreven resultaten geven immers aan dat de geïsoleerde cDNA's corresponderen met een ander lid uit de gen-familie.

Tsukagoshi et al., Gene 84 (1989) 319-327.

4

De correctie van de structuur van het 5' uiteinde van het *A. nidulans alcR* gen (Felenbok et al., 1988; Kulmburg et al., 1991) maakt de conclusies van Felenbok et al., (1989) met betrekking tot het werkingsmechanisme van dit regulatie gen uiterst discutabel.

Felenbok, B. et al., Gene 73 (1988) 385-396.
Kulmburg, P. et al., FEBS Lett. 280 (1991) 11-16.
Felenbok, B. et al., Proc. EMBO-Alko workshop on Molec.
Biol. of Filamentous Fungi, Helsinki, 1989, 73-83.

5

Op basis van de gepresenteerde resultaten worden door Andrianopoulos en Hynes voorbarige conclusies getrokken met betrekking tot de structuur van het 5' uiteinde van het *amdR* gen.

vector, draagt, gelet op de afleiding van dit woord (Gr. myth. $\phi\alpha\sigma\mu\alpha$, phasma= spook, fantoom) niet bij tot de maatschappelijke aanvaarding van het gebruik van moleculair biologische technieken.

Kahn, M. and Helinski, D. R., Proc. Natl. Acad. Sci. USA 75 (1978) 2200- 2204.
Brenner et al., Gene 17 (1982) 27-44

7

De veelal huiverige houding ten opzichte van (natuur)wetenschappelijk onderzoek binnen orthodox christelijke kring, gaat voorbij aan het feit dat de werkelijkheid, d.i. het onderwerp van dit onderzoek, zijn oorsprong heeft in God.

Keizer, A. Opbouw 35 (1991) 242-243
Verhey, J.W. Opbouw 35 (1991) 354-356

8

De, in het nieuwe logo van de Hoofdgroep Gezondheidsonderzoek gebruikte, lettercombinatie (GO) getuigt, gelet op de huidige uitstroom van medewerkers, van een sterk visionair inzicht bij de bedenkers van dit logo.

9

Het verdient aanbeveling bij de inrichting van (bank)loketten, de bevestigingsplaats en (met name) de lengte van het bevestigingskoord van het aldaar aanwezige schrijfgerei niet alleen af te stemmen op rechtshandige cliënten.

10

Het gebruik om een proefschrift te besluiten met een nawoord doet ten onrechte vermoeden dat promoveren het einde is.

Stellingen behorend bij het proefschrift "Functional elements in the promoter region of the glyceraldehyde-3-phosphate dehydrogenase gene of *Aspergillus nidulans*" van Peter J. Punt