

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

**Characterization of the alpha-Mannosidase Gene Family
in Filamentous Fungi**

by

Caleb Joshua Eades
B.Sc., University of British Columbia, 1994

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Biology

We accept this dissertation as conforming
to the required standard

~~Dr. W.E. Hintz, Supervisor (Department of Biology, University of Victoria)~~

~~Dr. B.F. Koop, Departmental Member (Department of Biology, University of Victoria)~~

~~Dr. N. Sherwood, Departmental Member (Department of Biology, University of Victoria)~~

~~Dr. C. Upton, Outside Member (Department of Biochemistry, University of Victoria)~~

~~Dr. C. Breuil, External Examiner (Department of Wood Science, University of B.C)~~

© Caleb Joshua Eades, 2001
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by
photocopying or other means, without the permission of the author

Supervisor: Dr. William E. Hintz

ABSTRACT

Protein N-glycosylation, which is ubiquitous in eukaryotes, is a complex pathway involving numerous gene families. Early stages of the glycosylation pathway show a high degree of conservation among eukaryotes, yet diversification of the number and size of gene families involved in the later stages of the pathway has led to the evolution of increasingly complex N-glycan structures and functions in various organisms. The overall purpose of this research project has been to characterize the diversity within the α -mannosidase gene family of filamentous fungi. The α -1,2-mannosidases are involved in mannose removal in the intermediate stages of the N-glycosylation pathways, and diversification of this gene family may have provided the first significant divergence in these pathways among major lineages.

Four novel α -mannosidases were identified and characterized from the filamentous fungus *Aspergillus nidulans*. These genes were designated Class II α -mannosidase, Class I α -1,2-mannosidase IA, Class I α -1,2-mannosidase IB and Class I α -1,2-mannosidase IC, based on their similarity to other Class I and Class II α -mannosidase sequences. The Class II α -mannosidase was highly similar to the rat ER/cytosolic and yeast vacuolar Class II α -mannosidases, and these three proteins formed a phylogenetically distinct subgroup, Class IIC. The Class I enzymes were highly related to each other, and to other fungal Class I α -1,2-mannosidases. Phylogenetic analysis indicates these genes duplicated and diverged subsequent to the divergence of fungi from insects and mammals. In addition to this research on *A. nidulans*, a single Class I α -1,2-mannosidase was identified and characterized from the Dutch Elm pathogen, *Ophiostoma novo-ulmi*, which was highly related to the *A. nidulans* Class I α -1,2-mannosidase IA and IC enzymes, and less so to the *A. nidulans* Class I α -1,2-mannosidase IB.

Analysis of the function and/or biochemical properties of these enzymes was examined using several methods. Disruption and overexpression of the *A. nidulans* Class IIC α -mannosidase did not have any noticeable effect on the growth or morphology of

the organism, indicating that this gene was not essential for growth. Biochemical characterization of the *A. nidulans* Class I α -1,2-mannosidase IC was initiated by recombinant secretion of the enzyme into culture media. Successful expression of the enzyme showed that the α -1,2-mannosidase IC did not exert any cytotoxic effects when overexpressed, suggesting that high levels of expression and purification should be feasible. Finally, disruption of the Class I α -1,2-mannosidase from *O. novo-ulmi* slightly altered the morphology of the organism, but was not lethal. The possible presence of multiple Class I α -1,2-mannosidases in this organism could explain the non-lethality of this mutation.

Elucidation of the N-glycosylation pathways of *A. nidulans* may be useful in host strain improvement for heterologous protein expression systems. Modulation of the N-glycosylation pathways to produce specific N-glycan structures would increase the utility of the host for the production of human therapeutic glycoproteins which require these N-glycans for efficacy. Additionally, investigation of the genetic components of the N-glycosylation pathways of the Dutch Elm pathogen may provide global antifungal targets with broad applicability in other fungi.

Examiners:

Dr. W.E. Hinz, Supervisor (Department of Biology, University of Victoria)

Dr. B.F. Koop, Departmental Member (Department of Biology, University of Victoria)

Dr. N. Sherwood, Departmental Member (Department of Biology, University of Victoria)

Dr. C. Upton, Outside Member (Department of Biochemistry, University of Victoria)

Dr. C. Breuil, External Examiner (Department of Wood Science, University of B.C)

Table of Contents

Abstract	ii
Table of Contents	iv
List of Tables	vi
List of Figures	vii
Acknowledgements	ix
CHAPTER 1 - General Introduction (excerpted from <i>Biotechnology and Bioprocess Engineering 5: 227-233, 2000</i>)	1
CHAPTER 2 - Identification and analysis of a Class 2 α -mannosidase from <i>Aspergillus nidulans</i> (<i>Glycobiology 8: 17-33, 1998</i>)		
2.1 Abstract	31
2.2 Introduction	32
2.3 Results	34
2.4 Discussion	55
2.5 Materials and Methods	59
CHAPTER 3 - Characterization of the Class I α -mannosidase gene family in the filamentous fungus <i>Aspergillus nidulans</i> (<i>Gene 255:25-34, 2000</i>)		
3.1 Abstract	67
3.2 Introduction	68
3.3 Materials and Methods	70

3.4 Results	75
3.5 Discussion	86
CHAPTER 4 - Expression and Secretion of the Class I α-1,2-mannosidase IC from <i>Aspergillus nidulans</i>	
4.1 Introduction	90
4.2 Materials and Methods	93
4.3 Results	98
4.4 Discussion	102
CHAPTER 5 - Cloning, Sequence Characterization, and Disruption of Class I α-1,2-mannosidase IA from <i>Ophiostoma novo-ulmi</i>	
5.1 Introduction	108
5.2 Materials and Methods	114
5.3 Results	122
5.4 Discussion	132
CHAPTER 6 - General Discussion	138
LITERATURE CITED	151

List of Tables

TABLE I: Oligonucleotide primer sequences 35

List of Figures

FIGURE 1: Types of N-glycans produced in eukaryotes	5
FIGURE 2: N-glycan biosynthetic pathway in the endoplasmic reticulum (ER) and Golgi apparatus	17
FIGURE 3: Full length sequence of the <i>Aspergillus nidulans</i> α -mannosidase.....	37
FIGURE 4: Sequence alignment of <i>Aspergillus nidulans</i> α -mannosidase, rat ER/cytosolic α -mannosidase, and yeast vacuolar α -mannosidase	42
FIGURE 5: Disruption/induced expression of the α -mannosidase gene.....	46
FIGURE 6: Confirmation of gene disruption	48
FIGURE 7: Expression of cytosolic α -mannosidase in <i>Aspergillus nidulans</i>	49
FIGURE 8: Matchbox sequence similarity matrix	52
FIGURE 9: Dendrogram prepared from multiple alignment of 18 α -mannosidase sequences	54
FIGURE 10: Kyte-Doolittle hydropathy plot of (A) <i>A. nidulans</i> α -mannosidase IA, (B) <i>A. nidulans</i> α -mannosidase IB and (C) <i>A. nidulans</i> α -mannosidase IC predicted amino acid sequences	81
FIGURE 11: Multiple sequence alignment of three α -1,2-mannosidase putative protein sequences from <i>A. nidulans</i>	82
FIGURE 12: Phylogram showing sequence relationships of Class I α -mannosidases	85
FIGURE 13: Expression vector for <i>A. nidulans</i> α -1,2-mannosidase IC	96
FIGURE 14: Secretion of <i>A. nidulans</i> α -1,2-mannosidase IC into culture media	101

FIGURE 15: Construction of <i>O. novo-ulmi</i> α -1,2-mannosidase disruption vector	120
FIGURE 16: Copy number of <i>O. novo-ulmi</i> Class I α -1,2-mannosidase	124
FIGURE 17: Kyte-Doolittle hydropathy plot of <i>O. novo-ulmi</i> predicted amino acid sequence	127
FIGURE 18: Diagnostic PCR screening of <i>O. novo-ulmi</i> disruption transformants	129
FIGURE 19: Phylogram showing sequence relationships of Class I α -mannosidases including <i>O. novo-ulmi</i> α -1,2-mannosidase	131
FIGURE 20: N-glycosylation synthesis pathway	142

Acknowledgements

I would like to thank various people in the Biology Department for their technical assistance and advice, which helped me achieve my research goals. In particular, I would like to thank Dr. Ben Koop and the students and staff in his lab for their assistance with automated DNA sequencing, multiple sequence alignments and generation of phylogenetic trees. I would also like to thank Graham Sinclair for his assistance in the protein expression work and for advice on various aspects of molecular biology, research and life in general. A very special thanks is owed to the Graduate Secretary Eleanore Floyd., who was always available to fix those seemingly unsolvable problems.

Thanks to various granting agencies which have made my research possible, especially to the Natural Sciences and Engineering Research Council (NSERC), and to the various funding sources made available through the Biology Department Scholarships and Fellowships program.

Thanks to my fellow lab colleagues, present and former, who have made the lab such a fun place to be, and a stimulating environment - particularly Brad Temple, Kirk Leifso, Mike Pinchback, Elisa Becker, Holly Williams and Paul de la Bastide. You have all provided a great deal of scientific inspiration and intellectual discussion with a healthy dose of practical jokes and general mischief to keep things from getting too serious. I have also greatly appreciated the support of my good friends Carolyn Bergstrom, Louise Hahn and David Western. You all mean a great deal to me.

It has been a tremendous pleasure working with Dr. William Hintz, my academic supervisor and mentor for the past 6 years. Will, you have been a great inspiration to me and it has been a joy learning from you how to be a good, inquisitive scientist. I have learned a great many things from you which I will carry with me during my career. You have been most generous in your support, both academic and financial. You have also been a good friend. Many thanks.

I would like to thank my mother Suzanne and my father John. You have always given me so much support and been so proud of my achievements. I am fortunate to have such caring parents. And to my older brother Luke, who has always had such interest in my academic and intellectual pursuits. I have always looked up to you. And Jessica, my younger sister, you are such a good person. I've drawn strength from my family and am glad to have such strong bonds with all of you.

Most of all, I would like to thank my partner Lori. Your support and love have made this all possible. We look out on the world together, shoulder to shoulder. I am yours, you are mine, we are what we are. Finally, to my children, Jamie and Ember - you bring joy to my life and are inspirations, both of you.

CHAPTER 1 – General Introduction (Excerpted from: Eades, C.J. and W.E. Hintz.

2000. Characterization of the α -mannosidase gene family in filamentous fungi: N-glycan remodelling for the development of eukaryotic expression systems. *Biotechnology and Bioprocess Engineering* 5: 227-233)

Protein glycosylation

Protein N-glycosylation is an important post-translational modification, in which oligosaccharides (glycans) are covalently linked to specific amino acids of proteins (Peberdy, 1994). N-glycosylation can have major effects on the structure and function of proteins, influencing stability, secretion, antigenicity, pharmacokinetics and biological activity (Goochee *et al.*, 1991; Karlsson, 1991; Opdenakker *et al.*, 1993). N-glycans also have significant roles in intracellular targeting, host-pathogen interactions, and other recognition events. Protein N-glycosylation involves a series of complex reactions resulting in an oligosaccharide attached to an asparagine (Asn) residue of the polypeptide. Two broad classes of N-glycans are found in mature glycoproteins - complex and high-mannose oligosaccharides. High-mannose N-glycans contain only mannose and N-acetylglucosamine, while complex N-glycans also contain galactose, sialic acid, and occasionally fucose (Knight, 1989), and contain the largest amount of structural variation (Kobata, 1992).

Diversity in glycosylation is found between species, within populations of the same species, among different cell types in an organism, and between different proteins within the same cell (Gagneux and Varki, 1999). The significance of this diversity is evident from

the conservation of numerous glycosidases and glycosyltransferases that are responsible for N-glycan synthesis in the endoplasmic reticulum (ER) and Golgi apparatus. This process is well characterized in mammals (reviewed in Kornfeld and Kornfeld, 1985) and yeast systems (reviewed in Herscovics, 1999), but remain poorly characterized in filamentous fungi. Research in filamentous fungi includes characterization of N-glycans produced by various fungal species (reviewed in Maras *et al.*, 1999), and the characterization of several of the enzymes involved in the N-glycan biosynthesis pathway. Further characterization of the N-glycosylation pathway in filamentous fungi is required to elucidate the synthesis and processing of N-glycans in these organisms, and for comparison to other eukaryotic systems.

Evolution and function of protein glycosylation

The remarkable complexity and diversity of the glycosylation machinery suggests that its products, the protein linked oligosaccharides, perform an important biological role (Gagneux and Varki, 1999). The sequence conservation, spanning several kingdoms, of the numerous glycosidases and glycosyltransferases, and the deleterious consequences of genetic mutations in the genes that encode these enzymes also points to significant role for this pathway (Campbell *et al.*, 1995; Chui *et al.*, 1997; Gagneux and Varki, 1999; Kornfeld, 1998; Li *et al.*, 1997). Numerous roles have been demonstrated for protein glycans, ranging from specific receptor mediated recognition type functions to more general physical and chemical roles (Drickamer and Taylor, 1998; Geisow, 1992; Lis and Sharon, 1993; Opdenakker *et al.*, 1993; Paulson, 1989; Rademacher *et al.*, 1988; Varki, 1993 and references therein). Analysis of the biosynthesis pathways for protein glycans,

and N-glycans in particular provides insight into the evolution of this pathway (Drickamer and Taylor, 1998; Gagneux and Varki, 1999).

Considering the large variety of sugar types (mannose, glucose, galactose, etc.), linkage types (α 1-2, α 1-3, β 1-4, etc.), and branch lengths that are possible, the number of potential glycan structures is staggering. In nature, the types of glycans represent a much smaller subset of this number (Fukuda, 1994; Kobata, 1992; Lis and Sharon, 1993). The types of glycans found on a particular glycoprotein often vary from site to site, and can also vary at the same site on different protein molecules (Lis and Sharon, 1993; Rademacher *et al.*, 1988). This 'microheterogeneity' of glycoproteins leads to the production of a number of different 'glycoforms' which may have different physical and biochemical properties, leading to functional diversity (Dwek, 1998; Lis and Sharon, 1993). The set of glycoforms may vary from cell type to cell type, and can also be affected by alteration of the environmental state of a given cell type (Andersen and Goochee, 1994; Liu, 1992). This variation is often attributed to differences in the cellular composition of glycosidase and glycosyltransferase enzymes which process and synthesize N-glycans in the ER and Golgi apparatus, differences that may be a result of genetic composition and/or expression levels of such enzymes (Cumming, 1991; Opdenakker *et al.*, 1993; Paulson, 1989). This type of heterogeneity allows for specific modulation of the various properties of glycoproteins which are not template driven, and can thus be responsive to cellular environment.

Despite the 'microheterogeneity' found on most glycoproteins and the potential number of carbohydrate structures that are theoretically possible, N-glycans tend to show some regularity of structure, which likely reflects functional constraints upon the evolution

of the N-glycan biosynthetic pathway. In mammalian cells, the N-glycans consist of a series of linked sugars, which are classified as 'core sugars' and 'terminal branches' (Figure 1) (Drickamer and Taylor, 1998; Paulson, 1989). The core sugars establish the basic branching pattern and are generally the same for all N-glycans and are also well conserved between major evolutionary lineages. The terminal sugars vary between N-glycans on the same protein, on different proteins within the same cell, or on proteins from different cell types (Drickamer and Taylor, 1998; Lis and Sharon, 1993; Paulson, 1989; Rademacher *et al.*, 1988). In general, more complex organisms such as mammals tend to add sugars such as galactose, GlcNAc, N-acetylgalactosamine (GalNAc), fucose and sialic acid, producing complex or hybrid type N-glycans (Figure 1). While these terminal additions can be found in large variety in complex organisms, simpler organisms contain less diverse terminal additions and produce mainly high mannose N-glycans, or modified versions thereof. Hemiascomycetous yeasts, for example, add large numbers of mannose residues (Herscovics and Orlean, 1993), whereas filamentous fungi add mainly mannose, but in some cases, galactose (Maras *et al.*, 1999) or sialic acid (Alviano *et al.*, 1999). The conservation of the core sugar structure among various lineages points to a more ancestral origin for these structures, while the increasing elaboration of the terminal branches over time reflects possible evolution and increased complexity of the N-glycan biosynthesis pathway (Drickamer and Taylor, 1998).

The N-glycan biosynthetic pathway involves three distinct stages in which can be seen evidence for the types of sugars (core vs. terminal) that are found on fully processed N-glycans. In the first stage, a large precursor oligosaccharide containing three glucose, nine mannose, and two N-acetylglucosamine (GlcNAc) is built on a membrane bound

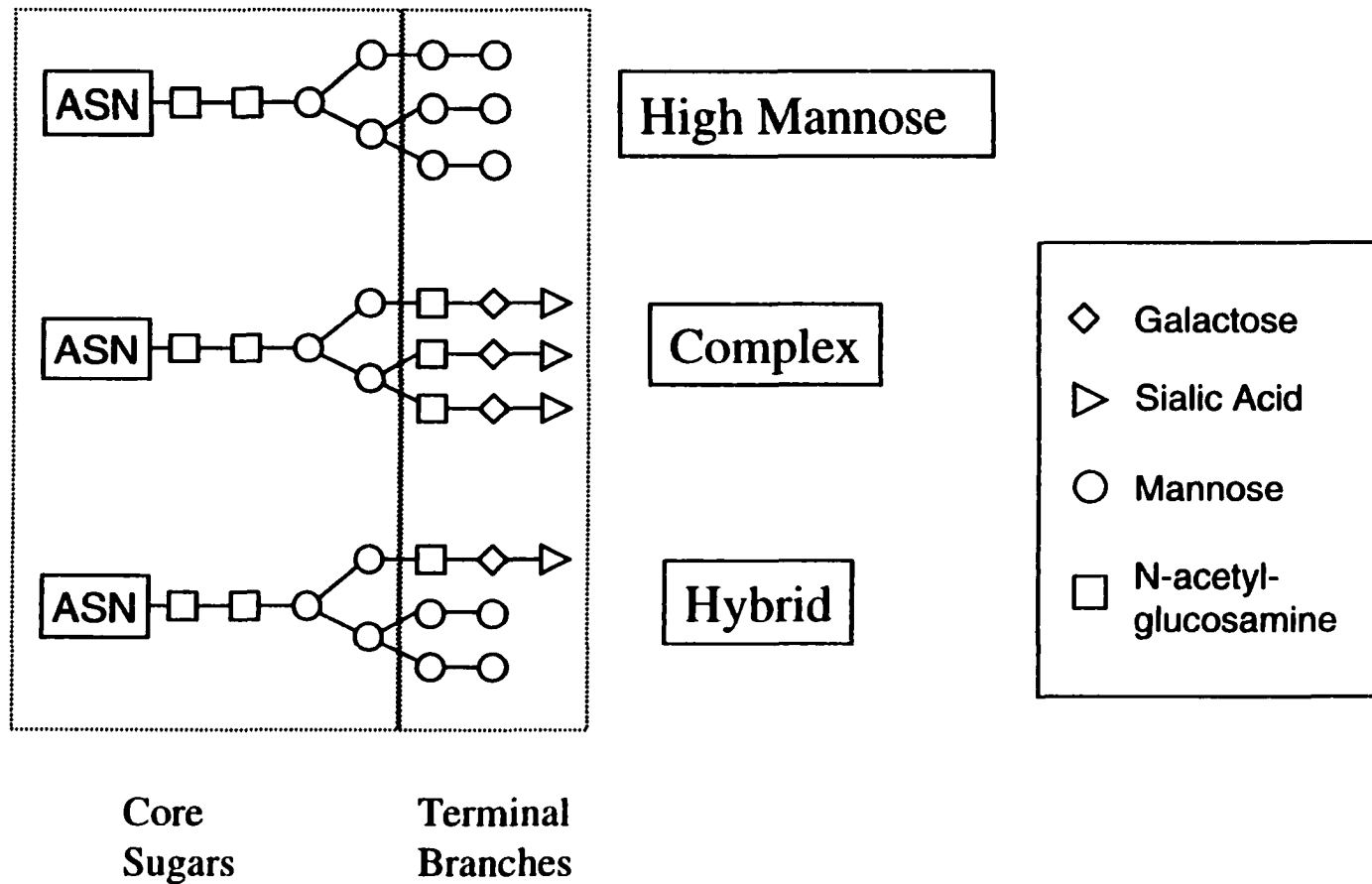


FIGURE 1: Types of N-glycan structures produced in eukaryotes. High mannose N-glycans contain only mannose in the terminal branches, whereas complex N-glycans have terminal branches containing N-acetylglucosamine, galactose and/or sialic acid. Hybrid N-glycans mannose share properties of high mannose and complex N-glycans

carrier and transferred *en bloc* to nascent polypeptides. During the second stage, glucose and some of the mannose are removed. In the final stage, various sugar molecules can be sequentially added to produce a diversity of final products (Kornfeld and Kornfeld, 1985). The production of final products thus involves transfer of a precursor oligosaccharide, removal of sugar molecules to produce the 'core' oligosaccharide, after which terminal sugars are added to the core. This circuitous and somewhat contradictory route suggests that perhaps the initial two stages represent a primordial pathway which produces the ancestral 'core', and the terminal sugar addition in the last stage of the pathway evolved later (Drickamer and Taylor, 1998). It has been suggested that these two distinct phases of the glycosylation pathway may reflect differing and evolving roles of N-glycans. In simpler organisms, such as yeasts, the high-mannose N-glycans serve a primarily structural role in the cell wall, while in more complex organism, the role of N-glycans is much more diverse and may be fine tuned by addition of various terminal branches (Gagneux and Varki, 1999; Marth, 1994; Paulson, 1989).

N-glycan roles – Folding, secretion and protease protection

It has been proposed that the attachment of large hydrophilic glycans to polypeptides in eukaryotes could have imparted specific physical properties to proteins which were selectively favorable (Drickamer and Taylor, 1998). It is suggested that perhaps the single most general role of oligosaccharides is as an aid in folding of the nascent polypeptide chain and in stabilization of the mature glycoprotein (Lis and Sharon, 1993). Glycosylation occurs as the protein is being synthesized and folded into its final form. The addition of large sugar chains could significantly affect the folding pattern of

these glycoproteins. Studies which have utilized the glycosylation inhibitor tunicamycin, as well as site-directed mutagenesis studies involving the removal of N-glycan attachment sites, have shown that reduction or complete removal of N-glycans often results in significant effects on the secretion levels and folding of many glycoproteins (Caplan *et al.*, 1991; Hickman and Kornfeld, 1978; Orlean, 1992; Parodi, 2000; Rasmussen, 1992; Riederer and Hinnen, 1991; Taylor and Wall, 1988). Deglycosylation increased the surface hydrophobicity of many of these proteins and caused them to assume more compact structures. As a result, many of these improperly glycosylated proteins aggregate in the ER and/or Golgi, while others are simply degraded. The effects of deglycosylation varies from protein to protein, and even from site to site within a protein. For example, selective removal of the N-glycan chains of human protein C revealed that the N-glycan found at Asn97 was the most important N-glycan for efficient secretion of this protein, whereas removal of other N-glycans from this protein had much less dramatic effects (Grinnell *et al.*, 1991).

Protein folding in the ER is facilitated by a number of molecular chaperones, such as the classical chaperones BiP (immunoglobulin heavy chain binding protein)/glucose-regulated proteins (GRPs), the unconventional chaperones calnexin (CNX) and calreticulin (CRT), and proteins which facilitate disulfide bond formation (Parodi, 2000). The effects of protein N-glycans in folding are not limited to providing the large hydrophilic groups which help maintain the glycoproteins in solution and prevent aggregation and degradation, but also direct further recognition and processing events. As newly synthesized proteins enter the ER, they are N-glycosylated with a large oligosaccharide molecule containing 3 glucose, nine mannose, and two GlcNAc molecules. Glucosidases

sequentially remove the glucose molecules in the ER. After the removal of the first two glucose molecules, the glycoprotein is retained in the ER by the interaction of the monoglucosylated oligosaccharide with CNX and CRT. It is proposed that this retention increases the interaction time of the glycoproteins with BiP/GRP allowing correct folding of the protein (Hammond *et al.*, 1994; Parodi, 2000). The CNX and CRT molecules themselves may also interact with the protein to specifically facilitate folding by specific molecular interactions. Glycoprotein-CNX/CRT interaction also decreases the folding rate and helps maintain the glycoproteins in solution for the prevention of protein aggregation, thus increasing the fidelity of the process (Hammond and Helenius, 1994; Parodi, 2000).

It has recently been proposed that protein N-glycans may actually be further involved in the transport of glycoproteins through the secretory apparatus, and in sorting of these glycoproteins to their ultimate cellular or extracellular destinations (reviewed in Hauri *et al.*, 2000). In addition to the previously mentioned CNX and CRT lectins involved in protein folding in the ER, several other lectins have been found in the secretory pathway. In mammalian systems, ERGIC-53 is a mannose-specific lectin which is involved in protein transport from the ER to the ER-Golgi intermediate compartment (ERGIC). Once glycoproteins are properly folded and glucose is fully removed, they no longer bind CNX:CRT, but the exposed high mannose N-glycans would be recognized by ERGIC-53, which would move these protein on to the ERGIC. It was shown that trimming of N-glycans by α -1,2-mannosidase is not required for ERGIC-53 lectin association or dissociation (Appenzeller *et al.*, 1999). Another mannose-specific lectin, VIP36, has also been found which is believed to be involved in sorting and transport of glycoproteins in the Golgi and possibly from the Golgi to the cell surface (Fiedler and

Simons, 1996). This lectin is specific for N-glycans containing as few as 5 mannose units (Hara-Kuge *et al.*, 1999; Yamashita *et al.*, 1999).

Certain proteins which are destined for intracellular locations also use N-glycan binding lectins for sorting. The best example of this is the mannose-6-phosphate receptors, which target proteins to the lysosome (Kornfeld, 1987). The mannose-6-phosphate receptor recognizes specific N-glycans containing six mannose groups and two terminal phosphate groups. Proteins containing these N-glycans are sequestered by the receptors and moved to the pre-lysosomal area. Upon dissociation from the receptor, the glycoproteins then move to the lysosome. In addition to cellular targeting, carbohydrates may function in targeting secreted proteins to specific areas of the organism, perhaps in conjunction with glycoprotein cell-surface receptor in the target cells (Paulson, 1989). This type of lectin-specific transport and sorting of glycoproteins in the secretory apparatus may explain why N-glycan processing first proceeds through a stage of sugar trimming reactions, prior to further elongation. These trimming reactions act as molecular signals for folding and transport, and it is only subsequent to this action that the N-glycans are free to be modified for other purposes. It should also be noted that many proteins appear to be able to fold and transport effectively in the absence of N-glycans. It is thus likely that this lectin-mediated folding and transport mechanism is glycoprotein specific (Hauri *et al.*, 2000).

Other studies have more directly illustrated the effects of glycosylation on protein folding and structure. Imperiali and Rickert (1995) observed two glycoproteins that assume a more compact, folded conformation upon glycosylation. The technique employed in this study was temporally sensitive, allowing the observation of conformation

changes at time intervals relative to that of protein synthesis. Rudd *et al.* (1995) have also shown structural changes associated with glycosylation in ribonuclease, plasminogen, and tissue plasminogen factor. Among other things, glycosylation increased the global dynamic stability of the proteins. This study also showed that ribonuclease is protected from the proteolytic action of pronase, which likely is due to shielding by carbohydrate side chains causing hindrance of the pronase, disallowing it access to the protein itself.

Dwek (1995) showed that removal of two galactose residues from one N-glycan in the F_c region of the IgG molecule destabilizes the hinge conformation of the F_c region of the antibody. Such effects may be the result of hydrogen bond formation, or steric interactions of the oligosaccharides with the protein. The fact that glycosylation is a co-translational event seems to indicate that such conformational effects must be exerted during the initial folding of the protein. Such studies demonstrate the significant specific and non-specific effects that N-glycans can exert on proteins during folding and secretion. The non-specific effects exerted by N-glycans are often due to the core sugar residues, and do not seem to depend on which particular sugars are attached.

N-glycan roles – Biological activity

Glycosylation can modulate the biological activity of proteins in many ways. Direct effects generally involve interaction of the oligosaccharide with the substrate molecule, by stabilizing the enzyme-substrate interaction, or by actually altering the enzyme activity. The glycans may also indirectly affect protein activity by altering such things as protein conformation, protein secretion, serum clearance, or protease susceptibility of the protein. Determining causality in this relationship can be very

difficult, especially as the complexity of such effects may cause conflicting results when comparing *in vitro* and *in vivo* studies.

There are few examples of oligosaccharides *directly* affecting the biological activity of a protein (ie. directly involved in the biological function of the protein). This may be due to the difficulty in pinpointing the exact source of such effects. Rudd *et al.* (1995) show that tissue plasminogen activator seems to be directly affected by variation in glycosylation. Occupancy of a particular N-glycan site at Asn-184 significantly decreases the activity of the protein towards its substrate, fibrin. It is suggested that this inhibition is due to decreased binding of the substrate. The authors also show that occupancy of other glycosylation sites in the protein can affect conformation of the protein, which may have effects on *in vivo* biological activity including serum clearance.

The roles of N-glycans for the *in vitro* and *in vivo* activities of erythropoietin have also been assessed (reviewed in Geisow, 1992). Studies have shown that glycosylation can affect the secretion, biological activity, and stabilization of the erythropoietin. Removal of the terminal sialic acids from the mature erythropoietin actually increased the *in vitro* activity of the protein, while removal of entire N-glycans resulted in significant reduction of *in vitro* activity (Dordal *et al.*, 1985; Takeuchi *et al.*, 1990) and had significant effects on secretion of the protein (Delorme *et al.*, 1992; Dube *et al.*, 1988). These *in vitro* studies do not accurately reflect the *in vivo* situation in which glycosylation is of paramount importance for stability and serum half-life of erythropoietin (Delorme *et al.*, 1992; Drickamer, 1991; Geisow, 1992; Takeuchi *et al.*, 1989). Although removal of sialic acid from these proteins can increase the *in vitro* activity, these asialo-glycoproteins are

removed from the bloodstream by a hepatic asialo-receptor, thus abolishing *in vivo* activity.

N-glycan roles - Receptor mediated recognition

Carbohydrate groups can also be involved in important biological recognition events. The high degree of structural diversity of these groups makes them ideal as specific recognition determinants in such events as intracellular protein-protein interactions, protein targeting, cell-cell interaction (ie. neural adhesion, host-pathogen interaction, antigenicity determinants). These signals may be extremely important in development and differentiation of cells, and in the numerous cellular interactions that are essential for life in complex organisms.

The sugar chains of glycoproteins and glycolipids are a major feature of cell surfaces. Since carbohydrates are such efficient carriers of information, it seems logical that these groups would act as recognition determinants for the attachment of other cells, microorganisms, hormones, antibodies or lectins. This is indeed the case, and there are numerous documented examples supporting this hypothesis. The adhesion of bacterial organisms to cells via carbohydrate receptors has been known for years. Bacteria are able to recognize and bind to very specific cell types. For example, whereas *Escherichia coli* is a common cause of urinary tract infections, it is seldom found in the upper respiratory tract. Likewise, group A *Streptococci* are common in the upper respiratory tract, but are seldom found in the urinary tract. It has been found that alteration of the carbohydrate surface structure of cells will change their specificity for microbial adhesion. Changing the

terminal structures, and sometimes the internal structures of specific surface glycans of mucosal cells will prevent the adhesion by the fimbriae of *E. coli* (Sharon and Lis, 1993).

Surface carbohydrates play a similar role in viral infection, plant toxicity, and symbiosis (Varki, 1993). As such, researchers have focused upon these sugars as potential targets for prevention and treatment of certain diseases. For example, inhibitors of adhesion may be designed through competition for binding sites. Paulson (1989) described the design of an inhibitor of the influenza virus using the detailed information about the sialic acid receptor of the virus. It was postulated that injection of a free receptor analogue would tend to occupy the binding sites, and prevent the adhesion of the virus. Surface carbohydrates could also mask the binding sites involved in viral infection. Such a case is described (Varki, 1993), in which the addition of a single O-acetylerster to the terminal sialic acid of the influenza receptor prevents the binding of the virus. This interplay between microbial recognition determinants and the masking of such determinants has many evolutionary implications

The level of involvement of protein linked carbohydrates in molecular recognition events varies. For example, the carbohydrate may be involved directly in recognition of a glycoprotein (ie. a hormone), or a surface glycoprotein of another cell (ie. microbe, white blood cell). It is possible that the recognition events merely 'corrals' the protein or cell, with subsequent receptor activation occurring via protein interactions. Alternatively, it is possible that the glycan directly activates the receptor itself. In this case, the oligosaccharide is directly involved in recognition. Carbohydrate side chains may also be indirectly involved in receptor function, by modulating the targeting or activity of the receptor. In this case, the glycan does not participate directly in recognition, but may alter

the glycoprotein in such a way as to influence recognition events. Carbohydrates may also serve to couple receptor systems to effector systems, as in the case of the adenylate cyclase system (Lis and Sharon, 1993).

The diversity of N-glycan effects has important implications for the recombinant production of proteins (Archer and Peberdy, 1997; Cumming, 1991; Hintz *et al.*, 1995; Jenkins *et al.*, 1996; Liu, 1992; Parekh and Patel, 1992). Erythropoietin is a commercially produced recombinant protein which has tremendous pharmaceutical use, however, improper glycosylation of this protein can affect the efficacy of this drug (Cumming, 1991; Geisow, 1992; Lis and Sharon, 1993; Takeuchi *et al.*, 1990). As such, it is important to understand the glycosylation pattern of this protein and choose an expression system accordingly. For instance, if this protein were to be heterologously produced in yeast, the high-mannose type N-glycans may cause improper folding, affect secretion, alter the binding specificity, cause degradation, or cause rapid clearance from the blood. This has necessitated the use of mammalian cell lines for the production of the recombinant protein, a much more expensive and labour intensive process. With a full understanding of the consequences of the glycosylation pattern, one may be able to specifically glycosylate the protein of choice in a much more efficient system. This could require host engineering to create 'tailored' N-glycans for specific applications.

Protein N-glycosylation is a multi-step pathway involving several gene families

Protein glycosylation is a post-translational modification of proteins which involves the attachment of sugar residues to newly synthesized polypeptides. Protein N-glycosylation occurs primarily in the ER and Golgi apparatus, and involves a series of

discrete catalytic steps. A diverse series of enzymes have evolved to carry out the complex steps of this pathway. It is becoming clear that for many of these catalytic steps, gene families have evolved to generate a number of similar genes to perform a diversity of specialized yet related functions, in effect fine-tuning the expressed products. Indeed, virtually every enzyme involved in the protein glycosylation pathway is a member of a multigene family. An extensive classification system has been developed to catalogue the related glycosidases and glycosyltransferases involved in carbohydrate processing (Henrissat, 1991; Henrissat and Bairoch, 1993, 1996; Henrissat and Romeu, 1995). It is necessary to understand the nature of the gene families that comprise the N-glycosylation pathway in order to begin to understand how to modify the pathway in such a way as to produce complex N-glycans. This requires the identification of the various enzymes involved in N-glycosylation, whether these enzymes are members of multigene families, and how the different members of the family may contribute to N-glycan processing. As a corollary this research can lead to an understanding of the evolutionary forces which have led to the formation of such families.

The N-glycosylation synthesis pathways have been fairly well characterized in mammalian systems (reviewed in Kornfeld and Kornfeld, 1985) and yeast expression systems (reviewed in Dean, 1999; Herscovics, 1999), but are not as well characterized in filamentous fungi. Protein N-glycosylation occurs when an oligosaccharide precursor ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) is transferred to newly synthesized proteins in the endoplasmic reticulum (ER) and Golgi apparatus (Kornfeld and Kornfeld, 1985). As the glycoprotein moves through the ER and Golgi to its final destination within or outside the cell, the N-glycan is modified by a series of glycosidases and glycosyltransferases to produce a large

diversity of structures. The types of modifications and final glycan structures formed are quite different between more complex higher eukaryotes and simpler lower eukaryotes.

In all eukaryotes the initial stage of N-glycan processing is the removal of three glucose molecules from the $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ by α -glucosidase I (Figure 2). In higher eukaryotes, up to four mannose residues are then removed by Class I α -1,2-mannosidases to produce $\text{Man}_5\text{GlcNAc}_2$, which is the precursor for complex N-glycan formation.

Complex N-glycans are formed in the Golgi by the addition of GlcNAc by N-acetylglucosaminyl transferase I (Gnt I), removal of two mannose residues by α -mannosidase II, followed by the addition of various sugars, such as galactose and sialic acid. In the lower eukaryote *Saccharomyces cerevisiae*, the removal of glucose is followed by the removal of one mannose residue to produce $\text{Man}_8\text{GlcNAc}_2$. Glycans with fewer than 8 mannose residues are not found in yeast. Mannosyltransferases then sequentially add a number of mannose residues to produce oligomannosidic N-glycans which typically contain core structures of up to 13 mannose but can also lead to the production of very large mannans containing up to 200 mannose residues (Dean, 1999; Herscovics, 1999; Herscovics and Orlean, 1993). This latter process is known as hyperglycosidic mannosylation.

The potential diversity of N-glycans and modification steps in filamentous fungi are less clear than those for yeast or mammalian systems. Filamentous fungi produce N-glycan structures which share properties of both yeast and mammalian systems. N-glycans containing as few as 5 mannose units ($\text{Man}_5\text{GlcNAc}_2$) have been found in filamentous fungi (Chiba *et al.*, 1993; Maras *et al.*, 1997b), suggesting processing of the $\text{Man}_9\text{GlcNAc}_2$ in a manner similar to higher eukaryotes. Some filamentous fungi, such as

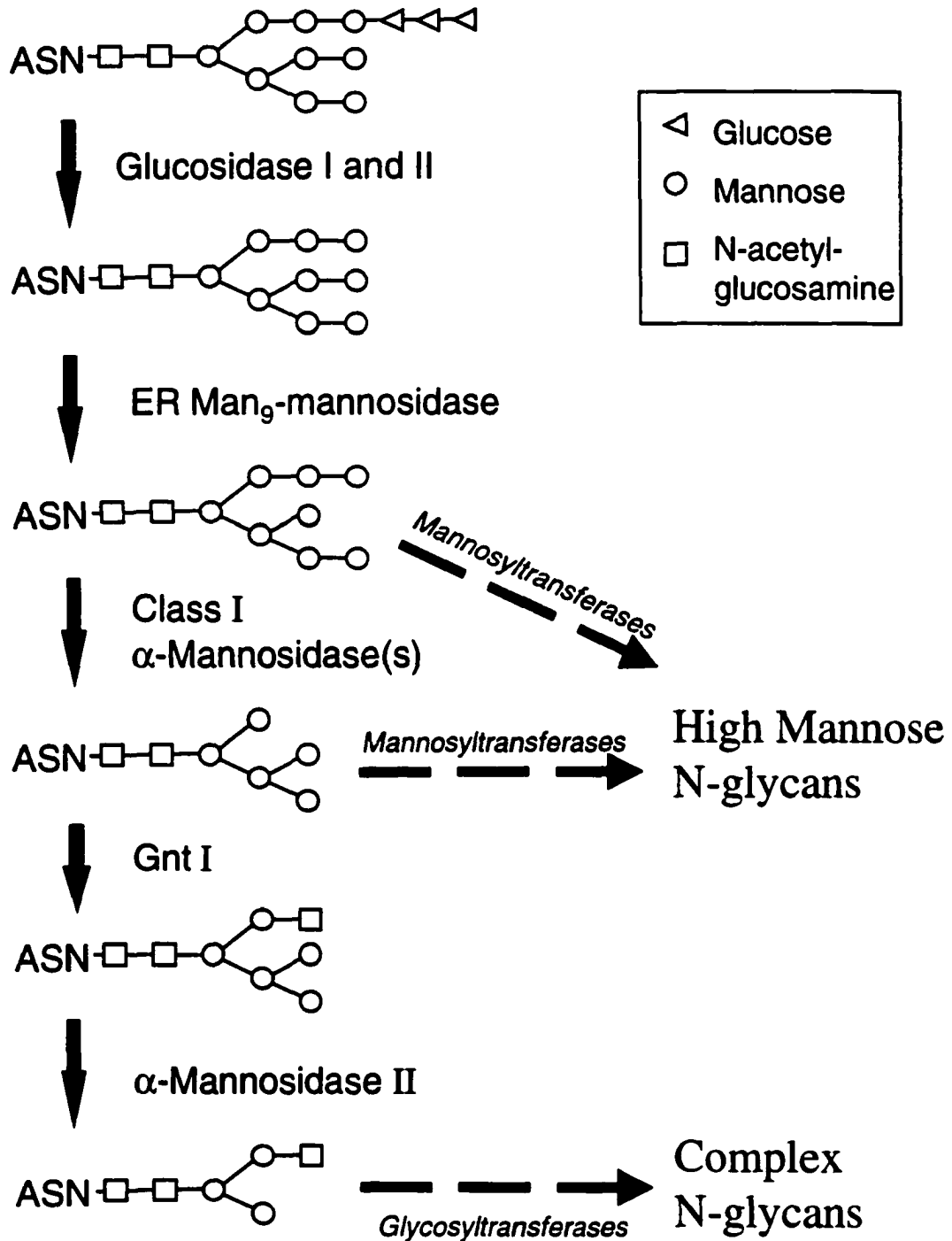


FIGURE 2: N-glycan biosynthetic pathway in the endoplasmic reticulum (ER) and Golgi apparatus. N-glycan biosynthesis proceeds through an initial stage of sugar removal by glucosidase I and II, and by various α -mannosidases. Subsequent addition of terminal branches is achieved by organism specific mannosyl- and glycosyltransferases.

certain species within the genus *Aspergillus*, also produce hyperglycosidic oligomannans, similar to those found in yeast (Maras *et al.*, 1999), however, this is not a common feature of filamentous fungi. The filamentous fungi thus appear to have an N-glycosylation system which is intermediate to yeast and mammalian systems. Like yeast systems, filamentous fungi only produce oligomannosidic N-glycans. Like mammalian systems, however, filamentous fungi produce N-glycans which can be fully trimmed by α -1,2-mannosidases to $\text{Man}_5\text{GlcNAc}_2$. This has important consequences for the remodelling of N-glycans in these expression systems, as the $\text{Man}_5\text{GlcNAc}_2$ serves as the substrate for all subsequent modifications leading to complex N-glycans in mammalian systems. An investigation of the α -mannosidase gene family is the first step in fully differentiating this pathway in filamentous fungi.

Evolution of gene families

The principles of gene evolution can be invoked to explain how the various gene families involved in protein glycosylation arose. Duplication of genes and even whole genomes can be responsible for the generation of multiple copies of genes having the same function and for the creation of novel proteins with diverse functions (Doolittle, 1995; Li, 1997; Ohno, 1970). Duplication events can involve anywhere from a few nucleotides to entire genomes. Equally important to the original duplication event is the maintenance and divergence of the duplicated DNA in the genome (Clark, 1994; Hughes, 1994; Ohta, 1989, 1990, 1994; Walsh, 1995).

Two genes that have descended from a common ancestral gene are known as 'orthologous' genes. Generally, orthologous genes follow from speciation events. Gene

duplication, however, will lead to multiple copies of the gene in a single species. Upon divergence, these duplicated genes are then called 'paralogous' genes (Li, 1997). These duplications have historically led to the evolution of large multigene families which are characteristic of eukaryotic genomes. The size and abundance of multigene families (and superfamilies) increase with the evolutionary complexity of organisms, thus their formation is an extremely important evolutionary process (Huynen and van Nimwegen, 1998).

Assuming that DNA duplications were selectively neutral, their invasion into a population would be influenced mainly by random genetic drift. If the duplication itself conferred some selective advantage to an organism (Clark, 1994), either through a direct advantage conferred by the duplicated gene product, or by the relaxation of selective constraints on the duplicated DNA, then duplications would have a greater probability of fixation. It has been hypothesized that duplications may confer a selective advantage since the duplicated locus would be more tolerant to deleterious mutations due the functional redundancy created by the extra gene copy (Clark, 1994; Hughes, 1994).

Large scale duplication events, such as whole or partial genome duplications can be as important for the expansion of gene families as single gene duplications. Most genes, especially in eukaryotic genomes, are involved in complex gene networks, such as regulatory networks and physiological pathways. Duplication events have the least disruptive effect on gene networks if they involve either a small number of genes in the network or all of the genes in the network (Wagner, 1994). Duplications involving ~40% of the genes in the network would be the most disruptive and likely would not be tolerated. Duplication of entire networks would occur with whole genome duplications,

or with chromosomal duplications involving tightly clustered gene networks.

Chromosomal duplications may arise from nondisjunction during meiosis, in which homologous chromosomes or sister chromatids fail to segregate, leading to aneuploidy and polysomy in the daughter cells. Whole genome duplications can arise by genome hybridization and polyploidy which may be followed by a period of chromosome loss (Li, 1997). Both of these scenarios could be tolerated by filamentous fungi and are likely factors in the evolution of gene families in these organisms.

The generation of duplicated gene networks may result in novel physiological or regulatory pathways, which may confer a selective advantage. It is hypothesized that *S. cerevisiae* may have utilized this process in the acquisition of anaerobic growth (Wolfe and Shields, 1997). Several large duplicated blocks with very similar gene arrangements have been found throughout the *S. cerevisiae* genome. It is thought that genome duplication occurred after the yeasts *Saccharomyces* and *Kluyveromyces* diverged and was followed by a period of degeneration, in which approximately 85% of the duplicated genes were lost and the remaining blocks were shuffled somewhat by reciprocal translocation events. Several pairs of duplicated genes are differentially regulated during aerobic and anaerobic growth, suggesting that genome duplication may have allowed *Saccharomyces* to adapt to anaerobic growth, a feature that is lacking in *Kluyveromyces*.

It is expected that the duplication of an entire genome would result in a large amount of redundant genetic material. In the absence of selection, genetic redundancies created by gene duplications are expected to be lost, either through acquisition of deleterious mutations leading to formation of pseudogenes, or through genetic divergence leading to genes with novel function. Such redundancies have been shown, however, to be

able to persist in populations for a long time (Hughes and Hughes, 1993). It is possible that genetic redundancies are maintained to increase the efficiency and fidelity of physiological pathways in the cell and to safeguard against any loss of information in such pathways (Tautz, 1992). This situation appears to have occurred in the N-glycan synthetic pathway. For example, many species carry multiple Class I α -mannosidase with overlapping functions in the N-glycan synthesis pathway (Moremen *et al.*, 1994). Genetically redundant material produced by large scale duplications may thus be maintained in populations, increasing the fidelity and flow of information in the cell, and providing genetic material for the evolution of proteins with novel functions. The diversification of physiological and developmental pathways will be more likely to occur if all genes involved in the pathway are duplicated simultaneously (Nadeau and Sankoff, 1997). Localized tandem duplications, however, occur more frequently and are readily tolerated. In fungi, functional gene clustering is quite common, which illustrates the potential role of localized tandem duplications in gene family evolution (Keller and Hohn, 1997). Both mechanisms, whole genome and tandem duplications, appear to have had a role in the diversification of gene families of fungi.

The α -mannosidase gene family

Research on the gene families of the N-glycan processing pathway illustrates the similarities and differences of the glycosylation pathways in various lineages. The α -mannosidase gene family is a diverse groups of genes conserved throughout eukaryotic evolutionary history. Members of this family have been found in mammals, insects, filamentous fungi and yeasts. Duplication events which led to the diversification of this

family from an ancestral gene appear to have occurred quite early in the eukaryotic evolutionary history. The α -mannosidases have a variety of cellular functions and localizations, and are involved in N-glycan processing in the endoplasmic reticulum and Golgi apparatus, as well as N-glycan degradation in the lysosome, vacuole, and cytoplasm (Daniel *et al.*, 1994; Moremen *et al.*, 1994). The α -mannosidases can be classified into two independently derived lineages, termed Class I and Class II (originally Class 1 and Class 2), based on protein sequence alignments as well as biochemical and physiological roles of the various gene products.

The Class I α -mannosidases are involved in the early stages of N-glycan processing in the ER and Golgi by catalyzing the removal of terminal α -1,2-linked mannose residues from N-glycan chains. This group of enzymes includes the ER Man₉-mannosidase, endomannosidase and Golgi mannosidase IA/IB. Several genetic pathways seem to exist for the removal of the α -1,2-linked mannose residues from N-glycans and there is significant genetic redundancy in this gene family (Daniel *et al.*, 1994). In most cases, there appear to be multiple paralogous genes which have evolved somewhat specialized, yet overlapping functions in each species. The second group of α -mannosidases is more heterogeneous and contains the lysosomal mannosidases, the Golgi mannosidase II and a distantly related group of enzymes, including the rat ER/cytosolic mannosidase (Bischoff *et al.*, 1990), yeast vacuolar mannosidase (Yoshihisa and Anraku, 1989).

Glycosylation in recombinant protein expression systems

Filamentous fungi, particularly those of the genera *Aspergillus* and *Trichoderma*, are widely used for the heterologous expression of proteins as they are capable of producing up to 20-30 grams of protein per litre of culture (Archer and Peberdy, 1997; Hintz *et al.*, 1995; Maras *et al.*, 1999; Punt *et al.*, 1994). Filamentous fungi produce abundant amounts of useful extracellular enzymes, such as (gluco)amylases, cellulases, pectinases, catalases, proteases, lipases, phosphatases and glucose oxidase, and the production of these enzymes accounts for a significant portion of the multibillion dollar annual market for industrial enzymes (Archer and Peberdy, 1997). Although the homologous production of these enzymes is very high, such yields are rarely reached with the production of heterologous proteins (Archer *et al.*, 1994; Archer and Peberdy, 1997). Several non-fungal proteins have been produced in filamentous fungi (Hemming, 1995), such as calf chymosin (Calmels *et al.*, 1991), human interferon and bacterial endoglucanase (Gwynne *et al.*, 1987). In addition, numerous fungal proteins have been heterologously expressed in other fungi. Generally, heterologous expression of fungal proteins is about 10-20% as efficient as homologous protein production, while heterologous expression of non-fungal proteins is much lower, often as low as 1% as efficient as homologous protein expression.

Several strategies have been employed to improve the yields of heterologous proteins, such as the use of strong promoter systems (Hintz *et al.*, 1995; Hintz and Lagosky, 1993), gene fusions to highly secreted proteins (Broekhuijsen *et al.*, 1993; Contreras *et al.*, 1991; Gouka *et al.*, 1997), and reduction of protease activity (Archer and Peberdy, 1997; van den Hombergh *et al.*, 1997 and references therein). For instance, Hintz

and Lagosky (1993) developed an expression system which utilized the inducible *alcA* promoter which normally drives expression of the alcohol dehydrogenase I gene of the ethanol regulon. This promoter is subject to carbon catabolite repression by the CreA repressor, but under glucose depleted conditions, this promoter can be induced to high levels of expression. Pathway specific induction is under the control of the transcriptional activator AlcR. By integrating multiple copies of the *alcR* gene into the *Aspergillus nidulans* genome, this host strain was capable of higher transcription of *alcA*-driven reporter genes. By providing limited amounts of glucose in the expression media, this promoter could be utilized for phased growth of the fungus - during the initial phase, the *alcA* promoter is carbon catabolite repressed and in the later phase, after glucose depletion, the promoter is induced to high levels. This approach allows the initial accumulation of biomass prior to expression of the desired protein.

One drawback to this approach is that glucose depletion can also induce several scavenger pathways, such as the production of extracellular proteases (Dunn-Coleman *et al.*, 1988; Hintz *et al.*, 1995; Hintz and Lagosky, 1993). To avoid this problem, Hintz and Lagosky (1993) developed a glucose derepressed promoter by mutation of the CreA binding sites in the *alcA* promoter. This modified promoter was used to express the human interleukin-6 (IL-6) protein which had been fused to the glucoamylase gene from *A. nidulans*, separated by a KEX2-like cleavage site. The KEX2-like cleavage allowed endoproteolytic cleavage of the glucoamylase protein, causing release of intact IL-6 into the culture media. It is believed that sequences present in the glucoamylase protein may facilitate efficient passage through the secretory apparatus and may provide protease protection to the secreted product (MacKenzie *et al.*, 1993). Utilizing this system,

expression of the human IL-6 was significantly increased, and the problem of co-expressed proteases was significantly reduced. This provides an excellent example of the types of approaches which can be utilized to increase the levels of heterologous protein production which will make filamentous fungi increasingly attractive as expression hosts.

An important consideration in the production of non-fungal enzymes is the fidelity of post-translational processing events, such as protein glycosylation (Archer and Peberdy, 1997; Hemming, 1995; Hintz *et al.*, 1995). This is especially true for the production of recombinant human therapeutic products such as epidermal growth factor (EGF), interleukin-6 (IL-6) and corticosteroid binding globulin (CBG) (Gwynne and Devchand, 1992). The post-translational addition of aberrant N-glycans to such proteins can result in reduced activity and/or stability, increased serum clearance, and can sometimes result in an adverse immune response (Goochee *et al.*, 1991; Jenkins *et al.*, 1996; Varki, 1993). To utilize filamentous fungi for the production of such specialized glycoproteins, it is preferable to produce glycoproteins which carry carbohydrate structures as similar to the natural product as possible. Understanding this process may allow manipulation of the N-glycosylation pathway to produce glycoproteins with 'correct' N-glycan structures. To fully realize the potential of filamentous fungi as highly flexible expression hosts it is necessary to examine the genetic components of their protein N-glycosylation pathways.

Remodelling the N-glycan pathway

The recombinant expression of glycoproteins requires consideration of the glycosylation machinery of the system to be used (Archer and Peberdy, 1997; Jenkins *et al.*, 1996; Meynial-Salles and Combes, 1996). The most widely used mammalian, insect,

yeast, and filamentous fungal expression systems all have significant drawbacks for the production of glycoproteins. Chinese Hamster Ovary (CHO) cell lines, for instance, produce complex N-glycans which lack terminal sialic acid residues which are necessary to prevent rapid serum clearance of these glycoproteins (Fussenegger *et al.*, 1999). Recombinant expression of functional α -2,6-sialyltransferase in this cell line causes production of N-glycans containing the requisite terminal sialic acid residues (Grabenhorst *et al.*, 1995). Despite this success in N-glycan engineering, the protein yields in CHO cell lines are significantly lower than other expression systems, hence it would be desirable to utilize an alternative expression system with high secretion capabilities. Insect cell lines, such as Sf9 cells are capable of much higher secretion levels, but N-glycan structures in these cells are generally not larger than $\text{Man}_3\text{GlcNAc}_2$ or $\text{GlcNAcMan}_3\text{GlcNAc}_2$, the precursors for complex N-glycans. Jarvis and Finn (1996) have attempted to engineer the glycosylation pathways of these cell lines to produce complex N-glycans by expression of subsequent glycosyltransferases in insect cells. Expression of the bovine β -1,4-galactosyltransferase into Sf9 cells resulted in the production of N-glycans with terminal galactose. These results demonstrate the feasibility of modifying the N-glycosylation pathways of particular expression systems to produce complex N-glycans which would allow the production of glycoproteins with specific N-glycan structures. Since yeasts and filamentous fungi are capable of very high secretion levels, it would be desirable to manipulate the pathways of these organisms for such a purpose.

The remodelling of the glycosylation pathway in yeasts is complicated by the fact that these organisms produce highly branched hypermannosylated structures that can contain up to 200 mannose residues (Herscovics, 1999; Herscovics and Orlean, 1993).

Yeasts also produce 'core' oligomannose N-glycans containing up to 13 mannose residues. In order to manipulate the pathway to produce complex N-glycans, it is necessary to remove the steps in the pathway which lead to hypermannosylation. Chiba *et al.* (1998) utilized mutant strains which were deficient in several of the glycosyltransferases necessary for the production of hypermannosylated N-glycans to bypass this problem. Processing reactions in the ER of yeasts, however, only reduce the $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ precursor to $\text{Man}_8\text{GlcNAc}_2$. To produce complex N-glycans, it would be necessary to further trim this product to $\text{Man}_5\text{GlcNAc}_2$. Overexpression of the *Aspergillus satoii* α -1,2-mannosidase which contained an ER 'HDEL' retention signal was successful in producing $\text{Man}_5\text{GlcNAc}_2$ in these mutant yeast strains. In order to further process these N-glycans, the addition of several additional steps will be necessary (Roy *et al.*, 2000).

Remodelling of the glycosylation pathway of filamentous fungi is being attempted in two fungal expression systems, *A. nidulans* and *Trichoderma reesei*. The production of complex type N-glycans relies upon the addition of GlcNAc to $\text{Man}_5\text{GlcNAc}_2$ by the enzyme Gnt I (Figure 2), the first committed step in the production of mammalian-type N-glycans. This enzyme is not found in filamentous fungi, hence Kalsner *et al.* (1995) inserted the mammalian Gnt I gene into the genome of *A. nidulans*. Expression of the Gnt I alone did not result in the production of N-glycans containing an additional GlcNAc ($\text{GlcNAcMan}_5\text{GlcNAc}_2$). This is likely because the substrate $\text{Man}_5\text{GlcNAc}_2$ was in limiting amounts. There may thus be a 'bottleneck' preventing the production of significant amounts of $\text{GlcNAcMan}_5\text{GlcNAc}_2$. Efficient removal of mannose in the ER and Golgi could provide the necessary precursors for the production of complex N-glycans. It is

expected that controlled overexpression of specific Class I α -mannosidases may clear the 'bottleneck' and permit production of complex N-glycans in *A. nidulans*.

Manipulation of the N-glycan processing pathway has progressed a little further in *T. reesei*. Maras *et al.* (1997b) were able to convert oligomannose glycans from cellobiohydrolase I (CBHI) to GlcNAcMan₅GlcNAc₂ by *in vitro* treatment with Gnt I. Only a small proportion of the N-glycans were converted, however, due to the fact that only a small fraction of the N-glycans provided a suitable substrate (Man₅GlcNAc₂) for Gnt I. Pre-treatment of the purified CBHI with α -1,2-mannosidase significantly increased the yield of complex N-glycans, illustrating the need for efficient production of suitable substrate. Maras *et al.* (1999) also reported the *in vivo* conversion of oligomannose N-glycans to complex N-glycans by heterologously expressed Gnt I, although the efficiency of conversion was low. Again, the conversion process may have been blocked by a 'bottleneck' preventing production of significant amounts of substrate for Gnt I. These strains may need to be further manipulated to clear this bottleneck, either by overexpression of α -1,2-mannosidases to produce Man₅GlcNAc₂, or by elimination of mannosyltransferase activity which may be converting Man₅GlcNAc₂ (or intermediates) into glycans which are unsuitable for conversion to complex N-glycans. Characterization of N-glycan processing gene families has other potential biotechnological uses. The substrate specificity of the various glycosidases purified from filamentous fungi are currently being used to sequence glycans and to aid in monosaccharide composition determination. The identification of novel glycosidases with differing substrate specificities will increase the number of tools available for this type of research.

Glycosylation and fungal pathogenicity

Protein glycosylation is known to be important in many aspects of fungal growth and/or pathogenicity. Many phytopathogenic fungi, for example, secrete cell wall degrading enzymes such as endopolygalacturonase and pectate lyase to aid in invasion of host plant tissues. Inhibition of N-glycosylation has been shown to reduce the secretion level and activity of these enzymes in certain fungal species (Dean and Anderson, 1991; Di Pietro and Roncero, 1996). Protein glycans are also directly involved in such pathogenic processes as host-tissue adhesion (Gow *et al.*, 1999; Hollenstein *et al.*, 1995; Masuoka and Hazen, 1997), formation of infection structures (Bircher and Hohl, 1997), and elicitation of host defenses (Basse and Boller, 1992; Basse *et al.*, 1993; West, 1981). In addition, changes in the cell wall carbohydrate composition of certain pathogenic fungi, such as *Sporothrix schenckii* altered the virulence of these organisms, pointing to a direct role for these molecules in pathogenicity (Fernandes *et al.*, 1999).

Recent studies of host-pathogen interactions have focused on the identification of individual pathogenicity determinants or specific elicitors of host defense responses. A criticism of these approaches is that they do not account for the complex nature of pathogenicity and often provide inconclusive results regarding the specific role of these factors. Protein glycosylation can simultaneously affect a large number of pathogenicity determinants and the general ability of a pathogen to invade host tissues. It is expected that sufficient differences exist in the glycosylation machinery of fungi, plants, and animals to provide ideal targets for the selective interruption of glycosylation in fungal pathogens. It is anticipated that the regulation of glycosylation will lower the ability of the fungus to invade the host and will provide a useful point of control. A significant portion of the

fungus cell wall consists of mannoproteins which are critical for the biology and pathogenicity of these organisms. Targeting the genes which are necessary for the production of these cell wall components may be an excellent approach for the development of antifungal drugs (Gow *et al.*, 1999; Tanner *et al.*, 1995). Such approaches will require characterization of the gene families which are involved in the glycosylation pathways of filamentous fungi.

Research objectives

The overall objective of this research project was to identify and characterize specific genetic components of the glycosylation machinery of filamentous fungi, with a focus on the α -mannosidase gene family. The cloning and characterization of this family was expected to provide insight into the evolution of the α -mannosidase gene family, especially in filamentous fungi, and to decipher the N-glycosylation synthesis pathway in filamentous fungi. Such information will enable the development of strategies for modifying the N-glycosylation machinery to produce specific N-glycan structures on recombinant proteins expressed in filamentous fungi. Characterization of the glycosylation machinery in filamentous fungi may also allow us to develop control strategies for fungal pathogens. The glycosylation pathways may provide an excellent global target for reducing the pathogenic fitness of these organisms.

CHAPTER 2 - Identification and analysis of a Class 2 α -mannosidase from

Aspergillus nidulans (Eades, C.J. and W.E. Hintz. 1998. *Glycobiology* 8: 17-33)

2.1 Abstract

A Class 2 α -mannosidase gene was cloned and sequenced from the filamentous fungus *Aspergillus nidulans*. A portion of the gene was amplified using degenerate oligonucleotide primers which were designed based on similarity between the *Saccharomyces cerevisiae* vacuolar and rat ER/cytosolic Class 2 protein sequences. The PCR amplification product was used to isolate the full length gene, and DNA sequencing revealed a 3383 bp coding region containing three introns. The predicted 1049 amino acid reading frame contained six potential N-glycosylation sites and encoded a protein of 118 kD. The protein sequence did not appear to encode a typical fungal signal sequence or membrane spanning domain. Although the cellular location of the *A. nidulans* mannosidase was not determined, experimental evidence suggested that it was located within a subcellular organelle. The Matchbox sequence similarity matrix indicated that the *A. nidulans* protein sequence was more highly similar to the rat ER/cytosolic ($R_{ij}=0.33$) and *S. cerevisiae* vacuolar α -mannosidases ($R_{ij}=0.43$) than the rat and yeast sequences were to each other ($R_{ij}=0.29$). These three enzymes were found to be distantly related to other Class 2 sequences, and compose a third subgroup of Class 2 α -mannosidases, as shown by ClustalW sequence alignment.

2.2 Introduction

Glycosylation, the process by which oligosaccharides are covalently linked to specific amino acids of newly synthesized proteins, can have major effects on protein structure and function. These include effects on the stability, antigenicity, and biological activity of glycoproteins (Goochee *et al.*, 1991; Opdenakker *et al.*, 1993) thus protein glycosylation can be a very important factor in choosing an expression system for the production of recombinant proteins. Prokaryotic expression systems, such as *Escherichia coli*, produce high levels of recombinant proteins but entirely lack glycosylation and other post-translational machinery (Kalsner *et al.*, 1995). Eukaryotic systems are preferred for recombinant protein production because they are capable of protein glycosylation and other post-translational modifications. Filamentous fungi of the genus *Aspergillus* are widely used for the expression of recombinant proteins and can produce as much as 20 grams recombinant protein per litre of culture (Hintz *et al.*, 1995). To take full advantage of the *Aspergillus* expression system, it would be desirable to produce glycoproteins which contain carbohydrate structures as similar as possible to the natural product. This is especially important for the production of recombinant human products of therapeutic interest such as epidermal growth factor (EGF), interleukin-6 (IL-6) and corticosteroid binding globulin (CBG) (Gwynne and Devchand, 1992). To work towards this goal, it is necessary to characterize the glycosylation pathway of secreted proteins and to understand the regulation of this process in *Aspergillus*.

Asparagine-linked (N-linked) protein glycosylation in higher eukaryotes is an ordered process which occurs in several stages (see reviews in Elbein, 1988; Herscovics and Orlean, 1993; Kornfeld and Kornfeld, 1985; Moremen *et al.*, 1994). Initially, an

oligosaccharide precursor consisting of three glucose, nine mannose and two N-acetylglucosamine molecules ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) is co-translationally transferred to the newly synthesized polypeptide in the endoplasmic reticulum (ER). This precursor is then sequentially processed as the protein progresses through the ER and the Golgi apparatus. In the ER, α -glucosidase I and II first remove the three glucose molecules. An ER-specific $\text{Man}_9\text{-}\alpha\text{-1,2-mannosidase}$ then removes a single mannose residue, producing $\text{Man}_8\text{GlcNAc}_2$. In the ER and Golgi, $\alpha\text{-1,2-mannosidases}$ remove a total of four mannose residues, yielding $\text{Man}_5\text{GlcNAc}_2$ which is the precursor for complex, hybrid, and high-mannose N-glycans. Following the addition of a single GlcNAc to $\text{Man}_5\text{GlcNAc}_2$ by GlcNAc transferase I (GnT I), mannosidase II removes two additional mannose groups, producing $\text{GlcNAcMan}_3\text{GlcNAc}_2$. Various transferases, such as GnT II, fucosyl transferase, galactosyl transferase, and sialyl transferase assemble the oligosaccharide into its final structure. In higher eukaryotes a variety of different carbohydrate units can thus be attached to a common precursor to form an array of distinct N-glycans.

It is generally accepted that the glycosylation machinery of lower eukaryotes is somewhat simpler than higher eukaryotes. Similar to higher eukaryotes, the initial precursor is processed to the $\text{Man}_8\text{GlcNAc}_2$ stage, however, in lower eukaryotes the $\text{Man}_8\text{GlcNAc}_2$ can be further mannosylated to yield N-glycans containing many mannose residues. For certain secreted and cell wall proteins, up to 200 mannose units may be added post-translationally (Herscovics and Orlean, 1993). The precise role of the $\alpha\text{-1,2-mannosidases}$ in this process remains unclear. The gene for the ER-specific $\alpha\text{-1,2-mannosidase}$ which trims the $\text{Man}_9\text{GlcNAc}_2$ molecule to the $\text{Man}_8\text{-oligosaccharide}$ has been characterized for the yeast *Saccharomyces cerevisiae* (Camirand *et al.*, 1991). This

enzyme does not, however, appear absolutely necessary for further oligosaccharide mannosylation, as the Man₉-oligosaccharide and the Man₈-oligosaccharide are equally suitable substrates for the α -1,6-mannosyltransferase which initiates outer chain formation (Romero and Herscovics, 1989). The α -1,2-mannosidase genes have also been cloned from *Aspergillus satoii* and from *Penicillium citrinum* (two closely related fungal species) (Inoue *et al.*, 1995; Yoshida and Ichishima, 1995). Mannosidases which further process the Man₈GlcNAc₂ to smaller forms, such as Man₅GlcNAc₂, may also exist although none have yet been found. A second α -mannosidase gene, located to the vacuole, has been characterized in *S. cerevisiae* (Yoshihisa and Anraku, 1989, 1990). We describe the cloning and sequence characterization of a similar α -mannosidase from the filamentous fungus *Aspergillus nidulans*.

2.3 Results

2.3.1 Isolation and sequence analysis of *Aspergillus* α -mannosidase

The *S. cerevisiae* vacuolar α -mannosidase sequence (Yoshihisa and Anraku, 1989) aligned well with the rat ER/cytosolic α -mannosidase sequence (Bischoff *et al.*, 1990) but not with other α -mannosidase sequences. Regions of shared homology were identified from this alignment of protein sequences. Conserved regions of the amino acid sequences having both a low degree of codon redundancy and a significant codon usage preference for *A. nidulans* were used to design degenerate oligonucleotide primers for the amplification of an α -mannosidase sequence from *A. nidulans* strain SM222 (Table I). A codon usage table was compiled from an analysis of twenty highly expressed *Aspergillus* genes and the best candidate for each amino acid was selected for primer design.

Primer	Sequence (5' => 3')
AspmanF	ggggtttaagcttactgtcacatcgatacagcctggctctggcccttcnnggac
AspmanR	ggggtttgaattcgggaggagtagccaaaggtgtcggggagccagaa
I1F	atgggtggagacgttctccgacagact
I1R	atcccagtggaattcaagccgttctct
I2F	ccctgggtggtgactccattcagcc
I2R	gaacaatcaacctgtctccgagg
I3F	gagctcgaagcgaagtgaatgg
I3R	gcgccctgggagctcggaggagg
KOconf-F	gacgatcccgaagatatgcgcg
KOconf-R2	cgcgccatcaggatccggtcacc
alcAprom-F	gaatagccaaaagtatccggaagccag
manR-conf	aacactccaatcctatcacctcgc
M3SL-F	<u>cgtaaggtgaccggatcc/gatggcgcgccg/gat/gaggcgacggtggtggtggcgggc</u>
M3SL-R	aggtgtcactggatccgtctcctgccattcgacttcgctttcg

TABLE I: Oligonucleotide primer sequences

Amplification of *A. nidulans* total genomic DNA with these primers yielded a 365 bp fragment. Sequencing of this fragment revealed a single open reading frame sharing significant homology with both the rat ER/cytosolic and *S. cerevisiae* vacuolar protein sequences, confirming that the PCR product represented a portion of the authentic *A. nidulans* α -mannosidase gene. The 365 bp PCR product was then used as a probe to screen an EMBL-3 library prepared with SM222 genomic DNA. A 9685 bp *BamHI* fragment was recovered containing the entire coding region of the *A. nidulans* α -mannosidase gene as well as several thousand base pairs of flanking DNA.

The DNA sequence of the *BamHI* fragment was determined and a coding region 3383 bp in length was tentatively identified (Figure 3). The proposed ATG start site was embedded in the sequence GCAACCATGG, which was similar to the optimal eukaryotic initiation codon (GCC{A/G}CCAUGG), and was similar to sequences found in other filamentous fungi (ANNATGC) (Gurr *et al.*, 1987). The DNA sequence was also determined for the 2892 bp region upstream of the putative ATG start site (Figure 3), and several motifs typical of fungal promoters were identified. The three translational reading frames of the *BamHI* fragment were compared to the rat and *S. cerevisiae* protein sequences and introns were initially identified by searching the primary sequence with 5' and 3' fungal intron consensus motifs (Gurr *et al.*, 1987). The intron search sequence included a 5' consensus donor site (GTNNR) and a lariat site (CTNA) positioned at least 8 bp upstream of the putative acceptor site (YAG). Three different introns were identified by the occurrence of these consensus motifs at sites where there was both a shift in the reading frame and extra residues in the expected coding region based upon the consensus sequence of the rat and yeast peptides. The positions of the introns were verified by

CAGCTTCTCGGCAACAATCTGGATAATGACAAGCACGTTGTCATGATCTTC

-2840 ACTCTCCTGTCCGATCACTATTTACGGCTCACCGAATCGAGGCACCAGCACCTTATCACCTATCTTGGTPTT
-2769 GCGAAGGACCAAAATCTTGGATAACGATATCTAAACTGGTTTCTATTGGTTTCCAAGCTGGATTGGCATCT
-2698 TCATGATGCAATCGCACCGTGATTTCCGGTGTTTATCTACTGGCAAACAGTCGACATGGTCCCTGCCAAATG
-2626 TACATTTCCGCACCCATCCGCACCACATCGGTGCCTTGCTTGCCTGTGCTTTGTCCCGCCAGGCTGCAG
-2556 TCGGCTGGTCAGTCTCGGGTGCAGTATGTACATAGAGGCGCGCTAGCACTGGCACTTGCAGGCAGAG
-2485 ATTGCTATTCCAATTGAACTTTTCTATTGCAATGCTTATCATGGTAGTCAGCAAACACTGAAGAAGCAAT
-2414 TATGCTCAACTTCAGCTCTGTCTCTGTACGCTCAGCCGGCTCTGTCTCATAGGCTCAGCCAGGCCGGCT
-2343 CAGGCTGGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCATCTTGTATC
-2272 TTGGCCCGGGTTGGACCCCGTTCTCGTCTGCTCAATCAGGAGTCAGGCTCAGAGGATTGATAAGCTGG
-2201 ACTTGAAGGAACAGCTAGTGTCTCAGGCCAAATCGGCAACTGGGCTCTGCACATGTGGACGCAGCTAAGC
-2130 TGGGGGATTTTCGCTTCAAGTTCGGGTGAGATACCGTGCATGCGTGTAGTGGATACGATGTACAGCTCTGAA
-2059 GCTATGGTGATATTGGCTTTTCTCCTGTTAGACGATCGTGTCTCATTTAAAGGGGGAGGGGAAAGGAGTCA
-1988 TCGGCAGAGAAAATGCAGTTACAGAAGCACGTTAAATGAGGCAACCGTTGATCTTTACGGGTATGTTTCGAC
-1917 GTTTCGAGCGGTGAGTCTGATCGTGTGCAACAATAAAGAAAGGAGGTGATTGGCATCAAGGTTCCAAGCC
-1846 GAAGCAGGCTGGAGGAGGTGCCAGGGCCAGTTCAAGCGAGGCAGGCTGGAGGTGGTGCAGTGCAGGCAAG
-1775 ATCTCCAATTCTACTTAATGCAATGCAACCCAGCAATGATCAAACCCCTCAGAGATCCAGGTTACCGCTC
-1704 GCTTAAAGGAGGCAACACGATTTCTGCTACTGGTCCATCCCACTGATGGCAAAATGATTCGATTTGAGAT
-1633 GGGCTAGTACAGTGGCGACAGAGGCTACTGCACGGCATACTTGGCCAGTGCCTGTTCTCCCAAACGAA
-1562 CGACCGCTATGCCTACAACATCATACGTCACGCTAATCGATTCCGCATCAATTAAGCGAGATCTCTCATTT
-1491 TACGTCGTCATCTCAGGCTCGCATGAGTGTGAGGTCATTACTACTTGTCTTCATCGATCTCAATGTCGA
-1420 TCAGATATGGATCTCTAATAACAACAGTCTTCCAATCCACGCTAAGCTCGTAGGCGGTTCTTGTGTCTTT
-1349 GGGGGCAGTGACCAGTCTGTCGAGCAGTACTATTGGCCAGAACTCATATGTAGTCTTCTAGGCTCGAG
-1278 AGAAGAGAGAGCGTCCGAGTACTGTTATCAGTTGGCATACCTTCCACTGCAACGAAAGCGTTCTGAGACAT
-1207 CACATCTCAAAGACGAATAACACAGAGGCAAGGTAGCTCGGTGCTCGAAACCAATGCCAAACACTTCAG
-1136 CTGACGGTTGGCTTCCGGATTCTTAAACAGCGGACCCCGCTCGATGGAATCCACGAAATCCATACAAGTT
-1065 TCTGACTTGTTCGAGCTTCCAGCCCGTTTCAAACCTGGTCTAGCTTCTCTGGCCCTCCATCGATGATCGC
-994 AGTGGGAAGATTTCACTCAGTAAAGTCGCTGCAATGTAGGTGGCAAGATTTCCCGGCAAAATCCAGACTGGG
-923 GGAACCTCGAGCCGGAGAAATCTGCGGTGTAAGTAGCCGCGTGCCTTGGCTTTGTTGGGCGTCAGGCTCTG
-852 CAGAGTGGAGGAGCATGCTCAAAGAGCTAATCCAAGAGAGACCGCACTAAACTGTTCCAGAGAGGAT
-781 TCGGCCGGAATCGTCCCTTCAAATCTGCTATCTGATATTGATGTCGAAAAACTTTCAAAGCGCCAAAAGTCG
-710 GAATGTTGATTATATCAGCTCGACTGCTCATTGTGACTGGGATGTACCTGTCTTCTTCTGCTTAAAGGAC
-639 TTGCTGAAACTGGGACGGGTTGAGCTGGCTTGGGGAATCGGAAGGCTTAAGGTAAGCACCCGGTATTGTA
-568 CTTGGCAGCAACGCACAGGCCAGGCTCAGAGCTCTCTACCCCTCCGTCATGTCATGTGCTATTTCCGAGAG
-497 ATGACAGGCCCGGCTCTGAGATGAGAGTGAATGCTGTTGATTTAGATCACAATGGGTTTAGGACGGGCC
-426 GTTCTTGGCAGGCCCACTAAGCTTGAATAAGCGAAAATCGTCAATGATGATATATAGAGAGCGCAGAGA
-355 TACATATCGGAAGGACTGTTTCCCAACTAGAGAACTTTCTTGTGAACAGTCGCCGAAAGGAAATCTAT
-284 ACAAAATACCTATCCTGTGCAAGCGTAAATATATCACGGCGTTGATGGGGGCTCAAATCTCGGATATCGTAG
-213 TCACGTTGGCCCCAGCTTCTGCCTTCTCCGATGCACCAACCAACGTCACCTGTTCTCTGAAATCCCTGCTC
-142 CAACCTCTCTGTTCCGCTCTCGAATCGTCTCTCTGTCTCTTTCCAAAACCTCTAGTTTTCGATCATCC
-71 CTTCAATTTCAAAGTATTGAAATTCAGTGTCTGTGCTTCAGCTATAGTTTACCAGTTCGGACCCGCAACC

FIGURE 3: Full length sequence of *Aspergillus nidulans* α -mannosidase. The DNA sequence was determined for the 3383 bp α -mannosidase coding region as well as a 2892 bp region upstream of the putative ATG start site. Within the coding region, three introns were found with characteristic fungal consensus sequences (underlined). Six potential N-glycosylation sites were detected (dotted underline). Within the upstream region, numerous CT-rich regions characteristic of fungal promoters were identified (double underline). There was a putative TATA region (TAATATAT) located 256 bp upstream of the start site and a CAAT site was found 60 bp upstream of this putative TATA region. A typical transcription start site (CAACCAAC) was identified 178 bp upstream of the coding region and a transcription terminator was found 11 bp downstream of the translation stop codon.

+1 ATG GGT GGA GAC GTT CTC CGA CAG ACT CCC AGC TCT CTA CCT CGC GCC CCT GAT
Met Gly Gly Asp Val Leu Arg Gln Thr Pro Ser Ser Leu Pro Arg Ala Pro Asp

55 GCG CCC GTC GGG CAA CGC ATT AGG AGC ATC TAC ACT GAC CGT CTA AAT CAG TTT
Ala Pro Val Gly Gln Arg Ile Arg Ser Ile Tyr Thr Asp Arg Leu Asn Gln Phe

109 ACA TCC AGA GGC CAG TAT CAA AAT CAG AAC TTA ATA GG GTATGTCCTTTAAAGGCTG
Thr Ser Arg Gly Gln Tyr Gln Asn Gln Asn Leu Ile Gly

164 ATCGGTCTTGTAAACTACTTCAAGTCCAAAGCTTCCCCGCATTACCTAAATCGTCTCCCACTCACGCTATTG
intron 1

235 CAAACCTTGACAG T AAA TTA TTC GAG GCC ACA AAT TCT GAC GAA GAT CAC GTC AAG
Lys Leu Phe Glu Ala Thr Asn Ser Asp Glu Asp His Val Lys

291 ATT TCT GTC TAC TCT GTC CCT GAC TTG CAA CGG CCA AGC TTT GAA GAA GCC ACC
Ile Ser Val Tyr Ser Val Pro Asp Leu Gln Arg Pro Ser Phe Glu Glu Ala Thr

345 TCT AAT GAG TTC CAG CCG ACG CAC ATC GGG GCA TCG TTC GGT CCC AGC TGG TCC
Ser Asn Glu Phe Gln Pro Thr His Ile Gly Ala Ser Phe Gly Pro Ser Trp Ser

399 ACG CAT TGG TTT GGG ATT CGC TTG ACG ATC CCC GAA GAT ATG CGC GAA AGA GAA
Thr His Trp Phe Gly Ile Arg Leu Thr Ile Pro Glu Asp Met Arg Glu Arg Glu

453 CGG CTT GAA TTC CAC TGG GAT GCA AAT AAT GAG GGC TTG GTC TGG ACT GAA GAT
Arg Leu Glu Phe His Trp Asp Ala Asn Asn Glu Gly Leu Val Trp Thr Glu Asp

507 GGC CAT CCG TTG CAG GGT CTC ACA GGT GGT GGA GAG CGG ATC GAG TGG ATT ATC
Gly His Pro Leu Gln Gly Leu Thr Gly Gly Gly Glu Arg Ile Glu Trp Ile Ile

561 CCT GAC GCC TGG GCC GAT GGA AAC GAG CAT GTC TTC TAT ATT GAG ATG GCG TGT
Pro Asp Ala Trp Ala Asp Gly Asn Glu His Val Phe Tyr Ile Glu Met Ala Cys

615 AAT GGG ATG TTT GGA AAT GCC CCT GGT GGT GAC TCC ATT CAG CCA CCC GCA CCG
Asn Gly Met Phe Gly Asn Ala Pro Gly Gly Asp Ser Ile Gln Pro Pro Ala Pro

669 GAC AAG TAC TAT ACC CTA ACC ACT GCG CGA ATT ACT GCA GTG AAT TTG GAG GCC
Asp Lys Tyr Tyr Thr Leu Thr Thr Ala Arg Ile Thr Ala Val Asn Leu Glu Ala

723 CGG GCG TTG TTC TAT GAT TTC TGG ATA ATT TCA G GTATGCCAAGTAACTCTATATCG
Arg Ala Leu Phe Tyr Asp Phe Trp Ile Ile Ser Asp
intron 2

780 CATGGGCTCTCATTGACTACTGCCAATAG AC GCC GCT CGA GAG TTT CCT GCT GAG GGC
Ala Ala Arg Glu Phe Pro Ala Glu Gly

841 TGG GAA TCC CAC GAG GCC AAT ATG GTG GCC AAT TCC ATC ATC GAT GCT TTC ATT
Trp Glu Ser His Glu Ala Asn Met Val Ala Asn Ser Ile Ile Asp Ala Phe Ile

895 GCT GGC AAT GGA AGT CTA GAA TCT ATC AAG GAA GGG CGG CGA ATC GCA CAA AAA
Ala Gly Asn Gly Ser Leu Glu Ser Ile Lys Glu Gly Arg Arg Ile Ala Gln Lys

949 TAC CTC GGA GAC AAG GTT GAT TGT TCA GAA GTG TAC GAT AGC GGC ACT CAG CCG
Tyr Leu Gly Asp Lys Val Asp Cys Ser Glu Val Tyr Asp Ser Gly Thr Gln Pro

1003 ATT GTA TAT GGG ATA GGG CAT TGC CAT ATC GAC ACC TGC TGG CTC TGG CCC TGG
Ile Val Tyr Gly Ile Gly His Cys His Ile Asp Thr Cys Trp Leu Trp Pro Trp

1057 GCC GAG ACC AAG CGG AAA GTT GCC AGG TCC TGG TCG AAT CAA TGC GAC CTT ATG
Ala Glu Thr Lys Arg Lys Val Ala Arg Ser Trp Ser Asn Gln Cys Asp Leu Met

1111 GAC CGC TAC CCC GAG CAC CGC TTT ACT TGC TCT CAG GCC CAG CAG TTC AAG TGG
Asp Arg Tyr Pro Glu His Arg Phe Thr Cys Ser Gln Ala Gln Gln Phe Lys Trp

1165 CTA AAG CAG TAC TAC CCG TCC GTA TTC GAC CGC GTC AAG GGC TGG GTG AAG AAG
Leu Lys Gln Tyr Tyr Pro Ser Val Phe Asp Arg Val Lys Gly Trp Val Lys Lys

1219 GGC CAC TTT CAG CCA ATC GGC GGT AGC TGG GTT GAG CAC GAT ACA AAC ATG CCG
Gly His Phe Gln Pro Ile Gly Gly Ser Trp Val Glu His Asp Thr Asn Met Pro

1273 AGC GGG GAG TCA CTG GTG AGG CAA TTT ATC TAC GGA CAA AGG TTG TTT GAG AGC
Ser Gly Glu Ser Leu Val Arg Gln Phe Ile Tyr Gly Gln Arg Leu Phe Glu Ser

1327 AAG TTC GGA GAG CGC TGC ACG ACT TTC TGG CTT CCG GAT ACT TTT GGC TAT TCC
Lys Phe Gly Glu Arg Cys Thr Thr Trp Leu Pro Asp Thr Phe Gly Tyr Ser

1381 ACG CAG ATT CCA CAG CTA TGC CGG CTA GCA GGC ATG AGC CGA TTC TTT ACG CAG
Thr Gln Ile Pro Gln Leu Cys Arg Leu Ala Gly Met Ser Arg Phe Thr Gln

1435 AAG CTT AGC TGG AAC AAC ATC AAC AAC TTT CCC CAT ACC ACG TTT CAA TGG GTT
Lys Leu Ser Trp Asn Asn Ile Asn Asn Phe Pro His Thr Thr Phe Gln Trp Val

FIGURE 3 (cont.): Full length DNA sequence of *Aspergillus nidulans* α -mannosidase.

1489 GCG CTT GAC GGC AGC CAG GTT ATG TGT CAT ATG CCT CCA TCA GAG ACA TAC ACG
Ala Leu Asp Gly Ser Gln Val Met Cys His Met Pro Pro Ser Glu Thr Tyr Thr

1543 GCG GAA GCT CAC TTT GGT GAC CTG AAG CGC AGT ATC ACC CAG CAC AAA TCT TTA
Ala Glu Ala His Phe Gly Asp Leu Lys Arg Ser Ile Thr Gln His Lys Ser Leu

1597 GAC AAG GAT AAC ACT TCG CTT CTT GTT TTT GGA AAA GGC GAC GGT GGT GGC GGG
Asp Lys Asp Asn Thr Ser Leu Leu Val Phe Gly Lys Gly Asp Gly Gly Gly Gly

1651 CCT ACG TTT GAG CAT CTA GAA AAG CTT CGT CGC TGT CGT GGC CTC AGC GAC AAA
Pro Thr Phe Glu His Leu Glu Lys Leu Arg Arg Cys Arg Gly Leu Ser Asp Lys

1705 ACC GGG CTA CTT CCG CGG GCA ACG ATG GGT ACA TCG GTT GAC GAT TTC TTT GCA
Thr Gly Leu Leu Pro Arg Ala Thr Met Gly Thr Ser Val Asp Asp Phe Phe Ala

1759 AAG CTC GAG AAG AAG GCT GCA GAT GGT ACA AAG TTT CCG ACA TGG TAC GGG GAG
Lys Leu Glu Lys Lys Ala Ala Asp Gly Thr Lys Phe Ala Thr Trp Tyr Gly Glu

1813 CTT TAC TTC GAG CTC CAC CGT GGC ACG TAT ACG ACG CAA GCG AAT AAC AAG CGC
Leu Tyr Phe Glu Leu His Arg Gly Thr Tyr Thr Thr Gln Ala Asn Asn Lys Arg

1867 AAC AAC CGC CGA GCT GAA TTT CTG CTC CGC GAA CTT GAG TTC CTA GCT ACG ATT
Asn Asn Arg Arg Ala Phe Leu Leu Arg Glu Leu Glu Phe Leu Ala Thr Ile

1921 GCA TCT ATA ACG CGA TCC GAC TAC GAG TAT CCC AAA AAA GAC ATT GAT GAA ATG
Ala Ser Ile Thr Arg Ser Asp Tyr Glu Tyr Pro Lys Lys Asp Ile Asp Glu Met

1975 TGG GAA GGG GTC CTT CTC TGT CAG TTT CAT GAC TGC TTG CCG GGT AGC TCC ATT
Trp Glu Gly Val Leu Leu Cys Gln Phe His Asp Cys Leu Pro Gly Ser Ser Ile

2029 GAA ATG TGT TAT GAC GAC TCA GAC GAG CTA TAC GCC CAG ATA TTC GAG ACG GGA
Glu Met Cys Tyr Asp Asp Ser Asp Glu Leu Tyr Ala Gln Ile Phe Glu Thr Gly

2083 TTG CGA GCA AGG GAG GCG GCT TTG GAG GCG CTA GGA TTT AGC ATG GCG GGG AGA
Leu Arg Ala Arg Glu Ala Ala Leu Glu Ala Leu Gly Phe Ser Met Ala Gly Arg

2137 GTT CCA GTG GCA ATT AAC ACC CTG CCT TGG AAT AGA TCC GAG GTG GTC AGA CTC
Val Pro Val Ala Ile Asn Thr Leu Leu Pro Trp Asn Arg Ser Glu Val Val Arg Leu

2191 CCT GCC GAC ATG GCC ACA AAT GGC TCC ATG CAT GCT CTA CTG AAC GGG GGT CCT
Pro Ala Asp Met Ala Thr Asn Gly Ser Met His Ala Leu Leu Asn Gly Gly Pro

2245 GGA GTC ATT CCC CAC TCG CAA GCC ATA GCG AGG GTC CCC GAG GTC AGC GTG ACC
Gly Val Ile Pro His Ser Gln Ala Ile Ala Arg Val Pro Glu Val Ser Val Thr

2299 GAG ATC AAG CCA GGC GTG TTC CGT CTT AAC AAC GGA AAA CTG CGC GTT GAT GTA
Glu Ile Lys Pro Gly Val Phe Arg Leu Asn Asn Gly Lys Leu Arg Val Asp Val

2353 GAA GAC GGT CTC ATC ACG TAT CTA TAT GAT ATT GAA GCA GAC CGC GAG GTT ATT
Glu Asp Gly Leu Ile Thr Tyr Leu Tyr Asp Ile Glu Ala Asp Arg Glu Val Ile

2407 GCT GCT GAC GGC AGA GCT GGC CAG CTG GTT ATC TTC GAC GAC AAG CCT CTT TAC
Ala Ala Asp Gly Arg Ala Gly Gln Leu Val Ile Phe Asp Asp Lys Pro Leu Tyr

2461 TGG CAA GCC TGG GAT GTC GAA GTC TTC CAC CTC GAT TCC CGG AAA GAA TTG CAG
Trp Gln Ala Trp Asp Val Glu Val Phe His Leu Asp Ser Arg Lys Glu Leu Gln

2515 TCT GGC AAG ACC ACC ATC GCG GAG AAC AAT CCC TAC CGT GTC AGT GTC GTT ACT
Ser Gly Lys Thr Thr Ile Ala Glu Asn Asn Pro Tyr Arg Val Ser Val Val Thr

2569 GAG ACA AAA ATC AGC GAA AAG AGC TGG GTC AAG ACG ACC ATC AGC CTC GCC GCC
Glu Thr Lys Ile Ser Glu Lys Ser Trp Val Lys Thr Thr Ile Ser Leu Ala Ala

2623 ACA GAA GGA AAC CGA CCA TCC TAC GTC GAG CTC GAA AGC GAA GTC GAA TGG CAG
Thr Glu Gly Asn Arg Pro Ser Tyr Val Glu Leu Glu Ser Glu Val Glu Trp Gln

2677 GAG ACT ATG AAG TTC CTT AAA GTC GAG TTT CCC GTC GAC ATC ACA AAC ACC GAA
Glu Thr Met Lys Phe Leu Lys Val Glu Phe Pro Val Asp Ile Thr Asn Thr Glu

2731 GCG TCT TAC GAA ACC CAA TAC GGC ATT ATC AGG CGC CCA ACT CAC TAT AAC ACA
Ala Ser Tyr Glu Thr Gln Tyr Gly Ile Ile Arg Arg Pro Thr His Tyr Asn Thr

2785 AG GTATATTTCCTCCCTTTCCCGCTCGCTCATATACTAAAGTTCTATAG C TGG GAT ATG GCT
Ser intron 3 Trp Asp Met Ala

2847 AAA TTC GAA GTC TGC GCA CAC AAA TGG GCC GAC CTC TCC GAG AAC GGG TAC GGG
Lys Phe Glu Val Cys Ala His Lys Trp Ala Asp Leu Ser Glu Asn Gly Tyr Gly

2901 GTC TCC ATC CTT AAC GAC TCC AAA TAC GGG TTC GCA ACT TGC GGA AAC CTG ATG
Val Ser Ile Leu Asn Asp Ser Lys Tyr Gly Phe Ala Thr Cys Gly Asn Leu Met

FIGURE 3 (cont.): Full-length DNA sequence of *Aspergillus nidulans* α -mannosidase.

2955 CGT CTT TCC CTC CTC CGA GCT CCC AAG GCG CCA GAC GCA CAC GCT GAC ATG GGT
 Arg Leu Ser Leu Leu Arg Ala Pro Lys Ala Pro Asp Ala His Ala Asp Met Gly
 3009 CGT CAT ACA ATC CGC TAC CGT ATC TTC CCA CAC AGC GGA CCT TTA GAC TTC CGG
 Arg His Thr Ile Arg Tyr Arg Ile Phe Pro His Ser Gly Pro Leu Asp Phe Arg
 3063 ACA GTC CGT GCT GGC TAT GAG TTC AAC CAT CCG CTC ATC CTC GAT ACC TCG TTA
 Thr Val Arg Ala Gly Tyr Glu Phe Asn His Pro Leu Ile Leu Asp Thr Ser Leu
 3117 GCT TTA CCC GAT GAG CAA AGT GTC AAC GCC GTT AGC GCG CTC TTC GTA AAA TTG
 Ala Leu Pro Asp Glu Gln Ser Val Asn Ala Val Ser Ala Leu Phe Val Lys Leu
 3171 AAC TAC CGG CTC TCC ATC TTT GAT CTC GAC GTT GTC AAG CGC GGC GAA GAC GAT
 Asn Tyr Arg Leu Ser Ile Phe Asp Leu Asp Val Val Lys Arg Gly Glu Asp Asp
 3225 GAG GAT ATT GAT GGT GAT CCT GAT ATT TCG CTT CCC AGC TCA AAA CGG AGC GGT
 Glu Asp Ile Asp Gly Asp Pro Asp Ile Ser Leu Pro Ser Ser Lys Arg Ser Gly
 3279 CAA AGT ATC ATT CTA CGG ATC TAT GAG AGT CTT GGT GGG AAG AGT AGG GGT GCA
 Gln Ser Ile Ile Leu Arg Ile Tyr Glu Ser Leu Gly Gly Lys Ser ArgG ly Ala
 3333 ATT CGG ACC GCT CTC CTG TAG AAAGGTGGGAAATAAAATGTCGTTTCGTATG
 Ile Arg Thr Ala Leu Leu *

FIGURE 3 (cont.): Full-length DNA sequence of *Aspergillus nidulans* α -mannosidase.

comparing the sizes of amplification products derived from reverse transcription (RT) PCR of *A. nidulans* mRNA and from genomic DNA using primer pairs designed to flank the putative intron sequences. RT-PCR amplification of the mRNA using the primer pair I1F/I1R (Table I) produced a fragment which was 103 bp shorter than amplification from genomic DNA. Sequencing of the RT-PCR fragment confirmed that the portion of the gene corresponding to the first intron (Intron I -103 bp) had been spliced out of the final transcript. Similarly, amplification with primer pair I2F/I2R and primer pair I3F/I3R (Table I) confirmed the presence of a 52 bp intron (Intron II) beginning 758 bp downstream from the start site and a 47 bp intron (Intron III) located 2785 bp downstream from the start site. This ORF produces a 1049 amino acid protein with a calculated molecular weight of 118 kD. Six potential N-glycosylation sites (Asn-X-Ser/Thr, where X is any amino acid except proline) were identified. The putative protein sequence aligned well with the rat ER/cytosolic and *S. cerevisiae* vacuolar sequences (Figure 4).

The *Aspergillus* peptide sequence was then analyzed to find regions such as ER signal sequences and transmembrane domains, which are typical of sorting or targeting signals. A Kyte-Doolittle hydropathy plot did not suggest the presence of any significant transmembrane regions. A PSORT search (Nakai and Kanehisa, 1992), which predicts protein localization based on amino acid composition, suggested that the *A. nidulans* α -mannosidase did not have an N-terminal signal sequence or any regions which would be typical of a vacuolar or lysosomal targeting signal.

```

1                               60
MGGDVLRQTPSSLPAPADAVPVGQRIRSIYTDRLNQFTSRGQ-YQNQNLIKLFPEATNS-D   Aspergillus
-----MSSEDIYDPQFKPVQGIYENRLRQFIDTGGDYHDLNLPKFDKKRISLD   Saccharomyces
      MAAAPFLK-----HWRTTFERVEKFSVP-IYFTDCNLRGRLFG-----D   Rattus
            o o          oo oo oo  o o o oo o oo          o
-----S-AP-P-----IY--RL-QF---G--Y-D-NL-G-LF-----D   Consensus

61                               120
EDHVKISVYSVP-----DLQRPST-----EEATSNEFQPTHIGASFGPSWSTHWF   Aspergillus
HDHVKVVWVQVSFERGSSPVSPDKRPSWKSIIERDKKGELEFREANINQPFGPSWSTTWF   Saccharomyces
SCPVTLS--FLTPERLPY-----EKAVQONFSPAQVGDSPGPTWWTWTF           Rattus
ooo o o o oo          ooo          ooo oo oo oo oooooooo oo
-DHVK-S-Y-V--ER-----RPS-----EKA--EF-PA-IG-SFGPSWST-WF   Consensus

121                              180
GIRLTIPEDMRER-ERLEFHWANNENGLVWTEGHPLOGLTGGGERIEWIIPDAW--ADG   Aspergillus
KVKIISLPEDWVKSNQQLLFQWDCSNEGIVIDPKTLIPVTAFSGGERTYVLPK----TSD   Saccharomyces
RVELVIPEVWVGKVEHLWCWESDG--ESLVW-RDGEFVQGLTKEGEKTSYVLSERLHAADP   Rattus
o o oooooo  o o o oo  oooooo  oo o oooo  oooooooo  oo
-V-L-IPEDWV--E-L-F-WD--NEGLVW--DG-P-QGLT-GGERTYVLP-----AD-   Consensus

181                              240
NEHVFIEMACNGMFGNAPGGDSIQPPAPDKYYTLTARITAVNLEARALFYDFWIISDA   Aspergillus
GKHFFYIEAGNNGMFGCGAGSTINPP-DDNRFHLRKADIVWPDLDARALYIDFWMLGDA   Saccharomyces
RSLTLYVEVACNGLLGAGKGSMAAP-DPEKMFQLSQAQKLVFHRDVHNLVLDLELLLV   Rattus
o oooo  oooooo  o oo o  o oo o  o o  o  o  oooooo  ooo o oo
--H-FYIE-ACNGMFG-G-GS-I--P-DP-K-F-L--A-I---LDARAL--DFW-L-DA   Consensus

241                              300
AREFPAEGWESHEANMVANSIIDAFIAGNLSLESIKEGRRIAQKYLGDKVCSEVY--DS   Aspergillus
ARELPGDSWQKHQARQLGNVAVMLFDPNDRS SVRKCRELLQREYFDSFLESSKVYEQGES   Saccharomyces
AKGLGEDNQRSFQALYTANQMVNICDPAQPETYPAEALASKFFGQR-----GG      Rattus
ooooo  o o oooo  oo  o ooo  o  o  o  o  oooooo  ooo o oo
ARELP-D-W-SHQA---AN---N-FDP---S-----E-L-----S           Consensus

301                              360
GTQPIVYGIGHCHIDTCWLWPFPAETKRKVARSWSNQCCLMDRYPEHRFTCSQAQFQKWLK   Aspergillus
QVLTNYYGIGNCHIDTAWLWPFPAETRRIKIVRSWSSQCTLMDRFPEYKFKVASQAQFQKWL   Saccharomyces
ESQHTIHATGHCHIDTAWLWPFKETVRKCARSWSTAVKLMERNTEFTFACSAQQLWVVK   Rattus
o  oooooo  oooooo  oooooo  oo ooooo  oo oooo  oo  o oooooo  ooooo
--Q--VYGIGHCHIDTAWLWPFPAET-RK-ARSW-S-QC-LMDR-PE--F-CSQAQFQKWLK   Consensus

361                              420
QYYPVFDVRVKGW-VKKGHFQPIGGSWVEHDTNMPSGESLVRQFIYQQLFESKFGERCT   Aspergillus
EDHPEFFNKVLIPIKIQSQFFAVGGTWEVENDTNI PSGESLARQFFFGQRFLLKHFLGKSK   Saccharomyces
NQYPLGLYAQLQEF-ACRGQFVPGGTWVMDGNLPSGEAMVRQFLQGQNFLLQEFQKMC   Rattus
oo  o  o          ooo  oooooo  ooo  oooooo  oooooo  oo  o
--YP--F--V-----GQF-PVGGTWVE-DTN-PSGESLVRQF--GQRFFL--FG--C-   Consensus

421                              480
TFWLPDTFGYSTQIPQLCRLAGMSRFFTKLQKLSWNNINFPHTTFQWVALDGSQVMCHMPP   Aspergillus
IFWLPDTFGYSSQMPQLCRLSGIDKFLTKLQKLSWNNINSFPHSTFNWAGIDGSQLLTHMPP   Saccharomyces
EFWLPDTFGYSAQLPQIMQCGIKRFLTKLQKLSWNLVNSFPHTTFWEGLDGSQVLVHFP   Rattus
ooooo  ooooo  o  oooooo  oooooo  oooooo  oo  o  oooooo  oooo
-FWL PDTFGYS-Q-PQLCRL-GI-RFLTQKLSWNNINSFPH-TF-W-GLDGSQVL-HMPP   Consensus

```

FIGURE 4: Sequence alignment of *Aspergillus nidulans* α -mannosidase, rat ER/cytosolic α -mannosidase, and yeast vacuolar α -mannosidase. Alignment was prepared using the ClustalW algorithm in the DNASTar sequence analysis package. Open circles represent amino acid identity between two of the three sequences, and closed circle represent perfect amino identity between all three sequences. The consensus sequence is derived from both of these sequence similarities. The sequence within brackets is the region which is conserved between all Class 2 α -mannosidases.

481	540		
SETYTAEAHFGDLKRSITQHKSLDKDNTSLLVFGKGDGGGGPTFEHLEKLRRCRGLSDKT		<i>Aspergillus</i>	
GNTYTADSHFGDVLRTAKQNKTPYYGSGMLLYGKGDGGGGPTEMLQKMRRIIRSMNRRN		<i>Saccharomyces</i>	
GDSYGMQGSVEEVLTVTNNRDKGRTNHSGFLFGFDGGGGPTQMLDRLKRLGNTDGGQP		<i>Rattus</i>	
o oooo oooooooo oooo o oo oooooo oooooo ooo oooo o			
G-TYTA--HFGDVLRT-TQNK-----N-SL-LFGKGDGGGGPT-EML-KLRR-R-----		Consensus	
541	600		
GLLPRATMGTSVDDFFAKLEKAA-DGTFATWYGELYFELHRGTYTTQANNKRNRRAE		<i>Aspergillus</i>	
GNVIPKLQVGITVDEFYDDILKRTNQGHDLPTWSGELYFHFHRGTYTSQAQTKMLRLSE		<i>Saccharomyces</i>	
RVQLSSPQLFTALER-----DSGQLCTWVGELFLELHNGTYTTAQLKKNRECE		<i>Rattus</i>	
o o ooo o oo o oo oooooo oooooo ooo oo o			
G-----T-DEF-----K---DG--L-TW-GELYFELHRGTYTTQAQ-KK-NR--E		Consensus	
601	660		
FLLRELEFLATIASITR-SDYEYPKKIDIDEMWEGVLLCQFHDCLPGSSIEMCYDSDLEY		<i>Aspergillus</i>	
IKLHDLIEWIAAKTSVLYPDSYKYPKQINELWENVLLCQFHDVLPGSCIEMVYKYEAVPM		<i>Saccharomyces</i>	
QILHDVELLSSALARS-AQFLYPAVQLQRLWRLLLNQFHDVVTGSCIQLVAEDAMNYY		<i>Rattus</i>	
ooooo oo oo o oo oooo oooooo oooooo oooooo o o			
--LHDLE-LA--AS-----Y-YP-KQI-ELWE-VLLCQFHDVLPGSCIEMVY-D----Y		Consensus	
661	720		
AQIFETGLRAREAALEALGFSMAGRVPVAINTLPWNRSEVVRPADMATNGSMHALLNGG		<i>Aspergillus</i>	
LHNVVKECTSLIDKTVQFLQSQSKADLVEMRTLWTSKPEKVSEECSLNGSYTSSVTGYDD		<i>Saccharomyces</i>	
EDIRSHGNTLLSAAAAALCAGEPFGPKGLRHYQHTALEAHRVGTQGLVGLTRLALVTP		<i>Rattus</i>	
o o o o oo oo o o o oooo o o o oo			
--I--G-T-L-AA--AL--S--G--V---TLTW---E-V-----T--AL----		Consensus	
721	780		
PGVIPHSAIARVPEVSVTEIKPGVFRLLNNGKLRVDVEDGLITYLYDIADREVIADGR		<i>Aspergillus</i>	
YIVLANGKLVIIICKKTGVITSITDETLGVEYLDTEHGRNKL-----		<i>Saccharomyces</i>	
SIGYAPAPTPTSLQPLLQQPVFVMQETDGSVTLDNQIIRVRLDPTGCLTSLVIVASGRE		<i>Rattus</i>	
oo o			
-IV-A-----L---L-----A----		Consensus	
781	840		
-----AGQLVIFDDKPLYWQAWDVEVFLHDSRKELQSGKTTIAENNPYRVSVVETKI		<i>Aspergillus</i>	
-----ANQFVIYDDKPLGWQAWDTELYSVNQYKYVTKPKVQLSCNTKEKCAVEAIFQ		<i>Saccharomyces</i>	
AIAEGALGNQFVLFDDVPLYWDAWDMYHLETRKPFVRGQAGTLAVGAEGLRGSAWFL		<i>Rattus</i>	
ooooooooooooooooooooo ooo oo o o o o o o			
-----ANQFVIFDDKPLYWQAWDVE-YHL--RK-V---K-T-A--N-----V-----		Consensus	
841	900		
SEKSWVKTISLAATEGNRPYSVELESEV---EWQETMKFLKVEFPVDITNTEASYETQ		<i>Aspergillus</i>	
ISEKCKIKSVISLNAVTDALSKVDISTTVENWDARNKFLKVEFPVNIRNDFASYETQ		<i>Saccharomyces</i>	
QISPNRSLSQEVVLDVGCYVRFHTEV-----HWHETHKFLKVEFPARVRSPOATYEQ		<i>Rattus</i>	
o o o o o o o o o o o o o oooooo ooo oooooo			
-----S-----G-----E-----W-ET-KFLKVEFPV-IRN--ASYETQ		Consensus	
901	960		
YGIIRRPHTYNTSWDMAKFEVCAHKWADLSENGYGVSIILNDSKYGFATCGNLMRLSLLRA		<i>Aspergillus</i>	
FGITKRPHTYNTSWDVAKFEVCHHKFADYSEYSKGVSIILNDCYGFSTHGNLMRLSLLRS		<i>Saccharomyces</i>	
FGHLQEAADPQQHSWDWARVEVWAHRWIDLSECDGLALLNCKYGTSVRGNVLSLLRA		<i>Rattus</i>	
ooo oooooo oooooo oooooo oooooo oooooo oooooo oooooo oooooo			
FGI--RPTHYNTSWD-AKFEVCAHKWADLSE---GVSIILNDCYGFST-GNLMRLSLLRA		Consensus	
961	1020		
PKAPDAHADMGRHTIRYRIFPHSGPLDFRTVRAGYEFNHPLILDTSALPDEQSVNAVSA		<i>Aspergillus</i>	
PKAPDAHADMGTHEIKYAIYPHRGALSSDTVKLAHEFNCFKYKLPKDIGMNFDDIISIS		<i>Saccharomyces</i>	
PKAPDVTADMGRHEFTYALMPHKSFSQEAQVIQAAYNLNFPLLALPAPGPAPDTTWSAFS		<i>Rattus</i>	
ooooooooooooooooooooo ooo oo o o oo o ooo o oo o o			
PKAPDAHADMGRHEI-YAI-PH-G-L---TV--A-EFN---L-LP---P-----S		Consensus	

FIGURE 4 (cont.): Sequence alignment of *Aspergillus nidulans* α -mannosidase, rat ER/cytosolic α -mannosidase, and yeast vacuolar α -mannosidase.

1021	1080	
LFVKLNRYLSIFDLVVKRGEDEDEDIDGDPDISLPSSKRSGQSIILRIYESLGGKSRGAI		<i>Aspergillus</i>
GDENVILSNIKRGEEDSAVKSNYSLNP-----RDEQSIVRVVYESLGGESFASL		<i>Saccharomyces</i>
VSSPAVVLETI-----KQAEKSHQHRTLVLRLYEAGSH		<i>Rattus</i>
o o oo oooo o oooooooo oo		
-----I--D-----KR-EQSI--R-YESLGGES-GS-		Consensus
1081		
RTALL-----		<i>Aspergillus</i>
NTTLNLKRIEKVDNLEMKVYKSLTATRDESNHAINRIPIKLRPFELASFRLYF----		<i>Saccharomyces</i>
VDCWLHTSLPVQEATLCLLLEQRDPTGHLSLQDNRLKLTFSPPQVRSLLLVLPAN		<i>Rattus</i>
oo o o		
-T-LL-----T--S-----		Consensus

FIGURE 4 (cont.): Sequence alignment of *Aspergillus nidulans* α -mannosidase, rat ER/cytosolic α -mannosidase, and yeast vacuolar α -mannosidase.

2.3.2 Disruption of α -mannosidase in *A. nidulans*

To determine whether this α -mannosidase was essential to normal cellular function, we interrupted the coding sequence of the *A. nidulans* gene. A disruption cassette p8 Δ KO (Figure 5a) containing stop codons in all three reading frames of the α -mannosidase gene was used to replace the endogenous copy of the gene. Homologous integration of the disruption cassette was confirmed by PCR amplification. The forward primer (KOconf-F) was specific for the α -mannosidase gene upstream of the recombination target while the reverse primer (KOconf-R2) was homologous to a unique region of the disruption cassette (Table I). This primer combination yielded a 1.15 Kb product for transformants that integrated the disruption cassette at the authentic α -mannosidase locus. This was confirmed by sequence analysis of the 1.15 Kb product. The PCR amplification product from the transformant AnKO-7 consisted of authentic α -mannosidase sequence immediately upstream of the disruption target as well as sequence derived from the disruption cassette including the stop codons and restriction sites unique to the disruption cassette. Integration of the stop codons was expected to result in the production of a truncated protein product consisting of the first 478 amino acids of the α -mannosidase protein having a calculated molecular mass of 58.6 kD.

As illustrated by the PCR analysis, the disruption cassette integrated at the authentic α -mannosidase locus, but the cassette could also integrate at several non-homologous sites which would not be detected by this method. To reduce potential phenotypic changes due to non-homologous integration events (ie. insertional mutagenesis), it was desirable to produce disruption transformants with a minimal number of non-homologous integrations. Transformants were examined by Southern analysis to

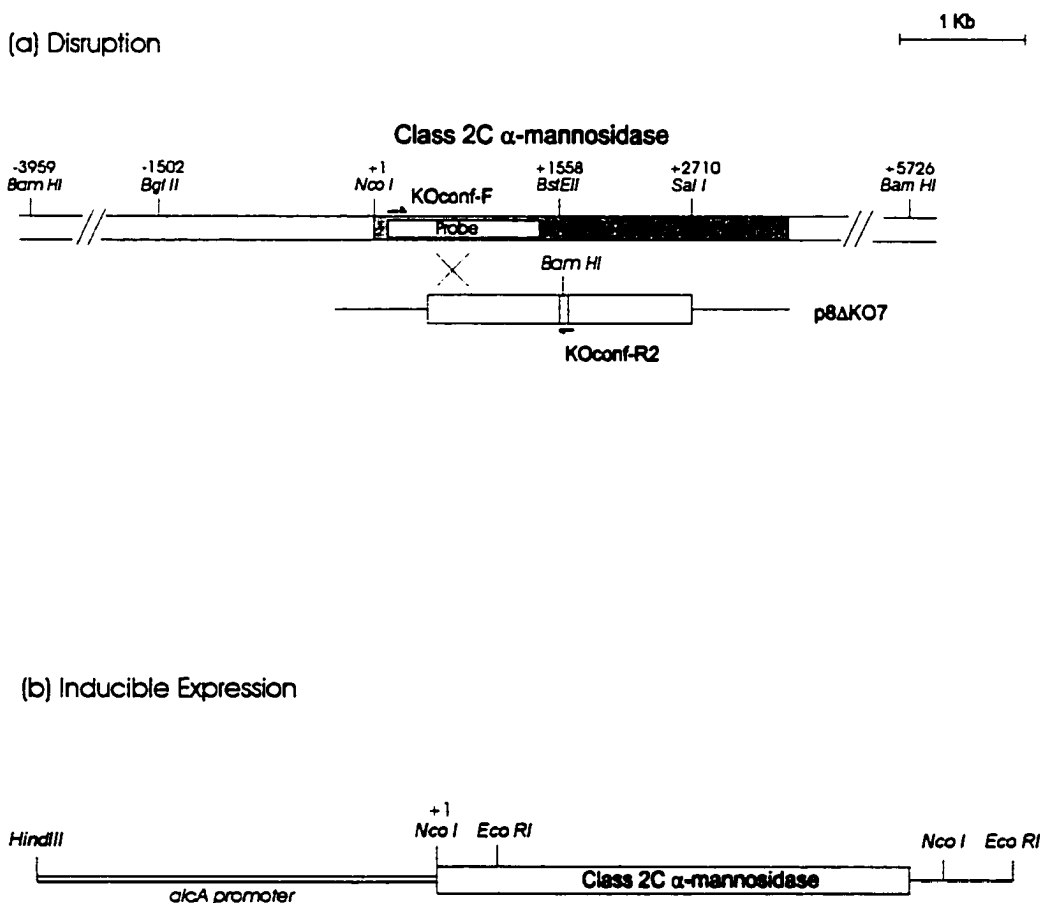


FIGURE 5: Disruption/induced expression of the α -mannosidase gene. The disruption cassette p8 Δ KO (a) was introduced into wild type T580 and *Aspergillus nidulans* transformants were screened for homologous integration of the cassette by PCR using the diagnostic primers KOconf-F and KOconf-R2. The disruption cassette contained a unique *Bam*HI restriction site which was used for diagnostic purposes, as well as stop codons in all three frames which caused disruption of the gene (boxed region). A fragment of the α -mannosidase gene (position + 424 to + 1507) was used for Southern hybridization to confirm proper integration of the cassette. The inducible expression cassette palcAman (b), was designed by direct fusion of the *alcA* promoter to the coding region of the α -mannosidase gene carried in the pUC19 vector. The plasmid was introduced into T580 to overexpress the α -mannosidase gene in *Aspergillus nidulans*. Integration of the plasmid into the genome was PCR screened using the forward primer alcAprom-F, specific for the *alcA* promoter, and manR-conf, specific for the α -mannosidase gene and overexpression was confirmed with mannosidase assays.

determine the number of copies of the cassette integrated at both the authentic and non-homologous sites. Homologous replacement of the authentic DNA with the disruption cassette was expected to result in the insertion of a novel *BamHI* site at position +1575 of the coding region of the gene (Figure 5a). When the radio-labelled α -mannosidase DNA was hybridized to wild type (T580) and the disruptant (AnKO-7) genomic DNA cut with *BamHI* and *BglII* (Figure 6), the control T580 DNA (Lane 1) showed a single 7.2 Kb fragment absent from An-KO7 (Lane 2). Three additional fragments of 3.0 Kb, 5.9 Kb, and 10.0 Kb were evident in the digested DNA of the disruptant. The 3.0 Kb fragment resulted from the truncation of the 7.2 Kb *BglII/BamHI* fragment due to integration of the disruption cassette at the α -mannosidase gene, while the 5.9 Kb and 10.0 Kb fragments likely resulted from non-homologous integration events. Since other transformants demonstrated numerous non-homologous integrations (not shown), transformant AnKO-7 was chosen for further analysis.

To confirm that the gene disruption did indeed eliminate α -mannosidase activity, enzyme activities of the wild type (T580) and the disrupted strain (An-KO7) were assayed using p-nitrophenyl- α -D-mannopyranoside (PNP) as a substrate. Enzyme activity could not be detected when wild type protoplasts were osmotically lysed by the addition of water, but low levels of activity were found when protoplasts and all subcellular components were lysed by the further addition of 1% TritonX-100. Since the levels detected were still quite low, the cell lysates were concentrated on a 100 kD filter to improve assay resolution. The α -mannosidase activity of the wild type remained constant after 38 and 48 hours of growth (Figure 7). No α -mannosidase activity was detected at

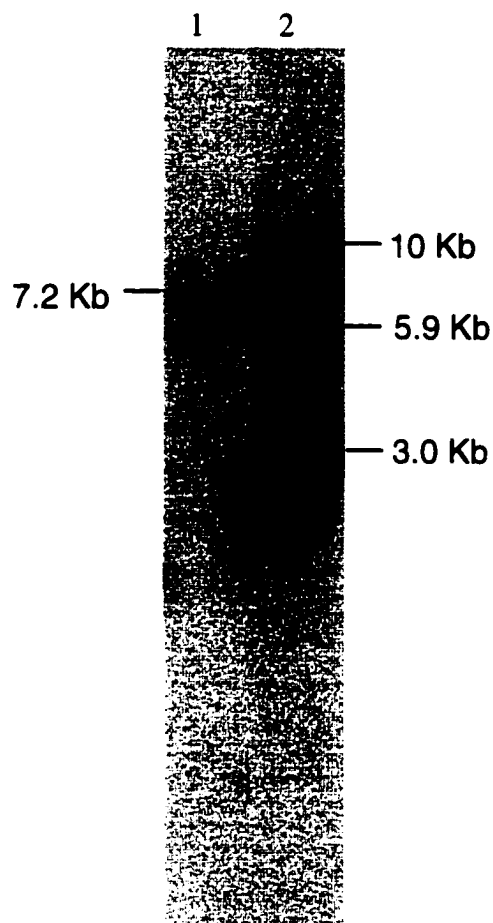


FIGURE 6: Confirmation of gene disruption. Total genomic DNA was digested with *BglII/BamHI* and probed with position + 424 to + 1507 of the α -mannosidase gene. The probe hybridized to a single 7.2 Kb fragment in the wild type (T580) (Lane 1) which was absent from the disruptant (AnKO-7) (Lane 2), indicating disruption of the α -mannosidase locus.

α -Mannosidase Expression in *Aspergillus nidulans*

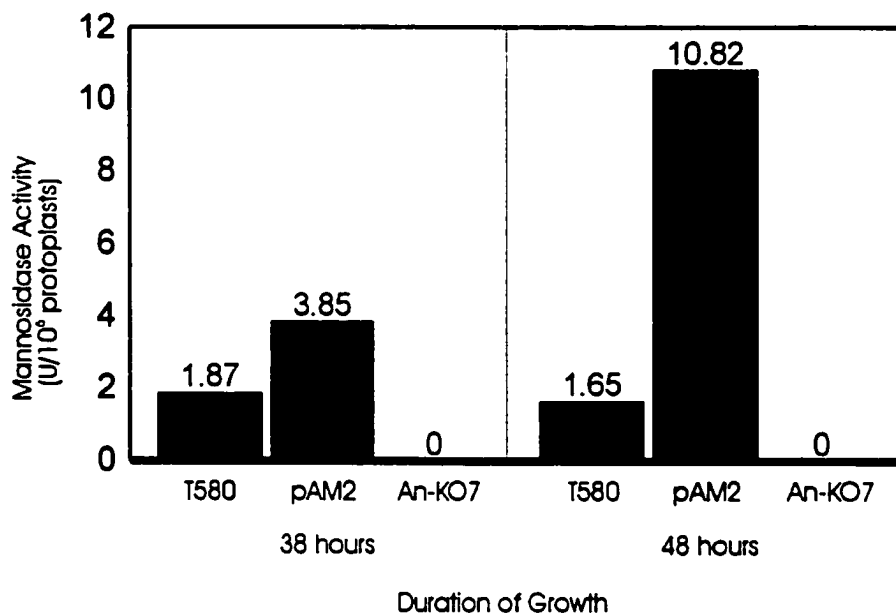


FIGURE 7: Expression of cytosolic α -mannosidase in *Aspergillus nidulans*. Three strains of *A. nidulans*, wild type (T580), disruption (AnKO-7) and overexpression (pAM2) were grown in YFT inducible media and mycelium was harvested at 38 hours and 48 hours of growth. Mannosidase activity was assayed using the substrate p-nitrophenyl- α -D-mannopyranoside (PNP). Activity is expressed as units of enzyme activity (U) per 10⁶ protoplasts.

either time point in the disrupted strain AnKO-7. The disrupted strain did not show any alteration in growth or morphology, as compared to wild type.

2.3.3 Induced expression of α -mannosidase in *A. nidulans*

To mitigate against possible toxic effects due to increased expression of α -mannosidase in *A. nidulans* we used a phased expression system to overexpress the *alcA*-driven α -mannosidase protein. The α -mannosidase gene was fused to the inducible *alcA* promoter (Figure 5b) and this expression cassette was introduced into *A. nidulans* strain T580 by cotransformation with the selectable marker pFB94. The *alcA* system permitted the accumulation of biomass during the first 24-36 h of growth. Upon glucose depletion *alcA*-driven expression was induced and expression levels monitored. Putative transformants were screened for integration of the expression cassette using an *alcA* promoter-specific primer, alcAprom-F, in combination with an α -mannosidase-specific return primer, manR-conf (Table 1). These primers did not amplify from the endogenous α -mannosidase sequence and were diagnostic for the integration of the overexpression cassette. The induced expression transformant pAM2 was identified using this primer pair. To verify induced expression of α -mannosidase, pAM2 was cultured in inducing media and assayed for α -mannosidase activity using the aryl substrate PNP. The transformant exhibited a twofold increase in activity after 38 hours of growth and a sixfold increase in activity after 48 hours, as compared to T580 (Figure 7). These transformants did not exhibit any toxic or deleterious effects, and did not show visible alterations in morphology or growth.

2.3.4 Sequence comparison with other α -mannosidases

The α -mannosidases have traditionally been divided into Class 1 and Class 2 sequences based on substrate specificity, inhibitor sensitivity, and sequence alignments (Daniel *et al.*, 1994; Moremen *et al.*, 1994). To classify the *A. nidulans* α -mannosidase, and to confirm previous classification schemes, a similarity matrix was generated for 18 complete α -mannosidase protein sequences, including the *A. nidulans* α -mannosidase, using the Matchbox alignment program (Figure 8). This program calculates the global similarity, using the BLOSUM62 scoring matrix (Henikoff and Henikoff, 1992), between each possible pair of sequences by comparing successive amino acid segments of length 9 from one sequence with all possible segments from the second sequence (Depiereux and Feytmans, 1992). The coefficient (R_{ij} , where $0 \leq R_{ij} \leq 1$) generated for each pair of sequences represents the proportion of segments of sequence i matching with at least one segment of sequence j . This alignment method provides a more objective value for sequence similarity than other programs, since it does not depend on an alignment of all sequences to compute similarity values, but rather, scans each possible pair of sequences and computes similarity coefficients independent of other sequences used in the alignment. Several blocks of sequences exhibited high similarity to each other, but low similarity to the rest of the sequences in the matrix, as indicated by the boxed regions in Figure 8. The *A. nidulans* α -mannosidase, rat ER/cytosolic α -mannosidase and yeast vacuolar α -mannosidase were highly similar (Box 1), with R_{ij} values ranging from 0.29 to 0.43. The *A. nidulans* α -mannosidase showed higher similarity to the *S. cerevisiae* vacuolar α -mannosidase ($R_{ij} = 0.43$) than to the rat ER/cytosolic α -mannosidase ($R_{ij} = 0.33$). This program did not find a significant global similarity between these sequences and the rest of

	CLASS 2 α -mannosidases									CLASS 1 α -mannosidases								
	Box 1			Box 2			Box 3			Box 4								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.00	0.33	0.43	0.04	0.02	0.03	0.03	0.03	0.03	0.02	0.01	0.02	0.01	0.01	0.03	0.02	0.01	0.01
2	0.32	1.00	0.29	0.03	0.05	0.03	0.03	0.03	0.04	0.01	0.01	--	0.02	0.01	0.01	0.02	0.02	0.01
3	0.42	0.27	1.00	0.05	0.03	0.03	0.04	0.04	0.02	0.02	0.01	0.02	0.01	0.01	0.02	0.02	0.02	0.01
4	0.04	0.03	0.05	1.00	0.33	0.33	0.13	0.16	0.14	0.01	0.01	--	0.01	0.01	--	0.02	0.01	0.03
5	0.03	0.05	0.03	0.34	1.00	0.79	0.14	0.14	0.13	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	--
6	0.04	0.04	0.04	0.33	0.77	1.00	0.14	0.14	0.13	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.02	0.01
7	0.03	0.03	0.04	0.12	0.13	0.13	1.00	0.36	0.36	0.03	0.02	0.02	0.03	0.02	0.01	0.01	0.01	0.02
8	0.02	0.03	0.04	0.14	0.12	0.13	0.35	1.00	0.89	0.01	0.01	0.02	0.03	0.02	0.01	0.01	0.02	0.01
9	0.03	0.03	0.02	0.13	0.11	0.11	0.35	0.89	1.00	0.02	0.02	0.03	0.03	0.02	0.01	0.01	0.02	0.02
10	0.04	0.02	0.03	0.01	0.02	0.04	0.05	0.02	0.03	1.00	0.71	0.93	0.70	0.48	0.20	0.16	0.19	0.15
11	0.02	0.03	0.03	0.02	0.02	0.03	0.04	0.03	0.04	0.97	1.00	1.00	0.89	0.65	0.27	0.23	0.25	0.23
12	0.03	--	0.03	--	0.02	0.03	0.04	0.04	0.05	0.96	0.75	1.00	0.71	0.47	0.20	0.15	0.19	0.14
13	0.02	0.03	0.02	0.02	0.03	0.03	0.04	0.05	0.05	0.70	0.65	0.68	1.00	0.48	0.22	0.17	0.20	0.14
14	0.02	0.01	0.01	0.02	0.02	0.02	0.04	0.03	0.04	0.48	0.48	0.46	0.49	1.00	0.19	0.15	0.20	0.14
15	0.06	0.03	--	0.01	0.01	0.01	0.03	0.02	0.03	0.23	0.23	0.23	0.25	0.23	1.00	0.13	0.14	0.12
16	0.03	0.04	0.03	0.03	0.02	0.03	0.02	0.03	0.03	0.21	0.21	0.19	0.22	0.20	0.14	1.00	0.85	0.13
17	0.02	0.04	0.03	0.02	0.03	0.03	0.03	0.04	0.04	0.24	0.23	0.22	0.25	0.24	0.14	0.85	1.00	0.13
18	0.02	0.03	0.03	0.05	0.01	0.01	0.04	0.03	0.03	0.17	0.19	0.16	0.17	0.15	0.12	0.12	0.13	1.00

FIGURE 8: Matchbox sequence similarity matrix. Coefficients of similarity (R_{ij}) of α -mannosidase protein sequences using Matchbox algorithm, where R_{ij} is the proportion of amino acid segments (length 9) of sequence i matching with at least one segment of sequence j . Boxed regions represent groups of sequences which show a relatively high degree of peptide sequence similarity to each other, but not to other sequences in the matrix. Box 2 and Box 3 also show some limited similarity (dotted lines), and are more similar to each other than they are to Box 1 or Box 4. Protein sequences are numbered as follows: 1) *A. nidulans* α -mannosidase; 2) rat ER/cytosolic α -mannosidase (Bischoff *et al.*, 1990); 3) *S. cerevisiae* vacuolar α -mannosidase (Yoshihisa and Anraku, 1989); 4) *Dictyostelium discoideum* lysosomal α -mannosidase (Schatzle *et al.*, 1992); 5) human lysosomal α -mannosidase (Nebes and Schmidt, 1994); 6) mouse lysosomal α -mannosidase (Merkle and Moreman, 1993); 7) *Drosophila melanogaster* Golgi mannosidase II (Foster *et al.*, 1995); 8) human Golgi mannosidase II (Misago *et al.*, 1995); 9) mouse Golgi mannosidase II (Moreman and Robbins 1991); 10) mouse golgi mannosidase IA (Lal *et al.*, 1994); 11) rabbit liver mannosidase I (Lal *et al.*, 1994); 12) human Man₉-mannosidase (Bause *et al.*, 1993); 13) mouse Golgi mannosidase IB (Herscovics *et al.*, 1994); 14) *D. melanogaster* mannosidase I (Kerscher *et al.* 1995); 15) *S. cerevisiae* Man₉-mannosidase (Camirand *et al.*, 1991); 16) *P. citrinum* mannosidase I (Yoshida and Ichishima, 1995); 17) *A. satoi* mannosidase I (Inoue *et al.*, 1995); 18) *Caenorhabditis elegans* mannosidase I (Wilson *et al.*, 1994).

the sequences in the matrix ($R_{ij} < 0.05$). Box 2 and Box 3, containing the lysosomal α -mannosidases and mannosidase II sequences, respectively, also exhibited high internal similarity, with R_{ij} values between 0.33 and 0.89. These two groups of sequences, however, also showed some moderate similarity to each other, as indicated by dotted lines in Figure 8 (R_{ij} values between 0.11 and 0.16). The sequences in Box 2 and 3 showed very low similarity with the other sequences in the matrix (Box 1 and 4), with R_{ij} values not greater than 0.05. The sequences in Box 1, 2 and 3 are all Class 2 α -mannosidases, as defined by Moremen *et al.* (1994). Box 4 consisted of all of the Class 1 α -mannosidases, and showed high internal similarity, while exhibiting a very low similarity to any of the Class 2 α -mannosidases. Certain pairs of Class 1 sequences showed extremely high similarity (ie. mouse and human kidney α -mannosidase I: $R_{ij} = 0.93$), whereas others showed more limited similarity (*P. citrinum* and *C. elegans*: $R_{ij} = 0.12$), but this could be correlated to species divergence. When sequences in Box 4 were compared to other sequences in the matrix, the R_{ij} was consistently less than 0.03.

To further resolve these sequence relationships, we compiled a dendrogram using the ClustalW algorithm (Thompson *et al.*, 1994). Figure 9 shows the most parsimonious tree generated by the multiple alignment of all 18 α -mannosidase sequences. The Class 1 and Class 2 sequences were found to be unrelated and were represented by two separate trees. Similar to the Matchbox program, the lysosomal and mannosidase II sequences were found to be related, but tended to fall into separate subgroups (Class 2A and Class 2B). In contrast to the Matchbox program, however, the *A. nidulans*, rat ER/cytosolic, and yeast vacuolar α -mannosidases were found to be distantly related to the rest of the Class 2 sequences, and formed a distinct subgroup (Class 2C).

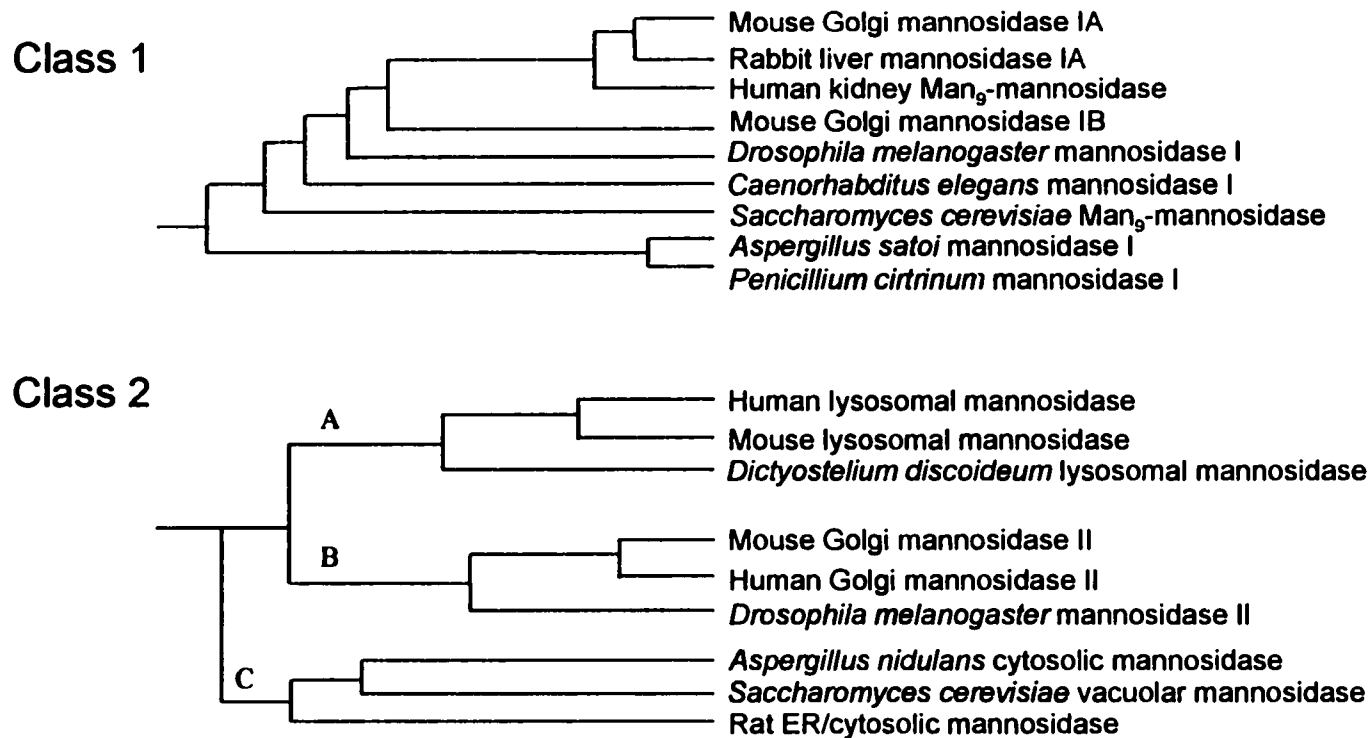


FIGURE 9: Dendrogram prepared from multiple sequence alignment of 18 α -mannosidase sequences. Alignments were performed using the DNASTar sequence analysis program which utilizes the ClustalW alignment algorithm. Two separate groups of sequences are shown representing groups of sequences which show significant homology. All 18 sequences were used in the initial alignment to define the two groups. Each of the two groups of sequences were then aligned to derive the sequence relationships within the groups. Class 2 sequences were further characterized by aligning the homologous clusters, and subdivided into three subgroups (2A, 2B, and 2C).

Since the Matchbox program did not find global similarity between the Class 2C and the Class 2A/2B sequences, yet the ClustalW dendrogram indicated that these sequences were distantly related, we further investigated this particular relationship between the Class 2 sequences. The Multiple Alignment Program (MAP) (Huang, 1994), which is more useful for identifying localized regions of similarity, revealed a single amino acid block which was similar between all Class 2 sequences. This block was found between amino acids 283 and 411 in the *A. nidulans* α -mannosidase (shown in Figure 4). This conserved 128 amino acid block confirmed that the Class 2C sequences were indeed related to the rest of the Class 2 sequences.

2.4 Discussion

The α -mannosidases play a key role in protein glycosylation, both in the modification of N-glycan chains prior to further elongation, and in the catabolism of oligosaccharides (Moremen *et al.*, 1994). Whereas mammalian glycoproteins generally contain complex type N-glycans, which may be necessary for the correct function of the glycoprotein, recombinant production of these same glycoproteins in lower eukaryotic expression systems leads to the addition of oligomannosidic N-glycans (Kalsner *et al.*, 1995). This can be a disadvantage of using fungal expression systems for the production of mammalian glycoproteins. We are currently characterizing the glycosylation pathways in the filamentous fungus *Aspergillus nidulans* so that we may better utilize this organism as an expression host. Our initial focus is on the role of the α -mannosidases in N-glycan processing.

Classification schemes have previously placed the α -mannosidases into two major groups, Class 1 and Class 2, based on biochemical properties, substrate specificity, inhibitor profiles, and sequence alignments (Daniel *et al.*, 1994; Moremen *et al.*, 1994). The *A. nidulans* α -mannosidase gene was expected to be a Class 2 α -mannosidase, as it was initially amplified based on similarity between the rat ER/cytosolic and *S. cerevisiae* vacuolar Class 2 α -mannosidases. To confirm this assignment, we performed a pairwise comparison of the *A. nidulans* α -mannosidase protein sequence to the complete protein sequences of 17 Class 1 and Class 2 α -mannosidases using the Matchbox algorithm (Depiereux and Feytmans, 1992). An advantage of this algorithm is that, unlike many sequence alignment programs, it does not rely on a user defined gap penalty. Sequence similarity was computed independent of the multiple sequence alignment thus producing the same result for any pair of sequences independent of the number of sequences included in the analysis. We found that the Class 1 sequences, while being highly similar to each other, were not related to the Class 2 α -mannosidases. The lysosomal α -mannosidases and mannosidase II sequences were found to be similar to each other, and tended to form two separate subgroups of higher similarity. Interestingly, the rat ER/cytosolic, *S. cerevisiae*, and *A. nidulans* α -mannosidases, while highly similar to each other, were not found to be globally similar to any of the other sequences. This analysis, while providing excellent overall sequence similarity coefficients, was not designed to locate localized regions of similarity, which may have been present in more distantly related sequences.

To identify localized regions of similarity, we utilized two alignment programs, ClustalW (Thompson *et al.*, 1994) and the Multiple Alignment Program (MAP) (Huang, 1994). Both of these programs detected small, yet significant regions of similarity between

all of the Class 2 sequences, supporting previous alignment based studies which suggested that these enzymes were all related (Daniel *et al.*, 1994; Moremen *et al.*, 1994). More recently, Henrissat and Bairoch (1993, 1996) also found that the Class 2 enzymes tended to belong to a single group of glycosylhydrolases (#38), whereas the Class 1 α -mannosidases belonged to a separate group (#47). The authors identified a protein domain which was common to all the Class 2 enzymes and likely represented a common catalytic domain. Our analysis, which included the *A. nidulans* α -mannosidase, emphasized the three subgroups of Class 2 α -mannosidases (2A, 2B, and 2C) which exhibited distinct sequence and biochemical properties. A dendrogram, compiled using the ClustalW alignment, showed that the Class 2C sequences were quite distantly related to the rest of the Class 2 α -mannosidases (Figure 9). Many conserved regions existed within the Class 2C enzymes (Figure 4) which were not found in the Class 2A or 2B enzymes. While the Matchbox algorithm was useful for objectively identifying overall sequence similarities and for defining the various groups of enzymes, more refined sequence alignment programs were required to detect the similarities between the more distantly related groups.

Although the Class 2C α -mannosidases showed a high degree of sequence similarity, the cellular locations and functions of the enzymes were variable. Although we were unable to conclusively determine the cellular location of the *A. nidulans* α -mannosidase, the protein was likely located in a subcellular organelle, since lysis of the cell and all subcellular components with TritonX-100 was necessary to release enzyme activity. The *A. nidulans* protein lacked a typical signal sequence or major transmembrane domain raising questions as to its method of localization. The rat and *S. cerevisiae* genes also did not contain typical signal sequences (Bischoff *et al.*, 1990; Yoshihisa and Anraku,

1990), though the *S. cerevisiae* protein was transported independently of the standard ER/Golgi sorting mechanism. As was reported for the *S. cerevisiae* vacuolar α -mannosidase (Yoshihisa and Anraku, 1989, 1990), the *A. nidulans* α -mannosidase was not essential for normal cellular function, as disruption of the gene did not have any visible effect on growth or morphology. The functions of the *S. cerevisiae* and *A. nidulans* α -mannosidases, however, were not determined. The rat ER/cytosolic enzyme, to which the *A. nidulans* gene showed a high degree of sequence similarity, is involved in oligosaccharide catabolism in both the ER and cytosol (Bischoff *et al.*, 1990; Grard *et al.*, 1996; Haeuw *et al.*, 1991; Moore and Spiro, 1994; Tulsiani and Touster, 1987), and it is possible that the *A. nidulans* α -mannosidase plays a similar role. In order to more fully understand the role of the *A. nidulans* α -mannosidase in protein glycosylation, more detailed analysis of the N-glycans produced by these transformants is required.

It is possible that there are yet other Class 2 α -mannosidases in *A. nidulans*. Since lower eukaryotes do not produce complex N-glycans, and mannosidase II is only necessary for complex N-glycan synthesis, it seems likely that this enzyme would not be found in *A. nidulans*. Since lysosomal α -mannosidase activity is necessary for oligosaccharide catabolism in higher eukaryotes, it would be reasonable to expect to find a similar activity in *A. nidulans*. The substrate p-nitrophenyl- α -D-mannopyranoside (PNP) has been used to assay Class 2 α -mannosidase activity. Surprisingly, when the *A. nidulans* Class 2C α -mannosidase was disrupted the PNP assay activity was reduced to zero. Based on examples of molecular sizes known for mannosidases in other species, the concentrated lysate was expected to also contain a lysosomal α -mannosidase (MW=115-121 kD). It is

possible that the activity of the lysosomal α -mannosidase was below the detection levels of our particular assay, but, alternatively, it is possible that *A. nidulans* contained a single Class 2 α -mannosidase. We are currently attempting to isolate a lysosomal α -mannosidase from *A. nidulans* by designing degenerate PCR primers from the known lysosomal α -mannosidase gene sequences.

Protein glycosylation is an important post-translational modification which can affect numerous properties of the protein and may be very important when using heterologous expression systems for the recombinant expression of proteins. Although the glycosylation pathways of higher eukaryotes are well characterized, other systems remain relatively undefined. We have cloned a Class 2 α -mannosidase from *A. nidulans* which appears to be involved in N-glycan catabolism. Induced expression of α -mannosidase resulted in the accumulation of two- to six-fold higher levels of the enzyme which did not affect the growth rate or morphology of the fungus, hence this system can provide a recombinant source of the *A. nidulans* α -mannosidase for purification. Biochemical characterization and cellular localization of this enzyme in *Aspergillus* will be necessary to determine its exact role in protein glycosylation. We are attempting to identify other genes in the glycosylation pathway in this organism, particularly α -mannosidases and mannosyltransferases, to further characterize these pathways in lower eukaryotes.

2.5 Materials and Methods

2.5.1 Strains, Media and Growth Conditions: The full length α -mannosidase gene was isolated from *Aspergillus nidulans* sporecolor mutant SM222. The *A. nidulans* expression host T580, a derivative of strain FGSC4 (Fungal Genetics Stock Center) is a uridine

auxotroph (ura^-) which was used for expression studies. Cultures were grown in CYM liquid media (10 g glucose; 2 g bactopectone; 1.5 g casamino acids; 1 g yeast extract; 10 ml 100X salt solution; 1 ml 1000X trace elements; 10 ml 100X vitamin solution; and 10 ml 100X adenine solution, per litre). Stock solutions (100X salt, 100X vitamin, 1000X trace elements, 100X adenine) were described in Kalsner *et al.* (1995). Non-transformed T580 strains were supplemented with 10 mM uridine. Protoplasts which integrated the *pyrG* selectable marker (pFB94) were selected for uridine prototrophy on minimal media (1 g fructose; 12 g threonine; 10 ml 100X salt solution; 1 ml 1000X trace elements; and 0.6 g $NaNO_3$, per litre) supplemented with 0.6 M sucrose as an osmoticum. Selected transformants for overexpression of the mannosidase were grown in liquid yeast-fructose-threonine (YFT) (5 g yeast extract; 2 g fructose; 12 g threonine; 10 ml 100X salt solution; 1 ml 1000X trace elements; and 0.6 g $NaNO_3$, per litre). Strains were maintained on CYM agar, and spore suspensions were obtained by washing cultured CYM agar plates with 8 ml 0.001% Tween 80. Mycelium for DNA isolations was obtained by inoculating 500 ml liquid CYM with 10^8 spores, and incubating 24 hours at $30^\circ C$ with constant agitation (200 rpm). Expression cultures were grown by inoculating 50 ml liquid YFT media with 10^8 spores, and incubating at $30^\circ C$ with constant agitation (200 rpm) for the duration of growth.

2.5.2 Oligonucleotide Primer Design: The forward primer (Table I) was designed by reverse translation of the protein sequence CHIDTAWLWPFXET, which was perfectly conserved between the rat and yeast sequences except at position X. The primer was made fully redundant at the three positions corresponding to amino acid position X. This

forward primer contained a short 5' tail with an embedded *HindIII* site preceding the coding sequence to enable directional cloning of the PCR product. The return primer was the complement of the reverse translation of the amino acid sequence FWLPDTFGYSS, which was perfectly conserved between the rat and yeast sequence except for the last serine (Figure 4). This primer was also tailed at the 5' end with a short sequence containing an embedded *EcoRI* site to facilitate the cloning.

2.5.3 DNA Isolation and PCR Amplification of α -Mannosidase: Total genomic DNA was extracted from finely ground freeze-dried mycelium. Approximately 400 mg of mycelium was vortexed with 2.5 ml of 50 mM EDTA; 0.2% SDS, centrifuged for 10 minutes and then 85 μ l of 3M KOAc; 5M acetic acid added to the supernatant. Following a 20 minute incubation on ice, the suspension was recentrifuged and DNA was isopropanol precipitated from the supernatant. After resuspension in 100 μ l TE (10 mM Tris pH 7.5; 1 mM EDTA), the DNA was extracted once with phenol, twice with chloroform/isoamyl alcohol (24:1) and ethanol precipitated. Each PCR reaction consisted of 10 - 100 ng of genomic DNA, 50 pmol of each primer, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 50 mM KCl, 0.01% gelatin, 0.1% Triton X-100, 200 μ M each of dATP, dCTP, dTTP, and dGTP and 2 units Taq DNA polymerase (Perkin-Elmer) in a final volume of 100 μ l. Amplification was performed in two stages using a RoboCycler (Stratagene) thermal cycler. Five cycles at a lower stringency (94^oC for 60 seconds, 55^oC for 120 seconds, 72^oC for 180 seconds) were followed by 25 cycles at higher stringency (94^oC for 60 seconds, 60^oC for 120 seconds, 72^oC for 180 seconds). The PCR product from the tailed amplification of the *A. nidulans* (SM222) nuclear DNA was directionally cloned (*EcoRI* /

HindIII) into pTZ18R prior to sequencing. The derived amino acid sequences of all three reading frames of the PCR-amplified DNA were compared to published α -mannosidase protein sequences to confirm that the PCR fragment represented a portion of the authentic α -mannosidase gene.

2.5.4 Library construction and screening: The α -mannosidase-specific PCR fragment was digoxigenin-labelled (DIG) by reduction of the dTTP concentration to 190 μ M and the inclusion of 10 μ M DIG-11-dUTP (Boehringer Mannheim) in a second PCR reaction using the cloned DNA as a target. The DIG-labelled DNA probe was used to identify the full length α -mannosidase gene by Southern hybridization. A partial library of *A. nidulans* (SM 222) sequences was constructed by digesting genomic DNA with *BamHI* and ligating the resulting fragments into the similarly digested lambda DNA vector EMBL-3.

Concatamers of the ligated DNA were packaged using the Gigapack II (Stratagene) *in vitro* packaging system. Approximately 10^5 recombinant lambda plaques were immobilised on nylon membranes (Genescreen Plus, Dupont) and hybridised with the DIG-labelled PCR fragment. A single strongly hybridising lambda clone was subcloned into the *BamHI* site of pTZ18R (*p* α -mann18) in preparation for sequencing.

2.5.5 Sequence Analysis: The DNA sequence was determined by the dideoxynucleotide method using the T7 sequencing kit (Pharmacia). Various restriction fragments of *p* α -mann18 were subcloned and individually sequenced using M13 universal primers.

Overlapping regions were sequenced with synthetic specific primers and in all cases, both DNA strands were sequenced. Genescanner sequence analysis software (Hastings Software Inc., Hastings-on-Hudson, NY) was used to identify open reading frames, generate

alignments and deduce polypeptides from the primary DNA sequence. FASTA searches of the GenPept and SwissProt databases (Genbank, Los Alamos) were made using the deduced polypeptides. The PSORT (Nakai and Kanehisa, 1992) program was used for prediction of protein localization sites in the deduced polypeptide. DNASTar sequence analysis software (DNASTar, Madison, WI) was used to compute amino acid similarities and to generate putative phylogenetic relationships. Intron positions were confirmed by reverse transcription (RT)-PCR of α -mannosidase cDNA. Total RNA was isolated from *A. nidulans* mycelium using a single-step guanidine isothiocyanate extraction (TRIzol reagent, Life Technologies). The RNA was treated with DNAase, extracted twice with chloroform containing 4% isoamyl alcohol, ethanol precipitated and reverse-transcribed to cDNA using oligo dT primers (Stratascript RT-PCR kit, Stratagene). This cDNA was subsequently amplified using primer pairs which flanked the putative splice sites and compared to fragments obtained by amplifying genomic DNA with the same primer pairs. Sequencing of the cDNA fragments was used to confirm the exact location and size of the predicted introns.

2.5.6 Construction of Disruption Cassettes: The disruption cassette p8 Δ KO (Figure 5a) was derived from a 2.2 Kb *EcoRI-Sal I* α -mannosidase fragment. This 2.2 Kb fragment was modified by the addition of diagnostic restriction enzyme sites, as well as stop codons (TGA) in each of the three reading frames. Unique sites were added by tailed PCR amplification. The 5' overhang of the forward primer M3SL-F (Table I) encoded unique *BstEII*, *BamHI*, and *AscI* restriction sites, and the three stop codons, while the remainder of the primer encoded authentic sequence which primed immediately downstream of an authentic *BstEII* site. The reverse primer M3SL-R contained an embedded *Sal I* restriction

site, and annealed just upstream of the authentic *Sal I* restriction site. The 1.1 Kb PCR reaction product was used to directly replace the 1.1 Kb *BstEII-Sal I* fragment of p8SR, using the *BstEII* and *Sal I* sites for directional ligation of the fragment into similarly digested p8SR vector. Restriction analysis of the plasmid and sequencing of the inserted DNA region confirmed that the restriction sites and stop codons were intact.

2.5.7 Construction and induction of inducible expression cassette: The inducible expression cassette (*palcAman*) was created by direct fusion of the full length α -mannosidase coding region with the inducible *alcA* promoter (Figure 5b) via *NcoI* insertion of the α -mannosidase gene into the expression cassette. Sequence analysis verified proper fusion of the promoter and coding region for expression of the gene. The *alcA* promoter is subject to carbon catabolite repression and transcription initiation is enhanced via the transcriptional activator AlcR in the presence of ethanol. In the presence of glucose, transcription from *alcA*-driven genes was repressed by interaction of the negative regulatory protein CreA with the *alcA* promoter (Hintz and Lagosky, 1993). The *alcA*-driven genes were regulated by controlling glucose levels during fermentation. Transformants were grown first in the presence of 1% glucose to repress the *alcA* promoter and prevent any possible toxic effects which may have arisen from overexpression of the gene. Selected transformants which had integrated the inducible expression cassette were then grown in glucose limiting conditions in the presence of an inducer, allowing for overexpression of the cytosolic α -mannosidase.

2.5.8 Protoplasting and Transformation: Protoplasts were prepared according to the method of Fincham (1989), using Sigma Lysing Enzyme for cell wall digestion.

Protoplasts of strain T580 (*ura⁻*) were cotransformed with 1 μ g of the selectable marker

pFB94, which converts transformed cells to uridine prototrophy, and with 1 μ g of the disruption or overexpression (p8 Δ KO or palcAman). Transformants were initially screened for integration of pFB94 by selection on minimal media and were then assayed for cotransformation by PCR analysis. Genomic DNA of putative transformants was prepared by the method of Cenis (1992), and screened using primer pairs diagnostic for the non-selectable cotransforming plasmids. Gene disruption was confirmed by Southern hybridization (Genescreen, Dupont).

2.5.9 Mannosidase Assays: Mycelium was grown in liquid YFT cultures, harvested on Miracloth (CalBiochem) by suction filtration, and washed once with maleate buffer (0.4 M MgCl₂, 0.05 M maleate, pH 5.8). Cell walls were digested with 10 mg/ml Novozyme 234 (Interspex) in maleate buffer with constant shaking (200 rpm) at 30°C. Protoplasts were purified by pre-screening through a 100 μ m mesh followed by centrifuging at 3000 RPM for 15 min at 4°C, washed twice with 10 ml ice cold STC (1.2 M sorbitol, 100 mM Tris-HCl pH7.4, 10mM CaCl₂), pelleted, and resuspended into a final volume of 1 ml STC. The protoplasts were lysed on ice by the addition of an equal volume of 10 mM Tris, pH 7.6; 1% TritonX-100 and the protoplast debris was pelleted by centrifugation (3000 RPM, 4°C). The supernatant proteins were size selected with an Amicon stirred ultrafiltration cell Model 8050, fitted with a 100 kD membrane. Samples of proteins >100 kD were concentrated to a final volume of 1.5 ml in 0.01 M sodium acetate pH 4.5. Mannosidase activity was quantified as described by Matta and Bahl (1972). Equal volumes of concentrate and 5 mM PNP (p-nitrophenyl- α -D-mannopyranoside) in 0.01 M sodium acetate, pH 4.0 were incubated at 37°C for 60 minutes. The reaction was terminated with

1 ml of 0.2 M sodium carbonate, pH 11.8 and the absorbance of the yellow chromogen measured at 420 nm. One unit of enzyme activity (U) was defined as that which releases 1 μ mol of p-nitrophenol per hour.

2.5.10 Sequence alignments: Matchbox sequence alignments were performed with 17 α -mannosidase sequences obtained from the GenBank database. Sequence similarities were identified using the web based Matchbox program

(<http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html>), utilizing the

BLOSUM62 similarity matrix, and default parameters. Sequence alignments using the

ClustalW algorithm were performed with DNASTar sequence analysis software (DNASTar, Madison, WI), utilizing the PAM250 similarity matrix and default alignment parameters.

This program was used to generate the most parsimonious phylogenetic tree for the

sequences. The Multiple Alignment Program (MAP), used to find localized regions of

similarity, using the web based server (<http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html>) with default parameters and the BLOSUM62 similarity matrix.

CHAPTER 3 - Characterization of Class I α -1,2-mannosidase gene family in

Aspergillus nidulans (Eades, C.J. and W.E. Hintz. 2000. *Gene* 255: 25-34)

3.1 Abstract

We describe the cloning and sequence characterization of three Class I α -1,2-mannosidase genes from the filamentous fungus *Aspergillus nidulans*. We used degenerate PCR primers to amplify a portion of the α -1,2-mannosidase IA gene and used the PCR fragment to isolate the 2495 nt genomic gene plus several hundred bases of flanking region. Putative introns were confirmed by RT-PCR. Coding regions of the genomic sequence were used to identify two additional members of the gene family by BLAST search of the *A. nidulans* EST sequencing database. Specific PCR primers were designed to amplify portions of these genes which were used to isolate the genomic sequences. The 1619 nt coding region of the α -1,2-mannosidase IB gene and the 1759 nt coding region of the α -1,2-mannosidase IC gene, plus flanking regions, were fully sequenced. All three genes appeared to encode type-II transmembrane proteins which are typical of Class I α -1,2-mannosidases. The deduced protein sequences were aligned with 11 published Class I α -1,2-mannosidases to determine sequence relationships. All three genes exhibited high similarity to other fungal α -1,2-mannosidases. The α -1,2-mannosidase IB exhibited very high similarity to the *Aspergillus satoii* and *Penicillium citrinum* α -1,2-mannosidases and likely represents an orthologue of these genes. Phylogenetic analysis suggests that the three *A. nidulans* Class I α -1,2-mannosidases arose from duplication events which

occurred after the divergence of fungi from animals and insects. This is the first report of the existence of multiple Class I mannosidases in a single fungal species.

3.2 Introduction

The filamentous fungus *Aspergillus nidulans* is often utilized as an expression host for the secretion of heterologous proteins. Many of these proteins contain complex N-glycans in their natural state. The N-linked protein glycosylation pathways have been fairly well characterized in mammalian systems (reviewed in Kornfeld and Kornfeld, 1985) and yeast expression systems, such as *Saccharomyces cerevisiae* (reviewed in Dean, 1999; Herscovics, 1999), but remain relatively poorly characterized in filamentous fungi. Since filamentous fungal expression systems are used to express a variety of genes from other organisms, it would be advantageous to regulate the type of N-glycans produced on heterologously expressed proteins in order to produce proteins containing N-glycans which are similar to the natural product (Hintz *et al.*, 1995; Kalsner *et al.*, 1995).

In all these expression systems, protein N-glycosylation involves the transfer of an oligosaccharide precursor ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) to newly synthesized polypeptides. In the endoplasmic reticulum (ER) and Golgi, α -glucosidases and α -mannosidase(s) modify the precursor into several intermediate forms. In higher eukaryotes (ie. mammals), α -1,2-mannosidases remove a total of four mannose residues, yielding $\text{Man}_5\text{GlcNAc}_2$ which is the precursor for complex, hybrid, and high-mannose N-glycans. In the yeast *S. cerevisiae*, however, an ER-specific Man_9 - α -1,2-mannosidase removes only a single mannose residue, producing $\text{Man}_8\text{GlcNAc}_2$ (Herscovics, 1999). Filamentous fungi produce N-glycan structures containing 5 mannose units ($\text{Man}_5\text{GlcNAc}_2$), suggesting further processing of

the $\text{Man}_9\text{GlcNAc}_2$ precursor (Chiba *et al.*, 1993; Maras *et al.*, 1997a). The mechanisms of synthesis of high mannose N-glycans in filamentous fungi seem to differ from yeast, and may be more similar to processes in higher eukaryotes.

The first committed step in the conversion of high mannose N-glycans to complex N-glycans is the transfer of GlcNAc to the $\text{Man}_5\text{GlcNAc}_2$ by GlcNAc transferase I (GnT-I). This enzyme generally requires $\text{Man}_5\text{GlcNAc}_2$ as a substrate. Fungi produce a variety of N-glycans from $\text{Man}_5\text{GlcNAc}_2$ to $\text{Man}_9\text{GlcNAc}_2$ (and variations thereupon), many of which are not suitable substrates for GlcNAc transferase I (Archer and Peberdy, 1997). Maras *et al.* (1997b) showed that *in vitro* removal of mannose with Class I α -mannosidase from N-glycans significantly improved the incorporation of GlcNAc into the N-glycans. Our early research has therefore focused on the initial processing of N-glycans by α -mannosidases in filamentous fungi. Manipulation of this region of the pathway may allow the development of a system for the *in vivo* expression of heterologous proteins containing complex N-glycans.

The α -mannosidases have been previously classified into two independently derived groups, Class I and Class II, based on biochemical properties, substrate specificity, inhibitor profiles, and sequence alignments (Daniel *et al.*, 1994; Eades *et al.*, 1998; Moremen *et al.*, 1994). The first group contains the α -1,2-mannosidases which are found in the ER and Golgi, including the ER Man_9 -mannosidase, endomannosidase and Golgi mannosidase IA/IB. The second group of α -mannosidases is more heterogeneous and contains the lysosomal mannosidases, the Golgi mannosidase II and a distantly related group of enzymes, including the rat ER/cytosolic mannosidase (Bischoff *et al.*, 1990),

yeast vacuolar mannosidase (Yoshihisa and Anraku, 1989), and the *A. nidulans* Class II mannosidase (Eades *et al.*, 1998).

Several α -mannosidases have already been characterized in filamentous fungi. The genes encoding Class I α -mannosidases have been isolated from *Aspergillus satoii* (Inoue *et al.*, 1995) and *Penicillium citrinum* (Yoshida and Ichishima, 1995) and a Class II α -mannosidase has been cloned from *A. nidulans* (Eades *et al.*, 1998). An ER resident α -mannosidase has been identified and cloned in *S. cerevisiae* (Camirand *et al.*, 1991) which removes a single mannose unit from the Man₉GlcNAc₂ oligosaccharide in the ER. Here we report the identification and sequence analysis of three novel Class I α -mannosidases from *A. nidulans*. Phylogenetic analyses of the deduced protein sequences suggest a recent duplication and divergence of these mannosidases.

3.3 Materials and Methods

3.3.1 Strains, Media and Growth Conditions: *Aspergillus nidulans* sporecolor mutant SM222 was grown in CYM liquid media, described by Eades *et al.* (1998). Strains were maintained on CYM agar, and spore suspensions were obtained by washing cultured CYM agar plates with 8 ml 0.001% Tween 80. Mycelium for protoplasting and DNA isolations was obtained by inoculating 500 ml liquid CYM with 10⁸ spores, and incubating 24 hours at 30°C with constant agitation (200 rpm).

3.3.2 Degenerate Oligonucleotide Primer Design: The forward primer was designed by reverse translation of the protein sequence GGLGESFYEY, and the reverse primer from the complement of the reverse translation of the sequence FXLAETLKLYL. These protein sequences were conserved between the α -1,2-mannosidase protein sequences of *S.*

cerevisiae (Camirand *et al.*, 1991), mouse IA (Lal *et al.*, 1994), human IA (Bause *et al.*, 1993), and rabbit liver (Lal *et al.*, 1994). A codon usage table compiled for *A. nidulans* was used to aid in nucleotide selection at degenerate sites.

3.3.3 DNA Isolation and PCR amplification: Total genomic DNA was extracted from finely ground freeze-dried mycelium of *A. nidulans* strain SM222 using the method described by Eades *et al.* (1998). Each PCR reaction consisted of 10 - 100 ng of genomic DNA, 50 pmol of each primer, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 50 mM KCl, 0.01% gelatin, 0.1% Triton X-100, 200 μM each of dATP, dCTP, dTTP, and dGTP and 2 units Taq DNA polymerase (Perkin-Elmer) in a final volume of 100μl. Amplification was performed in two stages using a Perkin-Elmer thermal cycler. Five cycles at a lower stringency (56°C) were followed by 30 cycles at higher stringency (63°C). The PCR products were eluted from 1% low-melting point agarose, cloned into T-Vector (Promega) using the T-Vector cloning system and sequenced with Universal Forward and M13 Reverse primers.

3.3.4 Library Construction and Screening: The cloned PCR products were used as a template to produce a radiolabelled PCR probe for library screening. The PCR reactions were performed as previously described, except that only 100 μM dCTP was used and 0.825 μM α-³²P-dCTP (5 μCi) was added to the reactions. The radiolabelled PCR products were purified with the Wizard PCR Prep Kit (Promega) and scintillation counted to assess radioactivity. A genomic library of *A. nidulans* (SM222) sequences was constructed by partially digesting genomic DNA with *Mbo*I and ligating the resulting fragments into lambda DNA vector EMBL-3 digested with *Bam*HI. Concatamers of the ligated DNA were packaged using the Gigapack II (Stratagene) *in vitro* packaging system.

Approximately 10^5 recombinant lambda plaques were immobilised on nylon membranes (Genescreen Plus, Dupont) and hybridized with the radiolabelled PCR probes. Single hybridizing clones were selected and rescreened. Lambda DNA was isolated with the Wizard Lambda DNA Extraction Kit (Promega), digested with restriction enzymes and subcloned into pUC18 cloning vector.

3.3.5 DNA Preparation and Sequencing of A. nidulans α -1,2-mannosidase genes:

Restriction fragments of the α -1,2-mannosidase IA containing the full length gene were cloned into pUC18 cloning vector. Initial manual sequencing of subclones was performed using the dideoxynucleotide method using a T7 sequencing kit (Pharmacia). Initial sequence data was obtained with the universal priming sites, and with specific sequence primers (primer walking). Final sequence data was provided by subcloning the fragments using various restriction enzymes and sequencing with an ABI373 automated fluorescent sequencer (Applied Biosystems, Foster City, CA). The α -1,2-mannosidase IB gene and the α -1,2-mannosidase IC were sequenced by cloning restriction fragments of the positive lambda subclones into BlueScript II cloning vector (Stratagene, La Jolla CA) and sequencing on an ABI373 sequencer, using the universal priming sites of the vector.

3.3.6 Identification of Introns:

Total RNA was extracted from fresh mycelium grown in 150 ml liquid media (~ 1 gram) using the guanidine isothiocyanate method (TRIzol, GibcoBRL). The RNA pellet was resuspended in 30 μ l H₂O/DEPC and 5 μ l was used for cDNA preparation and PCR amplification. Reverse transcriptase PCR was performed using the SuperscriptII reverse transcriptase kit (GibcoBRL) for first strand cDNA synthesis (oligo-dT primer), followed by PCR amplification of the cDNA. The cDNA was used to amplify regions containing putative introns for comparison with genomic DNA

amplification. RT-PCR products were eluted from agarose gels and cloned into T-Vector (Promega) and sequenced using universal primers.

3.3.7 Sequence Comparison: The Class I α -mannosidase protein sequences used in the sequence comparisons were obtained from GenBank as follows: *S. cerevisiae* (Camirand *et al.*, 1991 - Accession M63598), *A. satoi* (Inoue *et al.*, 1995 - D49827), *P. citrinum* (Yoshida and Ichishima, 1995 - D45839), *Drosophila melanogaster* (Kerscher *et al.*, 1995 - X82640), *Spodoptera frugiperda* (Kawar *et al.*, 1997 - AF005035), *Homo sapiens* IA (Bause *et al.*, 1993 - X74837), *Homo sapiens* IB (Tremblay *et al.*, 1998 - AF027156), *Mus musculus* IA (Lal *et al.*, 1994 - U04299), *Mus musculus* IB (Herscovics *et al.*, 1994 - U03457), *Sus scrofa* (Bieberich *et al.*, 1997 - Y12503). The outgroup was *A. nidulans* Class IIC α -mannosidase (Eades *et al.*, 1998 - AF016850). These sequences were aligned using the Multiple Alignment Program algorithm (Huang, 1994).

3.3.8 Phylogenetic Analyses Multiple alignments were converted into PAUP format and phylogenetic trees were generated on PAUP 4.0b* computer software (Swofford, 1998). After initial execution of the alignment file, trees were generated using maximum parsimony with a full heuristic search (100 replicates of random stepwise branch addition). The most parsimonious tree was rooted using the *A. nidulans* Class IIC α -mannosidase protein sequence as an outgroup. Confidence values for the tree branches were determined using the bootstrap option (100 bootstrap replicates using a full heuristic search).

3.3.9 Mannosidase Assays: Crude protein extracts were obtained from protoplasts and from culture filtrates. Secreted proteins were precipitated from 1.5 ml of culture filtrate by saturation with ammonium sulfate. After centrifugation, the protein extract was resuspended in 500 ml 0.01 M phosphate buffer (pH 6.0) and precipitated again using

ammonium sulfate (to ensure removal of sugar residues which would interfere with mannosidase assays). Proteins were resuspended in 100 μ l 0.01 M phosphate buffer (pH6.0) and 27 μ l of the extract was used in the mannosidase assay. Intracellular proteins were extracted from protoplasts, which were prepared as in Eades *et al.* (1998). Protoplasts were centrifuged and resuspended in 200 μ l 0.01 M phosphate buffer (pH6.0)/1% octylthioglucoside, and resuspensions were vortexed vigorously to ensure complete protoplast lysis. Cellular debris was removed by centrifugation, and 27 μ l of the lysate was used in mannosidase assays.

Mannosidase assays were performed using the disaccharide Man- α -1,2-Man- α -OCH₃ as a substrate in a coupled enzyme assay as described earlier (Scaman *et al.*, 1996), with some modifications. Digestion of the substrate was performed in a 30 μ l final volume containing 27 μ l of crude extract in 0.01M phosphate buffer (pH6.0) and 3 μ l 100 mM disaccharide Man- α -1,2-Man- α -OCH₃ incubated at 37°C for 3 hours. Detection of released mannose was achieved by addition of 30 μ l Tris-HCl (pH7.6) and 240 μ l of developing solution, containing glucose oxidase (55 U/ml), horseradish peroxidase (1 U/ml) and o-dianisidine dihydrochloride (70 μ g/ml), incubated at 37°C for 3 hours. Absorbance measurements at 450 nm determined final color change. Standard blanks included all components of the colorimetric reaction, plus the substrate. As a control, enzyme extracts which were not used in the mannosidase digestion were subjected to the colorimetric reaction, to determine the absorbance which is due to the extract itself, and not due to mannose release. These values were subtracted from the absorbance values of the assays. Free mannose was used as a standard. All assays were performed in triplicate

and the mean and standard deviation was calculated for each sample. Mannosidase activity was standardized by comparison with total protein in the crude enzyme extracts, and was defined as the amount of mannose released from substrate per μg of total protein per hour. Protein concentrations were determined by the Bradford method (Bradford, 1976) using BSA as a standard.

3.4 Results

3.4.1 Identification of three Class I α -1,2-mannosidases from *Aspergillus nidulans*

To identify Class I α -1,2-mannosidase genes in *A. nidulans*, degenerate PCR primers were designed from conserved blocks found in multiply aligned Class I α -mannosidase genes. Amplification of *A. nidulans* genomic DNA with degenerate oligonucleotide primers yielded a number of products, which were ligated into cloning vector and sequenced. The sequence of a 900 bp clone was used to search the GenBank database for sequence homology (BLAST search). High scoring matches were found with other Class I α -1,2-mannosidases, confirming that a portion of the Class I α -1,2-mannosidase had been amplified from *A. nidulans*. An α -1,2-mannosidase specific radiolabelled probe derived from this 900 bp clone was used to screen an EMBL-3 library of genomic sequences, and a single lambda clone was recovered which contained the entire α -1,2-mannosidase gene. Two non-overlapping *Bam*HI subclones were sequenced by manual and automated sequencing, resulting in 2- to 5- fold redundancy sequencing of the full length gene (α -1,2-mannosidase IA) and several thousand bp of flanking sequence (GenBank Accession # AF129497). The possibility of missing sequence at the junction of

the non-overlapping *Bam*HI subclones was eliminated by completely sequencing the original 900 bp PCR fragment, which spanned the *Bam*HI junction.

The deduced coding region of the *A. nidulans* α -1,2-mannosidase IA gene was used to search the *A. nidulans* EST Sequencing Project Database using the BLAST algorithm to determine if there were multiple α -1,2-mannosidase genes expressed in this organism. Two different short cDNA fragments were identified which showed significant similarity, but not 100% identity, to the α -1,2-mannosidase IA gene, and thus represented separate genes. A second set of PCR primers specific for the novel genes were designed according to the cDNA fragments. The amplification products from these primers (~200 bp) were radiolabelled and used to probe the *A. nidulans* genomic library to select the full length genes. A clone containing the full length α -1,2-mannosidase IB gene was recovered as a single 5.6 kb *Bam*HI clone. The gene and several hundred bp of flanking sequence were sequenced with 2- to 4-fold redundancy at each base position (GenBank Accession #: AF129496). A second lambda clone containing the full length α -1,2-mannosidase IC gene was recovered and two non-overlapping *Bam*HI subclones (4 kb and 6 kb) were isolated, which together contained the gene and flanking regions. Again, the sequence across the *Bam*HI subcloning junction was verified to eliminate the possibility of missing sequence. The gene and several hundred bp of flanking region were fully sequenced (Accession #: AF233287).

3.4.2 Sequence characterization of the α -1,2-mannosidase IA, IB, and IC genes

The DNA sequences of the three Class I α -1,2-mannosidases were analyzed to determine the amino acid coding sequence, including determination of the correct reading

frame, identification of potential intron sequences, and identification of the correct translational start codon for each gene. A BLAST search of the α -1,2-mannosidase IA gene revealed two open reading frames (ORFs) separated by a region of DNA containing several stop codons which could indicate the presence of an intron sequence. To verify the presence of an intron at this site, PCR products spanning the putative intron were amplified from reverse-transcribed RNA (RT-PCR), cloned into vector and sequenced. Comparison of the α -mannosidase IA sequence with RT-PCR sequence verified the presence of a 50 bp intron at the expected splice junction. The intron contained a 5'-splice site (5'-GTAAGT-3') which matched the consensus sequence for filamentous fungi (5'-GTANGT-3'), and a 3'-splice site (5'-TAG-3') which matched the consensus 5'-YAG-3' (Ballance, 1986, 1991; Gurr *et al.*, 1987). The intron also contained an internal lariat sequence (5'-GCTGAC-3'), located 15 bp upstream of the 3'-splice site, consistent with the consensus 5'-(G/A)CT(G/A)AC-3' for fungal introns.

The deduced amino acid sequence of the *A. nidulans* α -1,2-mannosidase IA gene was aligned with other published α -mannosidases to determine whether there might have been other introns which did not shift the reading frame and did not contain stop codons but did increase the size of the putative gene product. Two additional regions of the first ORF did not align with other published sequences and appeared as large 'gaps' in the multiple sequence alignment for all of the other α -1,2-mannosidase genes used in the alignment. To determine whether these sequences represented introns or encoded polypeptide sequence novel to the *A. nidulans* α -mannosidase, RT-PCR products were compared to genomic PCR products. The PCR amplification products derived from both RNA and genomic DNA, using primer pairs which flanked these regions, were the same

size, indicating that there were no introns present in these regions. Although introns can sometimes be present with less conserved splice sequences (Gurr *et al.*, 1987), these regions did not contain consensus splice motifs.

The α -1,2-mannosidase IB gene contained three open reading frames separated by two regions which contained several stop codons and caused a shift in the reading frame. These regions were searched for consensus intron splice sites. Both of the putative introns contained consensus 5'- and 3'- splice sites and a consensus internal lariat sequence. The two putative introns were verified by RT-PCR. In both cases, the RNA-derived amplification product was smaller than the DNA-derived amplification product, and the size difference corresponded to the predicted size of each intron. Sequencing of the RT-PCR products verified the presence of a 51 bp and 53 bp intron at the respective splice sites. While the position of each intron in the coding region correlated directly to two of the introns found in the *P. citrinum* α -1,2-mannosidase gene, the intron sequences themselves were not conserved. The α -1,2-mannosidase IC gene contained a single contiguous open reading frame and did not contain any consensus intron sequences. This gene did not appear to contain any intron sequences.

The authentic start codon (ATG) of the α -1,2-mannosidase IA gene was inferred by sequence context of the putative start codons, combined with protein sequence alignments with known α -mannosidase proteins. A potential start codon (ATG) occurred in frame 42 bp after the stop codon which defined the 5' end of the first ORF, while the next in frame ATG codon was 423 bp downstream. This first codon was thus a better candidate for the translational start codon. Translation originating at this start codon would produce a protein product with an N-terminus which was larger than other fungal

α -mannosidases, but similar in size to the N-termini of mammalian and insect Class I α -mannosidases. Although there is not a strong consensus sequence surrounding the translational start codon in filamentous fungi (Ballance, 1986), there is a preference for a purine at position -3 (97%). The putative start codon for this gene had a G at position -3, and thus conforms to this rule. Placement of the translational start codon can also be inferred from sequence context with respect to promoter elements and the transcription start site. There was a TATA-like element at position -47 of the putative start site and several pyrimidine rich blocks within 100 bp upstream of the proposed start codon. These pyrimidine rich regions are often found in fungal promoters and may influence the level of transcription (Ballance, 1986). The 5' non-translated region did not contain a CAAT-box upstream of the TATA-like element, but this is not unusual for fungal promoters.

The first potential translational start codon in the α -1,2-mannosidase IB gene occurred 42 bp into the first ORF. Comparison of the coding region of this ORF with the *A. satoi* and *P. citrinum* α -1,2-mannosidase genes showed that the position of the putative translational start site correlated with the start sites of these genes. This start codon also contained a purine at the -3 position, a TATA-like element at position -76, and several CT-rich blocks in the sense strand. The first potential translational start codon of the α -mannosidase IC gene occurred 38 bp into the ORF. The start site also contains a purine at position -3 and CCAAT motif at -221, but did not contain a clearly definable TATA box.

The putative coding region of the α -1,2-mannosidase IA gene encodes a 816 amino acid protein with a predicted molecular weight (MW) of 91 kD. This is somewhat larger than other Class I α -mannosidases, which range in size from 53 kD for the *P. citrinum* α -mannosidase I (Yoshida and Ichishima, 1995) to 73 kD for the *H. sapiens* α -

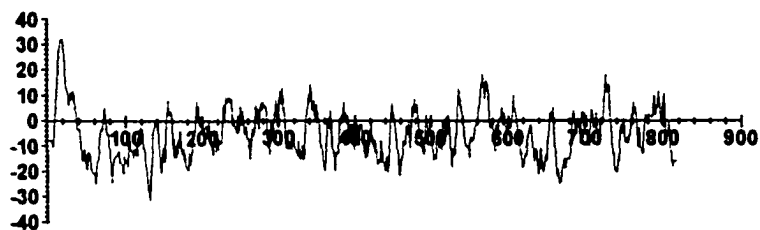
mannosidase IB (Bause *et al.*, 1993), *M. musculus* α -mannosidase IA (Lal *et al.*, 1994) and α -mannosidase IB (Herscovics *et al.*, 1994), and *S. scrofa* α -mannosidase I (Bieberich *et al.*, 1997). The coding region of the α -1,2-mannosidase IB gene encodes a 505 aa protein with a predicted MW of 56 kD, while the α -1,2-mannosidase IC gene encodes a 586 aa protein with a predicted MW of 65 kD. Both of these predicted sizes are within the range of currently identified Class I α -mannosidases.

The putative α -1,2-mannosidase IA, IB, and IC proteins contained several charged N-terminal amino acids representative of a typical signal sequence motif downstream of the putative start site. Kyte-Doolittle hydrophathy plots showed that the signal sequences were followed by highly hydrophobic regions approximately 15-16 amino acids in length, while the remainder of the C-termini were relatively hydrophilic (Figure 10). These proteins likely form type II transmembrane proteins, which is a characteristic of other Class I α -1,2-mannosidases.

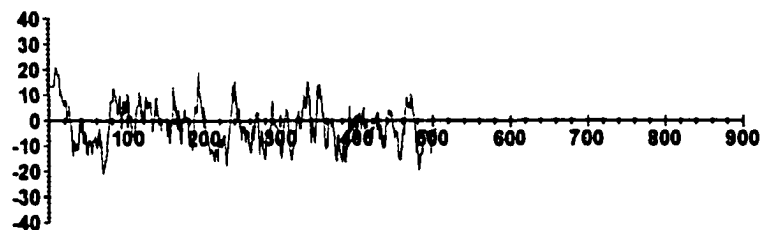
3.4.3 Multiple Sequence Alignment

The derived *A. nidulans* α -1,2-mannosidase amino acid sequences were compared with each other and with α -1,2-mannosidase protein sequences from other species to assess the relationship of these genes. Figure 11 shows the multiple sequence alignment of the three α -1,2-mannosidase protein sequences from *A. nidulans*. The N-termini are not well conserved, which is not unexpected, since the N-termini of Class I α -mannosidases encode the transmembrane domain and the stem region of the protein, neither of which contain catalytic activity. Several well conserved regions were found further downstream

IA



IB



IC



FIGURE 10: Kyte-Doolittle hydropathy plot of (A) *A. nidulans* α -mannosidase IA, (B) *A. nidulans* α -mannosidase IB and (C) *A. nidulans* α -1,2-mannosidase IC predicted amino acids sequences. Vertical axis shows the hydrophobicity of a given region of the protein (Kyte and Doolittle, 1982) with positive values representing hydrophobic regions of the protein, and negative values representing hydrophilic regions of the protein.

<i>A. nidulans</i>	IA	MFRARRSRIS	LVFAVIEVLL	IFHFSRLAVT	ISLQSWVPPP	PVDHHPNPPF	DQNLKDPYEN	DNSATGSGAP	70
<i>A. nidulans</i>	IB	-----	-----	-----	-----	-----	-----	-----	0
<i>A. nidulans</i>	IC	MPPRWSSSLIS	IT--AIFLVL	FELLHR----	-----	-----	-----	-----	24
<i>A. nidulans</i>	IA	PPALVEPEEY	QRPPLYTDS	DSPTPSKERL	DTPSN----	---VPSQEPE	FDAARLQTGA	QTQNKHEDDE	132
<i>A. nidulans</i>	IB	-----	-----	-----	-----	-----	-----	-----	0
<i>A. nidulans</i>	IC	-----	-----NT	DTPRAANRAT	NGPANGFARQ	QSICPSTPPQ	PPYNRSTGG	-----	66
<i>A. nidulans</i>	IA	DIVPIHWKP	MPERHFVSP	ALIKLPTGQS	KELPQLQ-AK	FKDESSDKM	QRLQQLDTIK	SAFLHAMNGY	201
<i>A. nidulans</i>	IB	-----	-----MR	TLALAAAFAG	FAAARVPAYA	I---TRPVM	RSDSRADAVK	EAFSHAWDGY	48
<i>A. nidulans</i>	IC	-----FNWGE	IPVRYPVSD-	-FIPLSTNSP	ATLPRICRSS	FPLQSSITKS	RQA----AVK	GAFQRAWTSY	125
<i>A. nidulans</i>	IA	KISAMGHDEV	FPLRGGFKDT	ENWGATLVD	ALDTLWIMDL	KEEFSMAVDY	VKKIDFT---	--TSTKKEIP	266
<i>A. nidulans</i>	IB	YNYAFPHDEL	HPISNGYGDS	RNHWGASAVD	ALSTAIMMRN	ATIVNQILDH	IAAVDYS---	---KTNAMVS	112
<i>A. nidulans</i>	IC	TTHAWKADEV	RPITAGSRNN	FGWGATLVD	NLDTLLIMGL	DEEFAAAVDA	LADIEFSPHS	SPSSSQSTIN	195
<i>A. nidulans</i>	IA	VFETIRYLG	GMLGAYDI--	-----SGH	KYDILLEKSV	ELADVLMDAF	DTPNRMFTLY	YKWSPEYASE	327
<i>A. nidulans</i>	IB	LFETIRYLA	GMISGYDLK	GPAAGLVDDS	RVDVLLQSQ	NLAEVLFKFAF	DTPSGVVPM	INITSGGND-	161
<i>A. nidulans</i>	IC	IFETIRYLG	GLLAAYDL--	-----TGC	RETFLLDKAI	QLGEMITYSF	DTENRMPVPR	WNLHKAGNGE	256
<i>A. nidulans</i>	IA	FRRGDFKAVL	AELGSLSEF	TRLAQLTKQD	KYDAIARIT	NEL--EKYQD	LTKLPLGLWPL	NLDASGCRRV	395
<i>A. nidulans</i>	IB	---GATTNGL	AVTGTLVLEW	TRLSDLTGND	EYARLSQRAE	DYLLHPEPAQ	YEPFPLIGS	AVNIADGKLA	248
<i>A. nidulans</i>	IC	PQRAAVQGVV	AELASSSEF	TRLSQLTGDM	RYFDAASRIT	DLL--DSQAG	HTRIPGLWV	SVNLQ-----	319
<i>A. nidulans</i>	IA	PGVSREPAAA	GQPVWSSDE	INSTSSVSFR	TRQIHEGGEF	VRHDNDSFET	GFPVSVDTRT	PPPQKQDCTGG	465
<i>A. nidulans</i>	IB	NG-----	-----	-----	-----	-----	-----	-----	250
<i>A. nidulans</i>	IC	-----	-----	-----	-----	-----	-----	---KGDLTRG	326
<i>A. nidulans</i>	IA	LNDQLSGIDK	FGLGALGDS	YEYLPKEYML	LGG--NNDQY	LNMYK----	AMDTVREYLV	YQPMKNNRD	529
<i>A. nidulans</i>	IB	-----HI	SWNGG-ADSY	YEYLIKMYVY	-----DPER	FGLYRDRWA	MAESSINHLA	SHPSTRP--D	303
<i>A. nidulans</i>	IC	-----ST	FSEGGADSA	YEYLGKTYRL	LGGVVGKGPQY	ERLARN----	ALDAGIRHL	FRPMPDHD	384
<i>A. nidulans</i>	IA	VRFLATVSMT	KSLDANPPGR	TFAYEGTML	TCFAGGLAI	GAKLFGLDKD	LKLGSLTDC	CVWAYEATKS	599
<i>A. nidulans</i>	IB	VTFLATYNEE	HQLGLT----	-----SQML	TCFDGGSFL	GGTLLDRQDF	VDFGLDLVAG	CHETYNSTLT	363
<i>A. nidulans</i>	IC	I-LLPGVAHA	TSSSVGLEPR	TE-----HL	ACFVGGMYAL	AGKLFNSQTY	LDTGRKLTDC	CINWYDNSPL	447
<i>A. nidulans</i>	IA	GIMPEAPQLV	PCKKGEPCW	DED---AYYM	AMDPYADKRP	ISHNKRAGP	EKGNWHVVAT	AESSSPQEDK	666
<i>A. nidulans</i>	IB	GIGPEQPSWD	PNGVPDSQKE	LFERAGFY--	-----	-----	-----	-----	391
<i>A. nidulans</i>	IC	GIMPEMFTVP	ACPSVAECPW	DETRGGIYTY	VRDGH----	-----	-----	-----	482
<i>A. nidulans</i>	IA	TQKSTTTEGR	HTGTTTGAGA	LSHEEFVTGK	ILNDRLPPGM	TGISARQYLL	RPEAIESVFI	MFLTGDPSPW	736
<i>A. nidulans</i>	IB	-----	-----	-----	-----	---INSGQYIL	RPEVIESFYI	AWRVTDGTY	419
<i>A. nidulans</i>	IC	-----	-----	-----	-----	-----	YEL RPEAMESIFY	MWRITGDEKY	505
<i>A. nidulans</i>	IA	REKGMKMPQA	VDKATKTELA	NSAISDVTVD	--NRPVDSM	ESFNLAEFLK	YFYLLFSDPS	LVSLEE----	800
<i>A. nidulans</i>	IB	LEWVWNAFTN	INKYCRATG	FAGLENVNAA	NGGGR-IDNQ	ESFMFAEVLK	YSFLTPAPED	DWQVQKSGSN	488
<i>A. nidulans</i>	IC	REAAWRMFTA	IEAVTKTEFG	NAAVRDVMVE	EGNVKREDSM	ESFMAEFLK	YLFLIPGETD	LVSLLD----	571
<i>A. nidulans</i>	IA	-YVLTREANP	EKRPRY-	-----	-----	-----	-----	-----	815
<i>A. nidulans</i>	IB	TFVYNTREANP	EKVYTPQ	-----	-----	-----	-----	-----	505
<i>A. nidulans</i>	IC	-WVYNTREANP	LRGAGS-	-----	-----	-----	-----	-----	566

FIGURE 11: Multiple Sequence Alignment of three α -1,2-mannosidase putative protein sequences from *A. nidulans*. Amino acids which are identical between the three sequences are in bold. Degenerate primer sites used in the first PCR amplification are shown with horizontal arrows. Vertical arrow shows the amino acid believed to be involved in dMM recognition, while asterisks show cysteine residues, conserved in all Class I α -mannosidases, which are likely necessary for enzyme activity.

of the non-conserved N-termini. The α -1,2-mannosidase IA protein contains several large unique spans which are not found in the α -1,2-mannosidase IB and α -1,2-mannosidase IC genes. The function and evolutionary significance of these 'extra' sequences is at present unknown.

A total of 14 Class I α -1,2-mannosidase protein sequences were compiled and a multiple alignment was performed. Again, the N-termini of these proteins were not well conserved in length or sequence composition, except that all encoded a type II transmembrane region of the protein. Several well conserved blocks were found throughout the remaining C-terminal region of the proteins. Several amino acids had been previously identified which were particularly important for the catalytic activity of the α -1,2-mannosidases. The disulfide bond formed between Cys³⁴⁰ and Cys³⁸⁵ of the *S. cerevisiae* was necessary for enzyme activity (Herscovics, 1999; Lipari and Herscovics, 1996). These two cysteine residues were conserved in all of the α -1,2-mannosidases including the three *A. nidulans* α -1,2-mannosidases (Figure 11). The amino acid which interacts with the inhibitor deoxymannojirimycin (dMM) has been identified as glutamine (Glu) in all species except *P. citrinum*, which has aspartate (Asp) at this position (Yoshida and Ichishima, 1995). A Glu residue was found at this position in the three *A. nidulans* α -1,2-mannosidases and there was further sequence conservation surrounding this residue, indicating that all three of these enzymes would likely be inhibited by dMM (Figure 11).

3.4.4 Phylogenetic Analysis

A phylogenetic analysis was performed from the multiple sequence alignment to characterize the relationships of the various genes in this gene family. Trees were

generated using maximum parsimony analysis tool in the PAUP software package and the most parsimonious tree was selected for further analysis (Figure 12). The sequences in the phylogram appeared as three major groups: mammalian, insect, and fungal. The exception to this was the *S. cerevisiae* α -1,2-mannosidase which showed low overall similarity to any of the other α -1,2-mannosidases. Members within these groups were highly related, and shared low, yet still significant, similarity to members of other groups. The mouse and human species each contained two α -1,2-mannosidase genes. These genes were highly related and likely resulted from a recent duplication event which occurred after the mammalian lineage diverged from insects and from lower eukaryotes. The *A. nidulans* genes also appear to have resulted from recent duplication events which occurred after the divergence of the various lineages, since these genes were more related to each other than to mannosidases in the other lineages. Bootstrap analysis was performed to assess the confidence values for each node of the tree (Figure 12). All branches of the tree are supported with high confidence, with bootstrap values ranging from 70 –100 (of 100) replications.

3.4.5 Determination of α -1,2-mannosidase activity in *A. nidulans*

In order to verify that α -1,2-mannosidases are actively expressed intracellularly, assays were performed to determine the secreted and/or intracellular α -1,2-mannosidase activity levels. To determine secreted α -1,2-mannosidase activity, crude protein extracts were obtained from liquid culture filtrates. Intracellular proteins were extracted from fungal protoplasts using a detergent buffer. Significant α -1,2-mannosidase activity was found in the intracellular protein extract (28.16 nmol mannose released from substrate/ μ g

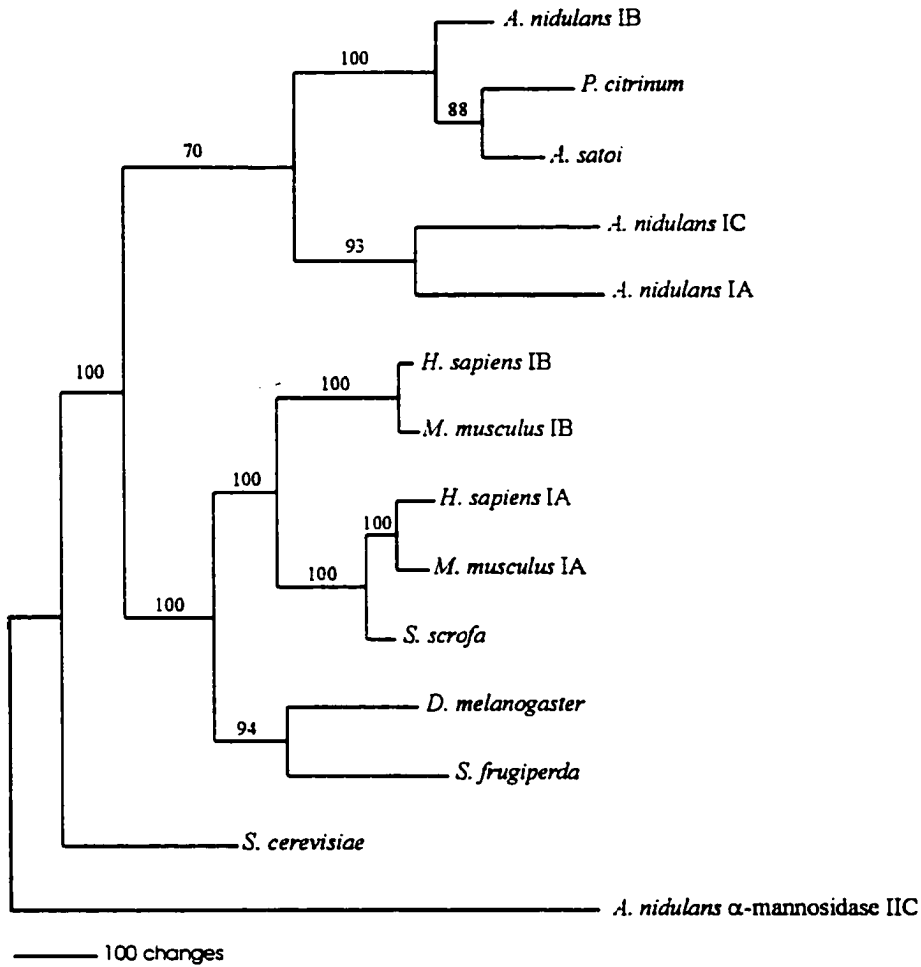


FIGURE 12: Phylogram showing sequence relationships of Class I α -mannosidases. Phylogram was generated from the multiple sequence alignment using PAUP 4.0b* analysis program. Numbers above the branches of the tree represent bootstrap confidence values, out of 100 bootstrap replications. Branch lengths represent the number of amino acid substitutions.

total protein/h; SD = 4.45), whereas very little activity was found in the extracellular extracts (0.77 nmol/ μ g/h ; SD = 0.74). This is consistent with the hypothesis that the Class I α -1,2-mannosidase genes from *A. nidulans* encode type-II transmembrane proteins which are expressed intracellularly.

3.5 Discussion

Protein N-glycosylation is an ordered process which occurs in all eukaryotic species. The early steps of this process, which include the transfer of the oligosaccharide to the nascent polypeptide and the removal of glucose residues, are almost identical in all eukaryotes (Kornfeld and Kornfeld, 1985). At this point, the pathways in various organismal lineages can become quite divergent resulting quite different N-glycans as the end products of the pathway (Herscovics, 1999). It appears that multiple α -mannosidases are involved in N-glycan processing in mammals and insects while a single α -mannosidase is found in the yeast *S. cerevisiae*. During the early stages of N-glycan processing, mannose removal in filamentous fungi appears to be more similar to higher eukaryotes such as mammals and insects than to yeast (Maras *et al.*, 1997a). Filamentous fungi, like higher eukaryotes, may utilize multiple α -mannosidases with defined specificities for this purpose.

We have identified three Class I α -1,2-mannosidase genes in a single fungal species, *Aspergillus nidulans*. All of these genes encoded gene products with similarity to other Class I genes, but were significantly different from each other. The α -mannosidase IB gene was highly similar at the protein level to the orthologous genes of *A. satoi* and *P. citrinum* genes and had two of three introns in identical positions as the *P. citrinum* gene.

The *A. satoi* gene sequence was derived from cDNA, thus the intron positions for this gene were not known (Inoue *et al.*, 1995). The protein products from these genes presumably share similarity in structure and substrate specificity. The α -mannosidase IA gene exhibited lower overall similarity to these three genes, which is indicative of a more distant relationship. The *A. nidulans* α -mannosidase IA gene is therefore considered to be a paralogue of the α -mannosidase IB gene. The *S. cerevisiae* Class I α -mannosidase was no more similar to the fungal mannosidase than it was to the insect or mammalian α -mannosidases, and it is difficult to draw any conclusions about its evolutionary derivation.

To date, only a single member of this gene family has been found in other fungal species, including the yeast *S. cerevisiae* (Camirand *et al.*, 1991), and the filamentous fungi *P. citrinum* (Yoshida and Ichishima, 1995) and *A. satoi* (Inoue *et al.*, 1995). This could be due to a lack of a concerted effort to search for these genes in other species. Only one Class I α -1,2-mannosidase gene has been sequenced in yeast which removes only a single mannose residue during N-glycan processing. This appears to be the only functional Class I α -mannosidase in yeast (reviewed in Herscovics, 1999). On the other hand, filamentous fungi remove several mannose residues during N-glycan processing (Maras *et al.*, 1997a), so it might be expected that multiple enzymes would exist to achieve this. This is the first definitive proof that this is indeed the case. It is unclear why multiple enzymes would exist to remove mannose residues from the N-glycan precursor. In mammalian systems, two Class I α -1,2-mannosidases cleave four mannose residues from $\text{Man}_9\text{GlcNAc}_2$, and several different isomeric intermediates are formed in this process. The murine α -mannosidase IA and IB, for instance, have different but overlapping specificity for the various intermediates which are formed during mannose

removal, and together are very efficient at fully removing the mannose residues to produce the $\text{Man}_5\text{GlcNAc}_2$ precursor for complex N-glycan formation (Lal *et al.*, 1998). Analysis of *D. melanogaster mas-1* mutants has also revealed a partially redundant N-glycan processing pathway which is likely due to the action of multiple α -mannosidases with overlapping specificities. Roberts *et al.* (1998) analyzed the N-glycans produced in *mas-1* mutants and found that while these mutants are able to synthesize the full range of N-glycans (Man_9 - Man_5), the ratio of the various N-glycans was quite different. In the *mas-1* mutants, there was an accumulation of intermediates such as Man_8 and Man_7 , while much less Man_6 and Man_5 were produced. It is conceivable that the three Class I α -1,2-mannosidases in *A. nidulans* also perform a similar complementary role in N-glycan processing. Purification of the *A. nidulans* α -1,2-mannosidases and determination of their substrate specificity will clarify their role in N-glycan processing.

The engineering of *in vivo* processing of N-glycans from lower eukaryotes to the complex N-glycans of higher eukaryotes requires two initial conditions: 1) GnT-I activity and 2) a suitable substrate (ie. $\text{Man}_5\text{GlcNAc}_2$) upon which GnT-I can act. The first step has already been achieved in *A. nidulans* by insertion and expression of the GnT-I gene (Kalsner *et al.*, 1995). The production of a suitable substrate for GnT-I activity may be more problematic, as there may be a 'bottleneck' preventing the production of significant amounts of $\text{Man}_5\text{GlcNAc}_2$. Efficient removal of mannose in the ER and Golgi would provide the necessary precursors for the production of complex N-glycans. The identification of three Class I α -mannosidase genes in *A. nidulans* is the first step in clarifying the process of mannose removal during N-glycan processing in filamentous fungi. Manipulation of this region of the pathway by controlled overexpression of the

three Class I α -mannosidases may clear the 'bottleneck' and allow production of complex N-glycans.

CHAPTER 4 - Expression and Secretion of the Class I α -1,2-mannosidase IC from *Aspergillus nidulans*

4.1 Introduction

The role of the Class I α -1,2-mannosidases in the post-translational processing of glycoproteins is well characterized in higher eukaryotes (Daniel *et al.*, 1994; Kornfeld and Kornfeld, 1985; Moremen *et al.*, 1994) and yeasts (Herscovics, 1999), but less is known about these enzymes in filamentous fungi. Single members of this gene family have been identified and characterized in several filamentous fungal species, including *Penicillium citrinum*, *Trichoderma reesei*, and *Aspergillus satoii* (Inoue *et al.*, 1995; Maras *et al.*, 2000; Yoshida and Ichishima, 1995). The recent sequence characterization of three Class I α -1,2-mannosidase (*mnsIA*, *mnsIB*, and *mnsIC*) genes in the filamentous fungus *Aspergillus nidulans* was the first report of multiple mannosidases in a single filamentous fungal species (Eades and Hintz, 2000b) and suggested the possibility of functional redundancy in the protein N-glycosylation pathway of filamentous fungi (reviewed in Eades and Hintz, 2000a) in a manner similar to that found in human, mouse, and *D. melanogaster* pathways (Lal *et al.*, 1998; Roberts *et al.*, 1998; Tremblay and Herscovics, 2000). These organisms also appear to contain multiple Class I α -1,2-mannosidases with overlapping specificities and somewhat different tissue expression patterns.

The biochemical characterization of fungal Class I α -1,2-mannosidases has helped to elucidate the role of these genes in protein N-glycosylation. Recombinant expression and purification of the *A. satoii* Class I α -1,2-mannosidase allowed refined assessment of the specific biological activity of this protein (Ichishima *et al.*, 1999). The enzyme was

directed to the extracellular media by replacement of the N-terminal region of the protein with the aspergillopepsin I secretion signal, a method which allowed for efficient secretion and purification in high yields. Substrate specificity analysis confirmed that the enzyme cleaved mannose- α -1,2-mannose bonds, but did not cleave mannose- α -1,3-mannose, or mannose- α -1,6-mannose. The *A. satoi* α -1,2-mannosidase was also able to digest the oligosaccharide Man₉GlcNAc₂, which was cleaved primarily to Man₅GlcNAc₂. It was anticipated that the primary substrate would be Man₉GlcNAc₂, but the enzyme had a higher preference for Man₈GlcNAc₂ as a primary substrate than Man₉GlcNAc₂. The digestion of these substrates produced several intermediate structures, primarily different isomers of Man₇GlcNAc₂, suggesting that the enzyme had a lower activity on Man₇GlcNAc₂. This points to a possible role for other α -1,2-mannosidase genes which may have differing substrate specificities, increasing the efficiency of the pathway and preventing the accumulation of intermediates in the processing of N-glycans from Man₉GlcNAc₂ to Man₅GlcNAc₂. This can be verified by identification of each of the Class I α -1,2-mannosidase genes and determination of the substrate specificities of the enzymes produced by these genes. The cloning of all three members of the Class I α -1,2-mannosidase gene family from *A. nidulans* was recently completed (Eades and Hintz, 2000b). Due to their high amino acid similarity, the *A. nidulans* α -1,2-mannosidase IB gene is presumed to have similar biochemical properties as the *A. satoi* α -1,2-mannosidase. We have thus chosen to examine the α -1,2-mannosidase IC from *A. nidulans* as it could have a distinct cellular function and/or substrate specificity.

Several properties of *A. nidulans* make it an attractive host for the overexpression of the α -1,2-mannosidase IC protein. An inducible promoter system exists with which it is

possible to selectively express proteins to mitigate against possible toxic effects of protein overexpression (Davies, 1991). As well, numerous selectable marker systems are available, and techniques for the introduction and integration of DNA are well established.

Filamentous fungi are able to secrete very high levels of protein, especially endogenous proteins (Archer *et al.*, 1994; Archer and Peberdy, 1997). Proteins expressed in their natural host are much less susceptible to degradation by proteases, and are more efficiently secreted. This will allow production of large quantities of the *A. nidulans* α -1,2-mannosidase IC by expression in *A. nidulans* which will be useful for *in vitro* studies. In addition, the effects of overexpression can be examined in an endogenous system, which may have bearing on the remodelling of N-glycan pathways in recombinant host expression systems.

The efficient expression of the *Aspergillus nidulans* Class I α -1,2-mannosidase IC protein in *A. nidulans* cells is described here. Expression is driven by an inducible promoter, and secretion into the extracellular media is facilitated by the replacement of the N-terminal transmembrane region with a secretion signal. This will permit easy purification of a recombinant form of the enzyme for further protein activity studies which will allow further understanding of the role of this enzyme in N-glycan processing. It will also be possible to determine potential cytotoxic effects of α -1,2-mannosidase overexpression. It is our goal to utilize these enzymes for the selective modification of N-glycans *in vivo*. Such work may require overexpression of these enzymes in the secretory pathway of the ER and/or Golgi apparatus. Identification of possible cytotoxic effects of α -1,2-mannosidase overexpression will help determine the feasibility of such work.

4.2 Materials and Methods

4.2.1 Strains and Media: The *A. nidulans* expression host T580 (ura⁻), a derivative of strain FGSC4 (Fungal Genetics Stock Center) was used for expression studies. Cultures were grown in CYM liquid media described in Kalsner *et al.* (1995) supplemented with 10 mM uridine. Protoplasts which integrated the *pyrG* selectable marker (pFB94) were selected for uridine prototrophy on minimal media (Kalsner *et al.*, 1995) supplemented with 0.6 M sucrose as an osmoticum. Selected transformants for overexpression of the mannosidase were grown in liquid yeast-fructose-threonine (YFT) (5 g yeast extract; 2 g fructose; 12 g threonine; 10 ml 100X salt solution; 1 ml 1000X trace elements; and 0.6 g NaNO₃, per litre). Strains were maintained on CYM agar, and spore suspensions were obtained by washing cultured CYM agar plates with 8 ml 0.001% Tween 80. Mycelium for DNA isolations was obtained by inoculating 500 ml liquid CYM with 10⁸ spores, and incubating 24 hours at 30°C with constant agitation (200 rpm). Expression cultures were grown by inoculating 50 ml liquid YFT media with 10⁸ spores, and incubating at 30°C with constant agitation (200 rpm) for up to 72h.

4.2.2 Construction of Expression Vector: The expression vector was created by replacing the N-terminal type-II transmembrane region of the *A. nidulans* α -1,2-mannosidase IC (*mnsIC*) gene (AF233287) with the synthetic signal sequence, MDRFLGRHLGLLRHCLRQ by tailed PCR amplification and fusion to the inducible *alcA* promoter. The forward primer 5'-

GGGAAACCAATGGACCGGTTCCCTCGCCGTCATCTCGGCCTTCTTCGCCACTGCC
TTCGCCAAGACACCACGCGCCGCCAATAGC-3' contained an *NcoI* restriction

enzyme site (*italics*) to facilitate fusion with the *alcA* promoter. The 5'-overhanging tail is underlined, and the region annealing to the DNA template is marked in bold. The forward primer was designed to anneal at position +91 of the coding region of the *mnsIC* gene, immediately downstream of the transmembrane domain of the original protein. The reverse primer 5'-GGAGGATGGGGACGAGTGGG-3' was designed to anneal to the reverse complement of the *mnsIC* coding region, from position +567 to +547, immediately downstream of an *EcoRI* restriction enzyme site that was used for cloning. The secretion signal replacement fragment was produced by PCR amplification using high fidelity *PfuTurbo* (Stratagene, La Jolla CA) polymerase. The proofreading capability of the *PfuTurbo* polymerase reduced the incidence of base mis-incorporation, however the enzyme also had a tendency to cleave the overhanging regions of the tailed PCR primers. To prevent tail cleavage and provide an extended substrate for subsequent amplification reactions, 5 amplification cycles were initially performed using *Taq* polymerase (Pharmacia) with the following reaction conditions. Each reaction contained 10-100 ng of genomic DNA, 50 pmol of each primer, 2 μ l 10X *Taq* Buffer (provided), 200 μ M each of dATP, dCTP, dTTP, and dGTP and 2 units *Taq* DNA polymerase in a final volume of 20 μ l. Pre-amplification with *Taq* polymerase resulted in the creation of sufficient 'tailed' template, such that the subsequent amplification using *PfuTurbo* would not result in tail cleavage. For the second stage of the amplification with *PfuTurbo*, 1 μ l of the *Taq* PCR reactions as a template for the new reactions. Each reaction consisted of 1 μ l template DNA, 2 μ l 10 *PfuTurbo* reaction buffer (provided), 50 pmol of each primer, 200 μ M each of dATP, dCTP, dTTP, and dGTP, and 2 units of *PfuTurbo* polymers. Amplification products were separated by gel electrophoresis on 1.5% agarose gels, and the 500 bp

fragment was eluted from the gel into 20 μ l dH₂O using the Qiagen Gel Extraction Kit. To facilitate cloning into T-Vector (Promega), single A overhangs were added to the eluted DNA by combining 0.5 units *Taq* polymerase, 50 μ M dATP, and 1 μ l *Taq* buffer with the eluted DNA and heating to 72°C for 10 minutes. The tailed products were ligated into T-Vector using the Promega T-Vector Ligation Kit, and ligations were electroporated into *E. coli* DH10B (Stratagene, La Jolla, CA). Positive clones identified by digestion with *Nco*I and *Eco*RI were sequenced with an ABI373 automated fluorescent sequencer (Applied Biosystems, Foster City, CA). The *Nco*I/*Eco*RI fragment from a clone containing the correct DNA sequence was inserted into the similarly digested pGUET vector. This vector contains the *alcA* promoter in the cloning vector pTZ19R, with an *Nco*I site at the 3' end of the promoter. This resulted in the direct fusion of the *alcA* promoter to the secretion signal linker of the *mns*IC gene. The *alcA*-secretion linker was then moved to BlueScriptII (Stratagene) cloning vector using the flanking *Hind*III and *Eco*RI sites. The remainder of the *mns*IC gene was then added to the *alcA*-secretion linker by first inserting the *Eco*RI/*Bam*HI fragment (the 5' end of the gene) of the *mns*IC coding region from the vector AN1C-B4-1. Finally, the remainder of the *mns*IC gene and 3'-flanking region was added by inserting the *Bam*HI/*Sac*II fragment from the vector AN1C-B6-12 (the 3' end of the gene) to create the final vector, ANICSEC-2 (Figure 13).

4.2.3 Protoplasting and Transformation: Protoplasts were prepared according to the method of Fincham (1989), using Sigma Lysing Enzyme for cell wall digestion.

Protoplasts of strain T580 (*ura*⁻) were cotransformed with 1 μ g of the selectable marker pFB94, which converts transformed cells to uridine prototrophy, and with 1 μ g of ANICSEC2. Transformants were initially screened for integration of pFB94 by selection

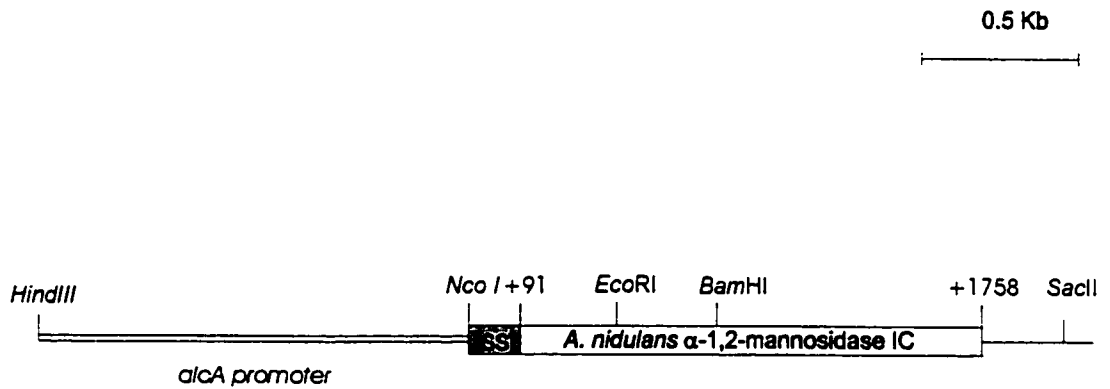


FIGURE 13: Expression vector for *A. nidulans* α -1,2-mannosidase IC. The coding region of the *A. nidulans* α -1,2-mannosidase IC from position +91 to +1758 was fused to a synthetic signal sequence (SS) and this fusion was directly linked to the *alcA* promoter. Transcription termination was via the endogenous α -1,2-mannosidase IC transcription terminator.

on minimal media, and positive transformants were then transferred to individual complete media plates. Conidia were collected using 8 ml 0.01% Tween-80, and conidia were used for subsequent inoculations. Transformants were screened for incorporation and expression of the ANICSEC-2 vector by assaying extracellular media for α -1,2-mannosidase activity.

4.2.4 Protein Expression: Fresh liquid cultures were prepared for protein expression by inoculation of 50 ml YFT media with 200 μ l of conidial suspensions. These cultures were grown on a rotary shaker (200 rpm) at 30°C for 24-72 hours. The YFT media contains limited amounts of glucose which represses the inducible *alcA* promoter. After 24-36 hours, the glucose is depleted and the cultures shift to fructose as a carbon source. This releases the glucose repression of the *alcA* promoter and allows expression of *alcA*-driven genes. One ml aliquots of media were removed from the cultures at various time intervals up to 72 hours, and centrifuged at high speed to remove the mycelium. To determine if transformant cultures were secreting high levels of protein, 80 μ l of the cleared aliquots was combined with 20 μ l 5X SDS-PAGE loading buffer, containing 0.0625 M Tris-HCl pH 6.8, 10% glycerol, 2% SDS, 0.72M β -mercaptoethanol and 0.025% bromophenol blue. Samples were boiled for 5 minutes and 35 μ l was loaded onto 4-20% Tris-HCl SDS-PAGE Mini-Protean II Ready Gels (BioRad). Gels were run for 1 hour at 100V, stained with 0.1% Coomassie Blue G-250; 40% methanol; 10% acetic acid for 30 minutes and destained in 40% methanol / 10% acetic acid for 2-4 hours.

4.2.5 Mannosidase Assays: Cleared aliquots of culture media were concentrated on NanoSep 30 kD microfilters (Pall Corporation). Protein samples were concentrated ten-fold by centrifugation of 1 ml of culture filtrate and resuspension in 100 μ l 0.01M phosphate buffer pH 6.0. Mannosidase assays were performed using the disaccharide

Man- α -1,2-Man-OCH₃ as a substrate in a coupled enzyme assay as described earlier (Scaman *et al.*, 1996), with some modifications. Digestion of the substrate was performed in a 30 μ l final volume containing 27 μ l of crude extract in 0.01M phosphate buffer (pH6.0) and 3 μ l 100 mM disaccharide Man- α -1,2-Man-OCH₃ incubated at 37°C for 3 hours. Released mannose was detected by addition of 30 μ l Tris-HCl (pH7.6) and 240 μ l of developing solution, containing glucose oxidase (55 U/ml), horseradish peroxidase (1 U/ml) and o-dianisidine dihydrochloride (70 μ g/ml) followed by incubation at 37°C for 3 hours. Absorbance measurements at 450 nm were used to determine final color change. Standard blanks included all components of the colorimetric reaction, plus the substrate and 27 μ l of concentrated protein samples. A standard curve of free mannose concentrations was determined for comparison. Specific α -mannosidase activity was standardized by comparison with total protein in the crude enzyme extracts, and was defined as the amount of mannose released from substrate per μ g of total protein per hour. Protein concentrations were determined by the Bradford method (Bradford, 1976) and comparison to a bovine serum albumin (BSA) standard curve.

4.3 Results

4.3.1 Secretion Vector Design

In order to analyze the activity and specificity of the α -1,2-mannosidase IC, it is necessary to purify significant amounts of the protein. This enzyme is typically a type-II membrane protein residing in the ER (Moremen *et al.*, 1994), although a secretable form of α -1,2-mannosidase was purified from *P. citrinum* (Yoshida *et al.*, 1993). In either case,

the natural production level of this enzyme by fungi is very low (Yoshida *et al.*, 1998) and purification is difficult. Recombinant expression of the α -1,2-mannosidase IC driven by a strong inducible promoter would drastically increase the level of production of this enzyme and make purification easier. To avoid the difficulty of purifying an intracellular enzyme, secretion of the enzyme to the extracellular media was also required. To do this, we replaced the N-terminal transmembrane domain of the protein with synthetic secretion signal, MDRFLGRHLGLLRHCLRQ and fused this recombinant protein in frame to the inducible *alcA* promoter (Figure 13). The coding region of the expression vector was followed by the endogenous transcription terminator. The secretion signal was created by tailed PCR, in which the tail of the PCR primer contained the secretion signal, as well as the necessary restriction enzyme sites for direct replacement of the transmembrane domain with the secretion signal. These restriction enzyme sites also facilitated the fusion of the coding region with the *alcA* promoter. The *alcA* promoter system is an inducible promoter which can be manipulated by changing growth media conditions. This system allows us to mitigate against possible toxic effects of the recombinant protein expression, by allowing the accumulation of fungal biomass under repressed conditions, followed by protein expression under inducing conditions.

4.3.2 Secretion of α -1,2-mannosidase IC in *A. nidulans*

The α -1,2-mannosidase IC expression vector was introduced into the genome of *A. nidulans* strain T580 by cotransformation with pFB94 vector, which contains the selectable *ura* marker. It has been shown that when two plasmids are used simultaneously in the transformation of fungal protoplasts (cotransformation), as many as 60% of

transformed protoplasts will integrate both of the plasmids into the genome (Wernars *et al.*, 1987). Transformants were selected for prototrophy on uridine deficient minimal media and transferred to complete media plates. Transformants were then grown in liquid YFT inducing media in shaking culture flasks. To determine whether any transformants were secreting high levels of α -1,2-mannosidase into the culture media, aliquots of culture media were run on SDS-PAGE gels to visualize secreted protein activity. Figure 14 shows that after 42 hours of growth there is very little secreted protein visible in the wild type T580, or the transformants #6-8. After 48 hours, the amount of secreted protein in transformant #7 increased significantly, with major protein bands appearing at \sim 90 kD and \sim 60-65 kD. The expected size of the α -1,2-mannosidase IC enzyme is \sim 65 kD, and thus this second protein band may be secreted recombinant α -1,2-mannosidase IC. The detection level of the Coomassie Blue staining procedure requires secretion levels of >20 mg specific protein / L of culture, and thus may not be sensitive enough to effectively identify secreted protein.

4.3.3. Mannosidase Activity in Culture Media

To confirm secretion of active recombinant α -1,2-mannosidase IC into the culture media of putative transformants, specific activity assays were performed. After 48 hours, aliquots of culture media were assayed for α -1,2-mannosidase activity to screen for expression of the secretion vector. In order to remove traces of culture media which would affect the activity assay, and to concentrate the enzyme, aliquots of culture media were washed with assay buffer and concentrated to 20% original volume using

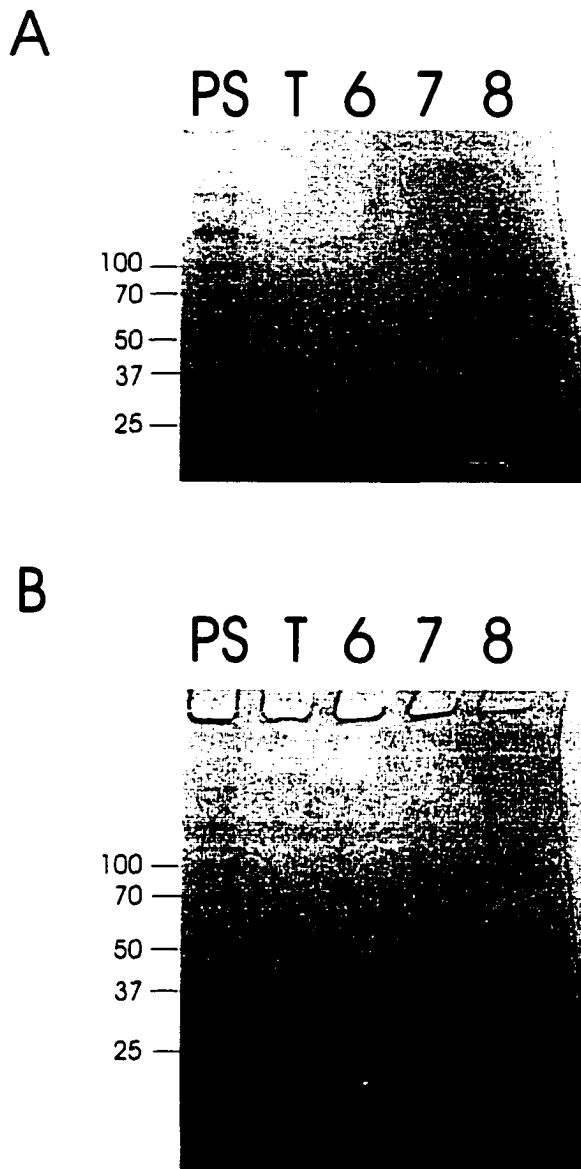


FIGURE 14: Secretion of *A. nidulans* α -1,2-mannosidase IC into culture media. Samples of cleared culture media from wild type T580 (T) and transformants 6, 7, and 8 were taken at 42 hours (A) and 48 hours (B). Protein standard (PS) sizes are shown in kilodaltons (kD). The black arrow indicates the ~65 kD protein which is likely secreted α -1,2-mannosidase IC.

microcentrifuge concentrators containing 30 kD membranes. Crude assays revealed several transformants with secreted α -1,2-mannosidase activity higher than wild type. Of these, transformant #7 showed the highest specific activity compared to wild type T580 strains. The specific activity of the culture media of transformant #7 was 50.0 Units (nmol mannose released/h/ μ g total protein) whereas the specific activity in the culture media of wild type T580 was 0.4 Units (nmol mannose released/h/ μ g protein). This assay confirms that transformant #7 secreted high levels of active recombinant α -1,2-mannosidase IC into the culture media.

4.4 Discussion

In order to begin to understand how various N-glycan structures are formed in filamentous fungi, attempts are being made to purify and characterize the various genes which have been identified in the N-glycosylation pathway of these organisms. Class I α -1,2-mannosidases have been purified and characterized in several filamentous fungi including *A. satoi* (Ichishima *et al.*, 1981; Yamashita *et al.*, 1980) and *P. citrinum* (Yoshida *et al.*, 1993). More recently, several recombinant α -1,2-mannosidases have been expressed and characterized in *A. satoi* (Ichishima *et al.*, 1999), *P. citrinum* (Yoshida *et al.*, 1998) and *T. reesei* (Maras *et al.*, 2000). These proteins share relatively high similarity at the amino acid level - the *T. reesei* α -1,2-mannosidase is 51.6% and 51.0% similar to the *A. satoi* and *P. citrinum* proteins, whereas the *A. satoi* and *P. citrinum* α -1,2-mannosidases are 70% similar. All of these recombinant α -1,2-mannosidases cleaved mannose- α -1,2-mannose linkages, but not mannose- α -1,3-mannose, or mannose- α -1,6-

mannose linkages, and shared similarity in pH optimum, temperature optimum, reaction kinetics and sensitivity to the inhibitor deoxymannojirimycin (dMM). All of these recombinant enzymes were able to degrade high mannose substrates such as $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$, but did not produce any structures smaller than $\text{Man}_5\text{GlcNAc}_2$, even under exhaustive conditions. When the substrate specificity of the *P. citrinum* and *A. satoi* recombinant enzymes was further analyzed, it was noted that there was an accumulation of glycan intermediates during the digestion of $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$, particularly when the enzyme:substrate ratios were reduced to more biologically significant levels (Ichishima *et al.*, 1999; Yoshida *et al.*, 1998). These enzymes readily digested $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_7\text{GlcNAc}_2$, which was then reduced to $\text{Man}_5\text{GlcNAc}_2$ at a much slower rate, resulting in the accumulation of $\text{Man}_7\text{GlcNAc}_2$. This reduced efficiency suggests a possible role for another Class I α -1,2-mannosidase in the N-glycan processing pathway.

Three different Class I α -1,2-mannosidase genes were recently identified and cloned from *A. nidulans* (Eades and Hintz, 2000b). One of these genes (*mnsIB*) encodes a putative protein product that bears high similarity to the *A. satoi* and *P. citrinum* Class I α -1,2-mannosidases and is likely the homologue of the genes encoding these proteins. To characterize this gene family and further understand the process of N-glycan trimming in filamentous fungi, we characterized the α -1,2-mannosidase IC from *A. nidulans* with the purpose of determining whether this enzyme exhibited significant similarities or differences with other recombinant α -1,2-mannosidases from filamentous fungi. An expression system was designed which allowed the protein to be secreted into the extracellular media from the organism from which the gene was cloned, *A. nidulans*. This is more advantageous

than heterologous expression of the protein in other organisms, such as yeasts. Yeast expression systems can often hypermannosylate proteins, which may alter their folding patterns and biochemical properties (Buckholz and Gleeson, 1991; Gellissen *et al.*, 1992), although this does not appear to be as significant a problem in some systems, such as *Pichia pastoris* (Cereghino and Cregg, 2000). Homologous protein expression is often much more efficient than heterologous protein expression, a phenomenon which may be due in part to the sequestration and degradation of heterologous proteins prior to secretion into the extracellular media (Archer and Peberdy, 1997; Gellissen *et al.*, 1992; van den Hombergh *et al.*, 1997).

In order to efficiently express a protein, it is often necessary to utilize a highly active promoter (Davies, 1991; Gwynne *et al.*, 1987). We utilized the *alcA* promoter of *A. nidulans* for the expression of the α -1,2-mannosidase IC from *A. nidulans*. The *alcA* promoter is a strong inducible promoter involved in the ethanol regulon in *A. nidulans*, which can be induced to very high levels of transcription (Davies, 1991; Hintz and Lagosky, 1993). This promoter, which normally drives the expression of the alcohol dehydrogenase I gene, is subject strong carbon catabolite repression in the presence of glucose by the *creA* negative control element. Under glucose depleted conditions, *alcA* is induced by ethanol utilizing the transcriptional activator *alcR*. It is thus possible to utilize a dual phase approach in which biomass is allowed to accumulate in the presence of glucose, followed by induction of the *alcA* promoter under glucose depleted allowing expression of the gene of interest. Using this system, we were able to mitigate against possible toxic effects resulting from the expression of the α -1,2-mannosidase IC protein.

A second consideration in the expression of the α -1,2-mannosidase IC protein was that this enzyme is hypothesized to be a type-II membrane bound protein localized to the ER/Golgi (Eades and Hintz, 2000b). It would be desirable to secrete this protein to the extracellular media in order to easily purify and characterize the protein. Thus, the N-terminal transmembrane region of the protein was replaced with a synthetic secretion signal. This secretion signal was 18 amino acids in length, and would presumably be cleaved prior to secretion of the promoter. Since the N-terminal regions of the Class I α -1,2-mannosidases are not necessary for the catalytic activity of these proteins (Moremen *et al.*, 1994), removal of this region will not affect the biochemical properties of the secreted protein.

The recombinant α -1,2-mannosidase IC protein was thus secreted in large quantities into the extracellular media of *A. nidulans*. The activity of this enzyme was verified by mannosidase assays. It will now be possible to easily purify this enzyme for further biochemical characterization. With the enzyme purified, it will be possible to determine the substrate specificities, inhibitor profiles, and kinetics of this enzyme and compare it to the Class I α -1,2-mannosidases from other fungal species. Since this protein has relatively low similarity to the *A. satoi*, *P. citrinum* and *T. reesei* proteins, it will be possible to determine whether the multiple Class I α -1,2-mannosidases which exist in filamentous fungi have similar, or distinctly different functions.

The N-glycosylation pathway of filamentous fungi appear to be similar to the mammalian N-glycan pathways. Both systems contain multiple Class I α -1,2-mannosidases (Eades and Hintz, 2000b; Herscovics *et al.*, 1994; Lal *et al.*, 1998; Lal *et al.*, 1994; Tremblay *et al.*, 1998; Tremblay and Herscovics, 2000) which specifically remove α -1,2-

mannose units from $\text{Man}_9\text{GlcNAc}_2$ during N-glycan processing. The types of N-glycan structures produced by these systems also bear similarity in that processing of the $\text{Man}_9\text{GlcNAc}_2$ precursor results in the removal of up to 4 mannose units to produce structures as small as $\text{Man}_5\text{GlcNAc}_2$ (Chiba *et al.*, 1993; Kornfeld and Kornfeld, 1985; Maras *et al.*, 1997a; Maras *et al.*, 1999). Subsequent processing in the mammalian glycosylation pathway further modifies the N-glycans to produce a variety of complex structures containing galactose, GlcNAc, and sialic acid. Filamentous fungi, however, do not produce complex N-glycans, but rather, produce a variety of mannosylated N-glycans which range from small structures (ie. $\text{Man}_5\text{GlcNAc}_2$) to higher mannose containing structures (Maras *et al.*, 1999).

The ability of purified α -1,2-mannosidases from filamentous fungi to digest $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$ (Ichishima *et al.*, 1999; Yoshida *et al.*, 1998) suggested that these enzymes could be involved in N-glycan trimming reactions in the ER and Golgi. However, the non-recombinant α -1,2-mannosidases from *A. satoi* and *P. citrinum* were purified from the culture media, indicating that these enzymes were secreted from the cell. In addition, Maras *et al.* (2000) predicted from the amino acid sequence that the *T. reesei* α -1,2-mannosidase contained a cleavable signal sequence that could also produce a secreted protein product, although no detectable α -1,2-mannosidase activity was found in the culture media of this organism. However, the presence in filamentous fungi of N-glycans as small as $\text{Man}_5\text{GlcNAc}_2$ suggests that Class I α -1,2-mannosidases are actively trimming N-glycans intracellularly in the secretory pathway. Several possibilities exist to explain the trimming of N-glycans in filamentous fungi. It is possible that the enzymatic activity found in the extracellular media of certain filamentous fungi results from

proteolytic cleavage of some membrane bound Class I α -1,2-mannosidase, resulting in the secretion of some α -1,2-mannosidase activity. The non-cleaved membrane bound α -1,2-mannosidase could remain in the secretory pathway in the ER and Golgi for the post-translational trimming of oligosaccharides in the N-glycan pathway. Alternatively, it is possible that the α -1,2-mannosidases from filamentous fungi are secreted proteins, and that mannose trimming occurs in the lumen of the ER and Golgi by non-membrane bound α -1,2-mannosidases. This trimming could also occur at the cell surface, a process that has already been suggested in mammalian cells (Porwoll *et al.*, 1998). However, this does not explain the fact that α -1,2-mannosidase activity is not found in the culture media of *T. reesei*, although mannose trimming is evident in this species. A third, and more likely possibility, is that multiple Class I α -1,2-mannosidases are found in any given fungal species, and that these different enzymes have differing substrate specificities and/or subcellular localizations to perform different or overlapping tasks. The duplication and refinement of α -1,2-mannosidase substrate specificity could allow diversification of the N-glycan pathway in these organisms for different functions. Localization of these enzymes in the secretory pathway would help elucidate such diversification.

We have been characterizing the genetic components of the N-glycosylation pathways in filamentous fungi. By manipulation of specific regions of the pathway, it may be possible to modulate the types of N-glycans which are produced by these organisms, which may make them more suitable for the heterologous expression of certain types of proteins, especially human therapeutic proteins. In order to produce proteins containing complex N-glycans, it is first necessary for the oligosaccharide precursor $\text{Man}_9\text{GlcNAc}_2$ to be completely trimmed to $\text{Man}_5\text{GlcNAc}_2$. This provides the substrate for subsequent

reactions leading to the production of complex N-glycans. Before we can begin adding the genetic components of the glycosylation pathway which will allow production of complex N-glycans in these organisms, we need to ensure that there are no 'bottlenecks' in the pathway preventing the full trimming of the $\text{Man}_9\text{GlcNAc}_2$ precursor molecule. It is hoped that overexpression of Class I α -1,2-mannosidases will effectively remove such 'bottlenecks' and drive the production of sufficient levels of substrate for subsequent reactions. We have determined that overexpression of the α -1,2-mannosidase IC gene is not likely to present any detrimental toxicity effects should we decide to overexpress the protein intracellularly. The overexpression of the α -1,2-mannosidase IC protein is the first step in the characterization of this protein, and the utilization of this protein for the *in vivo* and/or *in vitro* modification of N-glycan structures in filamentous fungi.

CHAPTER 5 - Cloning, Sequence Characterization, and Disruption of Class I

α -1,2-mannosidase IA from *Ophiostoma novo-ulmi*

5.1 Introduction

Recent studies on host-pathogen interactions have focussed on the identification of individual pathogenicity determinants, such as pathogenicity related (PR) genes, degradative enzymes, or elicitors of host defense responses. Such approaches generally do not account for the extremely complex and multivariate nature of pathogenicity. Successful invasion of a host by a particular pathogen requires that the pathogen recognize the host tissue and establish a nutritional relationship with the host (de Wit *et al.*, 1994). This requires recognition of the host, adhesion of the pathogen to the host tissue, activation of pathogen related genes in the pathogenic fungus (PR genes), and invasion of the host tissue, often via secreted degradative enzymes. The pathogen must also be able to avoid or suppress the host defense response. Most of these early molecular events involve interactions between the cell wall/plasma membrane and secreted glycoproteins. There is growing evidence that the carbohydrate moieties of these glycoproteins play an important role in the host-pathogen interaction.

Numerous examples exist which illustrate the role of N-glycosylation in fungal pathogen systems. For example, treatment of the phytopathogen *Phytophthora* with tunicamycin, a potent inhibitor of protein glycosylation, strongly reduced the adherence of the fungus to host surfaces and inhibited the formation of infection structures required for host invasion (Bircher and Hohl, 1997; Hollenstein *et al.*, 1995). Several studies have shown the importance of N-glycans in protein secretion (reviewed in Fiedler and Simons,

1995). Di Pietro and Roncero (1996) demonstrated that the secretion of endopolygalacturonase (endoPG) and pectate lyase was completely inhibited in *Fusarium oxysporum* by treatment with tunicamycin. Dean and Anderson (1991) conducted a similar study with a xylanase from *Trichoderma viride*. Additionally, West (1981) provided evidence that the carbohydrate moieties of endoPG may also be necessary for enzyme activity. It can thus be seen that manipulation of N-glycan formation can simultaneously affect a number of processes which are important for the effectiveness of many fungal pathogens.

The carbohydrate structures of mannoproteins are also very important for the physical and chemical properties of the fungal cell wall. The fungal cell wall consists of mostly glucan, mannoproteins, and chitin (Brul *et al.*, 1997; Ruiz-Herrera, 1992). The mannoproteins form a fibrillar layer within the cell wall, and are found in highest concentrations near the outer cell wall layer (Gow *et al.*, 1999; Klis *et al.*, 1997). In addition to the specific roles that the cell wall glycoproteins may play in the host-pathogen interactions, these glycoproteins are also necessary for the maintenance of the cell wall and growth by hyphal elongation. Several studies have shown that disruption of genes involved in the N-glycosylation pathway have significant effects on the properties of the cell wall (Dean, 1995; Gow *et al.*, 1999; Tanner *et al.*, 1995). Many of these glycosylation mutants showed severe morphological defects and also showed increased sensitivity to certain antibiotics, likely a result of increased permeability of the cell wall.

Cell surface carbohydrates have been directly implicated in the virulence of the opportunistic pathogen *C. albicans*, a yeast that causes candidiasis in humans. Masuoka and Hazen (1997) showed that the degree of cell wall protein mannosylation affected the

cell surface hydrophobicity (CSH) of *C. albicans*. CSH influences several important processes affecting disease development and virulence, including cell adhesion, morphogenesis and resistance to phagocytosis (Hazen, 1989). More recently, Masuoka and Hazen (1999) showed that the degree of cell wall mannosylation was determined primarily by the outer chain of cell wall protein N-glycans. Thus, manipulation of such outer chain structures could have a significant affect on the virulence of such organisms.

The role of cell surface carbohydrates has also been implicated in the pathogenicity of *S. schenckii*. This filamentous fungus is the etiological agent of sporotrichosis which causes chronic cutaneous infections in humans and animals (Crout *et al.*, 1977). The pulmonary form of sporotrichosis is particularly prevalent in immunocompromised patients. Recently, a direct correlation was observed between the cell wall carbohydrate composition, particularly rhamnose and mannose, and the virulence of the strains tested (Fernandes *et al.*, 1999). The rhamnose and mannose levels in the cell wall were observed to vary over culture time, and these differences were correlated to the ability of the fungus to cause disease symptoms in mice. This correlation led to the suggestion that the cell wall composition may influence the virulence of *S. schenckii*.

Sialic acid residues, which may be important for defense against host immune responses and for cell adhesion, have been found in *S. schenckii*, as well as in the pathogens, *Cryptococcus neoformans*, *Fonsecaea pedrosoi*, *Paracoccidioides brasiliensis*, and *C. albicans* (Alviano *et al.*, 1999; Rodrigues *et al.*, 1997; Soares *et al.*, 1993; Soares *et al.*, 1998; Soares *et al.*, 2000; Souza *et al.*, 1986). These anionic terminal sugars, typical of higher eukaryotic N-glycans, are not widely reported in fungi, thus it is interesting that these structures are found in these fungal pathogens. The sialic acids found

in these fungi have been shown to be important for the adhesion of these organisms to host tissues, and have also been implicated in avoidance of host immune clearance receptors (Alviano *et al.*, 1999). The manipulation of glycan structures may inhibit the ability of these organisms to properly add sialic acid, thus reducing their virulence.

The filamentous fungus *O. novo-ulmi* is the causal agent of Dutch Elm Disease (DED), which has decimated elm populations in most of eastern North America (Brasier, 1991; Hintz, 1999). The DED fungus is mainly vectored by the European elm bark beetle *Scolytus multistriatus* and the native elm bark beetle *Hylurgopinus rufipes* (Bowden *et al.*, 1994). Berbee and Taylor (1992a) have suggested a close relationship of *Sporothrix* and *Ophiostoma* based on 18S rRNA sequence divergence. The close relationship of *S. schenckii* and *O. novo-ulmi* was recently confirmed by comparison of chitin synthase I genes from these and other fungal species (Hintz, 1999). The close relationship of these organisms suggests that cell wall and secreted glycoproteins may be important in the virulence of *O. novo-ulmi*, and that studies in this plant pathogen may also have applicability in the more dangerous human pathogen. *O. novo-ulmi* offers many advantages as an experimental organism and a model for host-pathogen interactions. An efficient transformation system has been developed, allowing for ease of genetic manipulation (Royer *et al.*, 1991). Similar to *S. schenckii* this fungus is dimorphic and depends on both the budding yeast and filamentous hyphal phases for pathogenicity. The investigation of pathogenic determinants in this organism, and the development of molecular tools specific for the organism will allow the development of an important model system for studying host-pathogen interactions.

Protein N-glycosylation is an ordered process which occurs in all eukaryotic species. The early steps of this process, which include the transfer of the oligosaccharide precursor $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ to the nascent polypeptide and the removal of the three glucose residues, are almost identical in all eukaryotes (Kornfeld and Kornfeld, 1985). At this point, the pathways in various organismal lineages can become quite divergent resulting in quite different N-glycans structures as the end point in the pathway (Herscovics, 1999). In higher eukaryotes (ie. mammals), α -1,2-mannosidases remove a total of four mannose residues, yielding $\text{Man}_5\text{GlcNAc}_2$ which is the precursor for complex N-glycans. Filamentous fungi and yeasts, on the other hand, do not produce complex N-glycans, but produce high mannose N-glycans through a series of steps involving mannose removal by α -mannosidases followed by the addition of numerous mannose residues by mannosyltransferases to form the 'outer chain'.

The α -mannosidases have been previously classified into two independently derived groups, Class I and Class II, based on biochemical properties, substrate specificity, inhibitor profiles, and sequence alignments (Daniel *et al.*, 1994; Eades *et al.*, 1998; Moremen *et al.*, 1994). The Class II α -mannosidases are a heterogeneous group of enzymes that includes the lysosomal mannosidases, the Golgi mannosidase II and a distantly related group of enzymes, including the rat ER/cytosolic mannosidase (Bischoff *et al.*, 1990), yeast vacuolar mannosidase (Yoshihisa and Anraku, 1989), and the *A. nidulans* Class II mannosidase (Eades *et al.*, 1998). Some of these enzymes are involved in oligosaccharide catabolism in the cytosol and vacuole. The Golgi mannosidase II is involved the production of complex N-glycans. Thus, these enzymes seem to be restricted to higher eukaryotes.

The Class I α -1,2-mannosidases are involved in the initial trimming of mannose sugars early in the N-glycan processing pathway and are thus ubiquitous in eukaryotes. Several Class I α -mannosidases have already been characterized in filamentous fungi. Single members of the Class I α -mannosidase genes have been identified in *A. satoi* (Inoue *et al.*, 1995), *P. citrinum* (Yoshida and Ichishima, 1995), and *Trichoderma reesei* (Maras *et al.*, 2000). More recently, three different Class I α -mannosidase genes were identified in *A. nidulans* (Eades and Hintz, 2000b). Here we describe the identification, sequencing and analysis of a α -1,2-mannosidase from *O. novo-ulmi*.

5.2 Materials and Methods

5.2.1 Strains, Media and Growth Conditions: *O. ulmi* strain MH75 and *O. novo-ulmi* strain VA30 and were maintained on OCM agar plates and mycelium for DNA isolation was grown in liquid OCM (Royer *et al.*, 1991) at room temperature.

5.2.2 Oligonucleotide Primer Design: Degenerate oligonucleotide priming sites were selected by aligning four published α -1,2-mannosidase sequences – human (Bause *et al.*, 1993), rabbit (Lal *et al.*, 1994), yeast (Camirand *et al.*, 1991) and mouse (Lal *et al.*, 1994). Highly conserved blocks which showed low codon redundancy were chosen to design the PCR primers. The forward primer MANFOR was designed by reverse translation of the protein sequence GGLGESFYEY, and the return primer MANREV was designed from the complement of the reverse translation of the sequence FXLAETLKYLY. A codon usage table compiled for *A. nidulans* was used to aid in nucleotide selection at degenerate sites.

5.2.3 DNA Isolation and PCR amplification: Total genomic DNA was extracted from finely ground freeze-dried mycelium of *O. novo-ulmi* strain MH75. Approximately 400 mg of mycelium was vortexed with 2.5 ml of 50 mM EDTA; 0.2% SDS, centrifuged for 10 minutes and then 85 μ l of 3M KOAc; 5M acetic acid added to the supernatant. Following a 20 minute incubation on ice, the suspension was recentrifuged and DNA was isopropanol precipitated from the supernatant. After resuspension in 100 μ l TE (10 mM Tris pH 7.5; 1 mM EDTA), the DNA was extracted once with phenol, twice with chloroform/isoamyl alcohol (24:1) and ethanol precipitated. Each PCR reaction consisted of 10 - 100 ng of genomic DNA, 50 pmol of each primer, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 50 mM KCl, 0.01% gelatin, 0.1% Triton X-100, 200 μ M each of dATP, dCTP, dTTP, and dGTP and 0.2 units Taq DNA polymerase (Perkin-Elmer) in a final volume of 100 μ l. Amplification was performed in two stages using Perkin-Elmer thermal cycler. Five cycles at a lower stringency annealing temperature (56°C) were followed by 30 cycles at higher stringency (63°C). The PCR products were eluted from 1% low-melting point agarose, cloned into T-Vector (Promega) using the T-Vector cloning system and sequenced with Universal Forward and M13 Reverse primers.

5.2.4 Library Construction and Screening: The *O. novo-ulmi* α -mannosidase gene fragment was used as a template for a second round of PCR with the incorporation of radionucleotides to produce a probe for library screening. The PCR reactions were performed as previously described, except that the concentration of "cold" dCTP was reduced by 50% and 0.825 μ M α -³²P-dCTP (5 μ Ci) was added to the reactions. The radiolabelled PCR products were purified with the Wizard PCR Prep Kit (Promega) and label incorporation assessed by scintillation counting.

A library of *O. novo-ulmi* (MH75) genomic sequences were constructed by digesting genomic DNA with *MboI* and ligating the resulting fragments on to *BamHI* pre-digested arms of lambda DNA vector EMBL-3. Concatamers of the ligated DNA were packaged using the Gigapack II (Stratagene) *in vitro* packaging system. Approximately 10^5 recombinant lambda plaques were immobilised on nylon membranes (Genescreen Plus, Dupont) and hybridized with the radiolabelled α -1,2-mannosidase specific probe DNA. Single hybridizing clones were selected and rescreened. Lambda DNA was isolated with the Wizard Lambda DNA Extraction Kit (Promega), digested with restriction enzymes and a 3.6 Kb *EcoRI* fragment containing the full length α -1,2-mannosidase gene was identified by Southern blot (Southern, 1975) and subcloned into pUC18.

5.2.5 DNA preparation and sequencing: Sequencing of the 3.6 Kb insert was accomplished using the “shotgun” sequencing method. Approximately 50 μ g of the *EcoRI* insert was purified from a 1.5% low melting temperature agarose gel (Wizard Preps, Promega), randomly sheared and the resultant fragments were ligating into the M13mp19 vector. Shearing was accomplished by fragmentation of the DNA in a medical nebulizer at 30 bps for 2 minutes followed by digestion with 50U Mung Bean nuclease to remove internal single strand gaps. The DNA was repaired by the addition of 40U T4 DNA Polymerase and 25U Klenow fragment. The DNA was then electrophoresed on a 2% NuSieve agarose gel, and fragments ranging from 0.7-1.0 Kb were excised, purified and blunt-end ligated into M13mp19. Ligated DNA was transformed by electroporation into *E. coli*.

Recombinant plaques were selected on agar plates supplemented with X-Gal and IPTG grown overnight at 37°C. The DNA templates were purified using the Qiagen Kit

(Qiagen) and diluted to 100 ng/ μ l. The automated fluorescence sequencing was performed on a ABI 373A automated sequencer (Applied Biosystems, Foster CA). Sequence data was generated for approximately 100 plaques, and the sequences were aligned and compiled using Seqman in the DNASTar program (DNASTar, Madison, WI), resulting in 7-fold redundancy at each base position.

5.2.6 Identification of Introns: Putative introns were located by examination of the DNA sequences for intron consensus sequences in combination with reading frame shifts.

Putative introns were verified by comparison of genomic DNA sequence of the α -1,2-mannosidase gene to reverse transcribed mRNA. Total RNA was extracted from mycelium grown in 150 ml liquid media using the guanidine isothiocyanate method (TRIzol, GibcoBRL). Approximately 1 gram of fresh mycelium was harvested by suction filtration was homogenized with 6 ml of TRIzol reagent, and incubated at 30°C for 10 minutes. The homogenate was extracted with 1.2 ml chloroform, and the mixture was shaken vigorously for 15 seconds, incubated at 30°C for 2 minutes and subjected to centrifugation at 3000 rpm for 15 minutes. The upper aqueous phase was removed and 3 ml isopropanol was added to precipitate the RNA. After incubation at 30°C for 10 minutes, the RNA was pelleted by centrifugation and washed with 70% ethanol/diethyl pyrocarbamate (DEPC). The RNA pellet was resuspended in 30 μ l H₂O/DEPC and 5 μ l was used for cDNA preparation and PCR amplification.

Reverse transcriptase PCR was performed using Superscript II reverse transcriptase (GibcoBRL) for first strand cDNA synthesis using an oligo-dT primer followed by PCR amplification of the cDNA using α -1,2-mannosidase specific primers. The oligo-dT primer was annealed to the RNA by adding 500 ng oligo-dT to

approximately 1-3 μg RNA in a final volume of 10 μl $\text{H}_2\text{O}/\text{DEPC}$. The mixture was heated to 70°C for 10 minutes and then quickly chilled on ice. Reverse transcription was then started by addition of 4 μl First Strand Buffer (250 mM Tris-Cl pH8.3; 375 mM KCl; 15 mM MgCl_2), 2 μl 0.1 M dTT, 1 μl 10 mM dNTPs, and 1 μl (200U) Superscript II reverse transcriptase followed by incubation at 42°C for 50 minutes. After first strand cDNA synthesis, the RNA was digested by the addition of 1 μl (2U) RNase H and incubation at 37°C for 20 minutes.

The cDNA was used to amplify regions containing putative introns for comparison with genomic DNA amplification. Primers were designed from genomic sequences flanking the putative intron (OU-IF1 and OU-IR1). PCR amplification of the cDNA was performed in Perkin-Elmer thermal cycler for 30 cycles (94°C, 1.5 min.; 57°C, 1.5 min., 72°C, 2.5 min.) and each reaction contained: 2 μl first strand reaction, 50 pmol of each primer, 10 μl 10X *Taq* Buffer, 200 μM each of dATP, dCTP, dTTP, and dGTP and 0.2 units *Taq* DNA polymerase (Perkin-Elmer) in a final volume of 100 μl .

5.2.7 Design of Disruption Cassette: The disruption cassette OU1A-Dis was constructed by flanking a selectable marker gene with DNA fragments derived from the α -1,2-mannosidase. The hygromycin phosphotransferase (HPH) gene encoding hygromycin resistance under control of the *Aspergillus nidulans gpdA* promoter was contained on the pAN7-1 vector (Punt *et al.*, 1987) and was used as a selectable marker. The marker was flanked 5' and 3' with 1 Kb of PCR amplified α -1,2-mannosidase sequence. The PCR primers used in the amplification reaction were 'tailed' with 5' overhanging regions containing appropriate restriction sites for directional cloning into the disruption vector. The primer pair OUDIS1-F and OUDIS1-R were tailed with the restriction sites *Bgl*II and

BstEII, respectively and were used to amplify the region -695 to + 366 of the α -1,2-mannosidase gene using standard PCR amplification conditions (Sec. 5.2.3). This amplification product was digested with *BglII* and *BstEII*, gel purified and ligated with an equivalent amount of similarly digested pAN7-1 vector. The original restriction sites in the pAN7-1 were at position +138 (*BglII*) and +396 (*BstEII*). The stage 1 construct which contained the 1100 bp mannosidase fragment upstream of the *gpdA-HPH* marker was digested with *BamHI* and *HindIII* to facilitate the insertion of a second flanking region downstream of the the marker. This second region, corresponding to positions +1401 to +2109 of the α -1,2-mannosidase gene was also produced by 'tailed' PCR. The 'tailed' primers OUDIS2-F and OUDIS2-R, containing the restriction site *HindIII* and *BamHI*, generated a 708 bp PCR product. The amplification product was digested with *HindIII* and *BamHI* and ligated with similarly digested OUDIS1 vector. The gene disruption construct (OU1A-Dis) contained both the 1100 bp and the 708 bp mannosidase regions flanking the *gpdA*-driven *HPH* gene and was used for disruption experiments (Figure 15).

5.2.8 Disruption of the *O. novo-ulmi* α -1,2-mannosidase gene

Protoplasts of *O. novo-ulmi* strain MH75 were prepared by the method of Fincham (1989) using Sigma Lysing Enzyme and were transformed with 1 μ g of the gene disruption construct OU1A-Dis (non-linearized) and plated on OCM containing 0.6M sucrose as an osmoticum and 200 μ g/ml hygromycin B (Sigma Chemicals). Transformed colonies were selected and re-plated on two successive rounds of hygromycin selection media to ensure transformant stability and purity. DNA for PCR verification of gene disruption was prepared according to Cenis (1992).

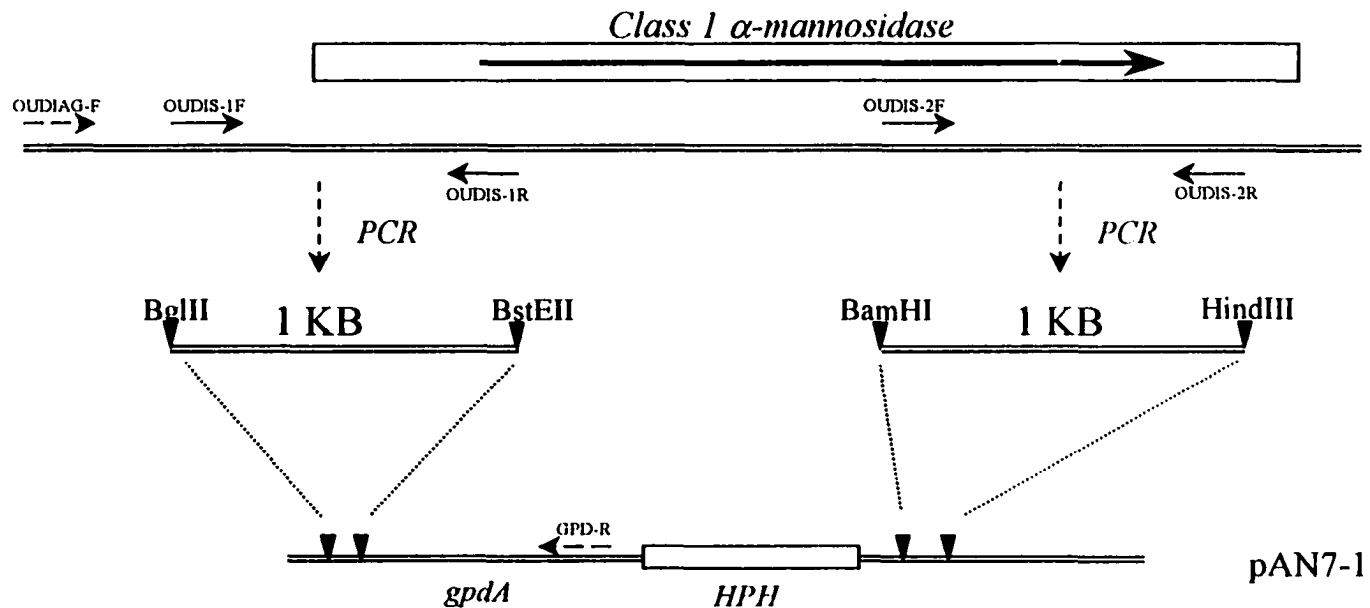


FIGURE 15: Construction of *O. novo-ulmi* α -1,2-mannosidase disruption vector. The tailed PCR primer pairs OUDIS-1F/1R and OUDIS2F/2R were used to amplify regions of the α -1,2-mannosidase gene. Tailed amplification products were directionally cloned into the vector pAN7-1, which contained the hygromycin phosphotransferase (HPH) gene driven by the glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) promoter. Integration of the disruption vector was screened using the diagnostic primers OUDIAG-F, which is upstream of the integration site and is not found in the disruption vector, and GPD-R, which is found in the *gpdA* promoter and not the α -1,2-mannosidase gene.

Gene disruptions were verified using diagnostic PCR primers. The forward diagnostic primer, OUDIAG-F was specific for sequence upstream of the α -1,2-mannosidase gene and not contained in the disruption construct and the return diagnostic primer was specific for the *gpdA* promoter which would only be found in the disruption construct, but not the wild-type genome. Amplification would only be expected if these two primers were in proximity due to integration of the disruption construct at the α -1,2-mannosidase locus. PCR amplification with the diagnostic primers OUDIAG-F and GPD-R was performed in a Perkin-Elmer thermal cycler for 30 cycles using standard conditions.

5.2.9 Sequence Alignments and Comparison: Protein sequences for Class 1 α -mannosidases were obtained from the GenBank database (National Center for Biotechnology Information – NCBI). Sequences used included *S. cerevisiae* (Camirand *et al.*, 1991 - Accession M63598), *A. nidulans* IA, IB, IC (Eades and Hintz, 2000b – AF129497, AF129496, AF233287) *A. satoi* (Inoue *et al.*, 1995 - D49827), *P. citrinum* (Yoshida and Ichishima, 1995 - D45839), *T. reesei* (Maras *et al.*, 2000 – AF212153) *D. melanogaster* (Kerscher *et al.*, 1995 - X82640), *S. frugiperda* (Kawar *et al.*, 1997 - AF005035), *H. sapiens* IA (Bause *et al.*, 1993 - X74837), *H. sapiens* IB (Tremblay *et al.*, 1998 - AF027156), *M. musculus* IA (Lal *et al.*, 1994 - U04299), *H. sapiens* IC (Tremblay and Herscovics, 2000 – AF261655), *H. sapiens* ER-Man9 (Tremblay and Herscovics, 1999 – AF148509) *M. musculus* IB (Herscovics *et al.*, 1994 - U03457), and *S. scrofa* (Bieberich *et al.*, 1997 - Y12503). The outgroup was *A. nidulans* Class IIC α -mannosidase (Eades *et al.*, 1998 - AF016850). Multiple alignments of the full length sequences were performed using the MAP alignment program (Huang, 1994) at the Baylor College of Medicine web site (BCM Search Launcher) with default parameters.

Phylogenetic relationships were determined using the more conserved C-terminal portion of each gene. Phylogenetic trees were generated using PAUP 4.0b* computer software (Swofford, 1998). After initial execution of the alignment file, trees were generated using maximum parsimony with a full heuristic search (100 replicates of random stepwise branch addition). The most parsimonious tree was rooted using the *A. nidulans* Class IIC α -mannosidase protein sequence as an outgroup. Confidence values for the tree branches were determined using the bootstrap option (100 bootstrap replicates using a full heuristic search).

5.3 Results

5.3.1 Identification of *O. novo-ulmi* Class I α -1,2-mannosidase

The *O. novo ulmi* Class I α -1,2-mannosidase was identified by utilization of degenerate PCR primers designed from conserved regions of the protein sequences from available Class I α -1,2-mannosidases. Conserved regions were identified by multiple sequence alignment of *S. cerevisiae* (Camirand *et al.*, 1991), mouse IA (Lal *et al.*, 1994), human IA (Bause *et al.*, 1993), and rabbit liver (Lal *et al.*, 1994). Selection of conserved sequences for primer design relied on high sequence identity between the four proteins in combination with a low codon redundancy. A codon preference table for the fungus *A. nidulans* was used to predict the codon redundancy in the primer. Both *A. nidulans* and *O. novo-ulmi* are Ascomycetous fungi and likely have a similar codon preference. Two primers, ManFor and ManRev, were designed which contained the least codon redundancy, and the highest annealing temperature using the computer program OLIGO. Assuming at least three base mismatches, a target annealing temperature in the range of

62-64°C was used. Amplification of *O. novo-ulmi* total genomic DNA with these two primers yielded a 750 bp product, which was close to the expected size, since the primer positions corresponded to a span of approximately 250 amino acids in the four homologous sequences. Upon sequence determination of this PCR product and searching the GenBank database for sequence homology (BLAST search), several high scoring matches were found with other Class 1 α -1,2-mannosidases, confirming that we had amplified a portion of the Class 1 α -1,2-mannosidase from *O. novo-ulmi*.

A genomic library of *O. novo-ulmi* sequences was screened and a 3.6 Kb *EcoRI* fragment containing the full length gene was subcloned and the entire sequence determined by the “shotgun sequencing” method. Copy number was determined by Southern analysis. Wild-type DNA was digested with one of several restriction enzymes (*EcoRI*, *SalI* and *SacII*) and probed with the 750 bp radiolabelled PCR fragment. The PCR fragment did not contain any internal *EcoRI*, *SalI* or *SacII* restriction sites. The probe hybridized to a single restriction fragment in each of the digests (Figure 16), indicating that this gene is found in single copy in the haploid *O. novo-ulmi* genome.

5.3.2 Sequence analysis

The α -1,2-mannosidase gene sequence *O. novo-ulmi* was analyzed to identify the translational start codon and open reading frame (ORF), including any introns, and to identify consensus sequences in the flanking regions of the genes. Reading frames were initially identified by sequence homology with published α -1,2-mannosidase sequences. The entire 3.6 Kb nucleotide sequence was compared to sequences in the GenBank database by BLASTX search, which compared the translated amino acid sequences of all

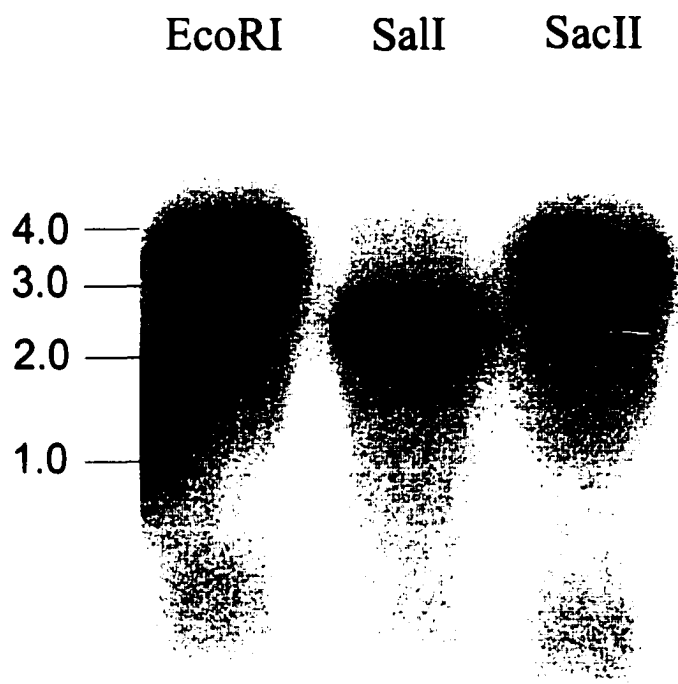


FIGURE 16: Copy number of *O. novo-ulmi* Class I α -1,2-mannosidase. Southern blots of genomic DNA was hybridized with the radiolabelled 750 bp amplification product of the PCR primer pair ManFor and ManRev. Genomic DNA was digested with either *EcoRI* (lane 1), *SalI* (lane 2) or *SacII* (lane 3). DNA ladders were used to determine sizes of the hybridizing factors (Kb).

three reading frames to protein sequences contained in the database. A single reading frame was found encoding a 625 amino acid polypeptide which contained a high degree of similarity with several α -1,2-mannosidase protein sequences in the database.

Alignment of the *O. novo-ulmi* α -1,2-mannosidase sequence with published sequences allowed detection of potential introns in the coding sequence. The full length *O. novo-ulmi* gene aligned very well with the mannosidase sequences from other filamentous fungi (*A. satoi*, *P. citrinum*, *T. reesei*), however, the *O. novo-ulmi* sequence did contain an apparent insertion of 37 amino acids in length, indicating the possibility of an intron at this location. This region did not contain any typical intron splice sequences and no shifts in the translational reading frame were noted. Reverse transcription PCR amplification of *O. novo-ulmi* cDNA followed by sequence analysis confirmed that there was no intron at this location in the *O. novo-ulmi* α -1,2-mannosidase gene and that this region does indeed represent a DNA insertion into the coding region of the gene.

The N-termini of Class 1 α -1,2-mannosidases are not very well conserved in length or composition. The location of the N-terminus of the *O. novo-ulmi* gene product was predicted based on sequence alignments, and the presence of an AUG start codon in the region corresponding to the expected length of the primary reading frame. The DNA sequence environment of the AUG start codon was compared to eukaryotic (especially fungal) consensus sequences to predict the correct start codon. The identified translation start site was embedded in the sequence CGTCATGGC. This sequence bears some similarity to other start sites in filamentous fungi, although there does not appear to be a strong consensus in these organisms (Ballance, 1986). The G at +4 and C at +5 are strongly conserved in other filamentous fungi, as is the C at -1 and the C at -4. The G at

-3 is inconsistent, since most eukaryotes (including filamentous fungi) usually have an A at this position, however, this site is not conserved in the *S. cerevisiae* α -1,2-mannosidase (Camirand *et al.*, 1991). Other supporting data for this ATG encoding the translation start site included the presence of a potential TATA-box found 129 bp upstream of this start site (TATAAA). No obvious CAAT motif was found in the vicinity of the start site, although this is not unusual. Fungal promoters also tend to contain numerous CT-rich regions (MacKenzie *et al.*, 1993), and several of these motifs were found upstream of the putative start site. The coding region of the presumptive gene was 1875 bp in length and encoded a peptide of 625 amino acids with an estimated molecular weight of 70 kD. The 3' downstream sequence did not contain a typical consensus transcription termination signal (AAUAAA), though the sequence was extremely AT-rich, which is characteristic of the downstream untranslated region of many fungal transcripts (Gurr *et al.*, 1987).

The N-terminus of the protein sequence was characteristic of a eukaryotic signal sequence and type II transmembrane domain. The first ten amino acids (*n*-region) were highly charged (net charge +4) and were followed by a large hydrophobic domain (*h*-region) of 14 amino acids as identified by Kyte-Doolittle hydrophathy plot (Figure 17). The hydrophobic domain was immediately followed by 5 polar residues (*c*-region). This is consistent with other Class 1 α -1,2-mannosidases which all contain a signal sequence and a type II transmembrane region, though the length and composition of these motifs are not well conserved.

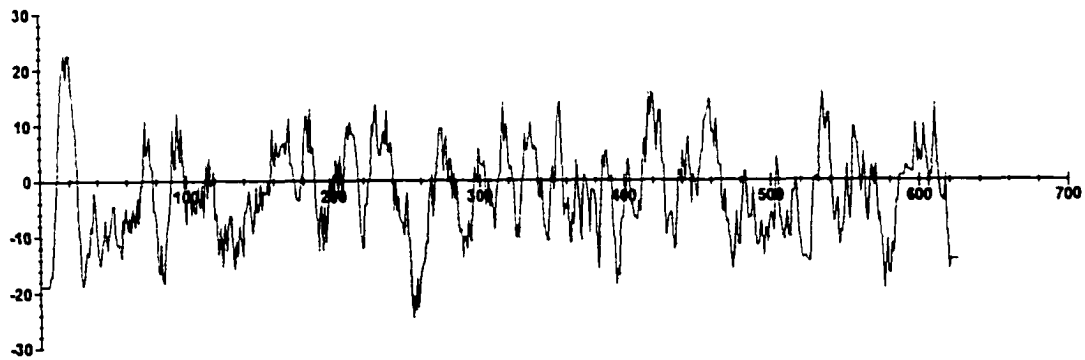


FIGURE 17: Kyte-Doolittle hydropathy plot of *O. novo-ulmi* predicted amino acid sequence. Vertical axis shows the hydrophobicity of a given region of the protein (Kyte and Doolittle, 1982) with positive values representing hydrophobic regions of the protein, and negative values representing hydrophilic regions of the protein.

5.3.3 Gene Disruption

A common method of inferring gene function *in vivo* is to disrupt the gene function by interruption of the translational reading frame of the gene. Disruption of the *O. novo-ulmi* α -1,2-mannosidase gene was accomplished by replacement of an internal region of the gene with the hygromycin phosphotransferase (*HPH*) gene. A disruption vector (OU1A-DIS) was designed which contained a *gpdA*-driven *HPH* gene flanked on the 5' and 3' ends by 700-1000 bp of α -1,2-mannosidase sequence. Wild type *O. novo-ulmi* was transformed with the disruption vector and initial screening on hygromycin selected for those protoplasts which had incorporated the *HPH* marker. To identify transformants in which homologous integration of the disruption vector resulted in replacement of the α -1,2-mannosidase gene, DNA was isolated from putative transformants and used as a template in a diagnostic PCR amplification.

The forward diagnostic PCR primer was designed to hybridize upstream of the targeted integration site in the genome, while the reverse PCR primer was designed to hybridize to the *gpdA* promoter of the disruption vector (Figure 15). This diagnostic PCR would produce a 1.3 kilobase amplification product only if homologous integration and gene replacement had occurred at the α -1,2-mannosidase gene locus. Using this method, one of fourteen of the putative transformants yielded the 1.3 kb amplification product indicating homologous integration of the disruption cassette (Figure 18). The amplified DNA was sequenced to verify that this PCR product was indeed derived from this integration event. The amplified DNA fragment contained α -1,2-mannosidase sequence located upstream of the integration site, immediately followed by α -1,2-mannosidase sequence included in the disruption cassette followed by the *gpdA* promoter sequence

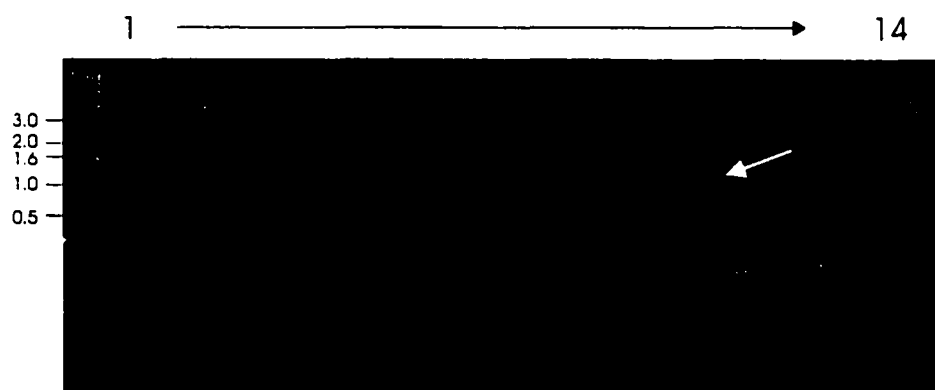


FIGURE 18: Diagnostic PCR screening of *O. novo-ulmi* disruption transformants. Gel electrophoresis of amplification products using the diagnostic PCR primers OUDIAG-F and GPD-R. The arrow indicates a PCR fragment at the expected size of 1.3 kb. DNA standards (St.) for size comparisons are shown in kilobases (kb).

from the pAN7-1 plasmid. Disruption of the Class I α -1,2-mannosidase in *O. novo-ulmi* was not lethal to the organism, although it did slightly alter the phenotype. The mutant strain exhibited a 'fluffy' morphology, containing more aerial hyphae than the wild-type strain. The mutant also exhibited a reduced radial growth rate compared to the wild-type.

5.3.4 Sequence comparison

The full length *O. novo-ulmi* α -1,2-mannosidase protein sequence was compared to other published α -1,2-mannosidase sequences obtained from the GenBank database, including homologous genes from mammals (human, mouse, pig), insect (*D. melanogaster*, *S. frugiperda*), yeast (*S. cerevisiae*), and filamentous fungi (*A. nidulans*, *A. satoi*, *P. citrinum*, *T. reesei*). The mannosidase protein sequences were relatively well conserved with respect to overall gene length, with some variability in the N-terminal regions. The yeast and fungal protein sequences were slightly smaller in length (511 to 549 aa) than the rest of the mannosidase sequences (641 to 670 aa). The *O. novo-ulmi* sequence was 625 aa in length, more similar in length to the mammalian and insect sequences than to the other fungal sequences. Multiple sequence alignments revealed that the C-terminal regions of the proteins were very well conserved, whereas the N-termini, particularly the first 50-100 amino acids were not well conserved between the various protein sequences.

A phylogenetic tree was generated from the multiple sequence alignment which provided an accurate measure of the relatedness of these gene sequences (Figure 19). Alignments were performed using the C-terminal portion of the genes containing highly conserved regions. As expected all of the fungal gene sequences, excluding yeast, grouped

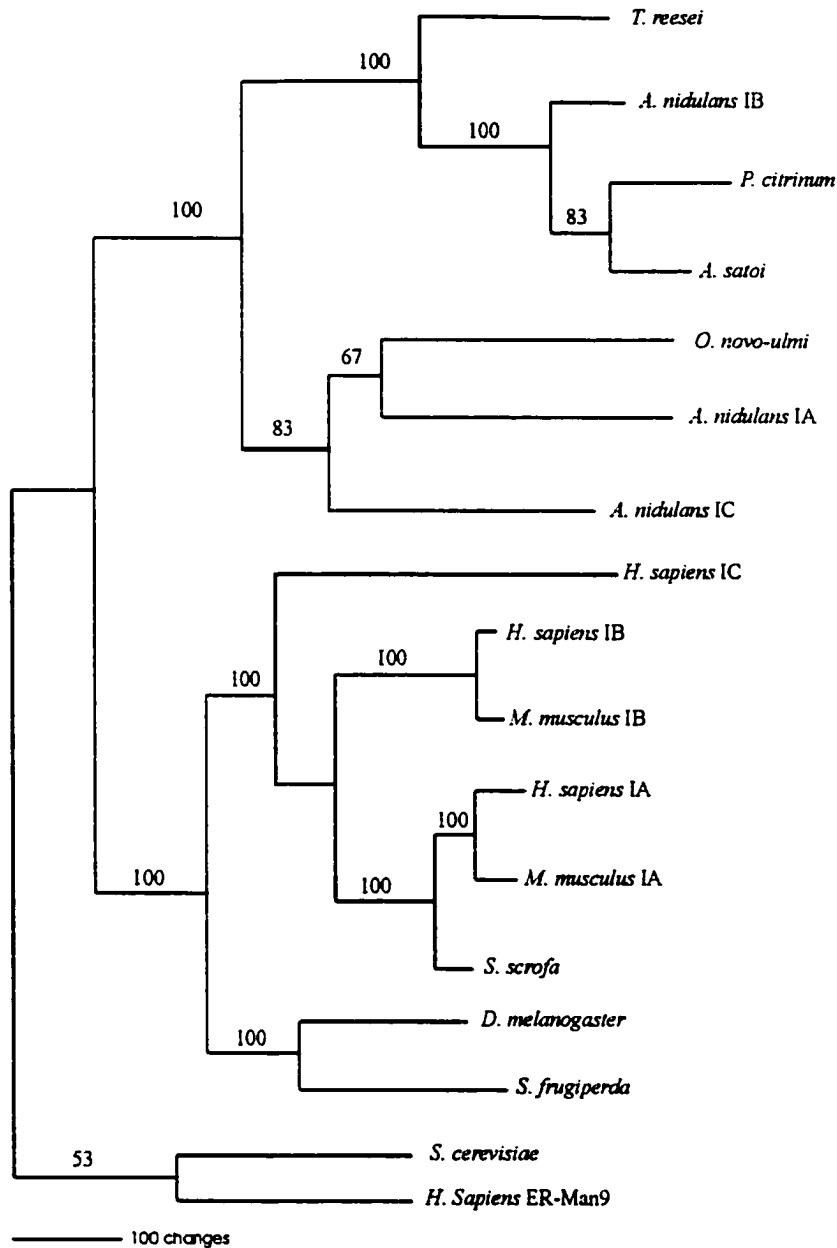


FIGURE 19: Phylogram showing sequence relationships of Class I α -mannosidases, including *O. novo-ulmi* α -1,2-mannosidase. Phylogram was generated from the multiple sequence alignment using PAUP 4.0b* analysis program. Numbers above the branches of the tree represent bootstrap confidence values, out of 100 bootstrap replications. Branch lengths represent the number of amino acid substitutions.

closely together, with the *O. novo-ulmi* sequence being closely grouped with the *A. nidulans* α -1,2-mannosidase IA and IC enzymes. The *A. nidulans* α -1,2-mannosidase IB, *A. satoi*, *P. citrinum*, and *T. reesei* α -1,2-mannosidases were monophyletic. The insect sequences also grouped together, as did the majority of the mammalian sequences. The human ER Man9-mannosidase and the yeast ER Man9-mannosidase also grouped together, although these enzymes appeared to be distantly related to the remainder of the Class I α -1,2-mannosidases. Bootstrap analysis was performed to assess the confidence values at each node in the tree. This analysis shows that the relationship of the *O. novo-ulmi* α -1,2-mannosidase, and the *A. nidulans* α -1,2-mannosidase IA and IC has low certainty, although the *A. nidulans* IB, *A. satoi*, *P. citrinum*, and *T. reesei* monophylogeny is supported with high confidence.

5.4 Discussion

Investigation of the genetic components of the glycosylation pathway in *O. novo-ulmi* may provide ideal targets for antifungal drug development. It would be ideal to identify genetic targets early in the fungal glycosylation pathways, such that inhibition or disruption of these genes may cause maximal effect. However, the viability of such targets is offset by the fact that the early stages of the glycosylation pathways are well conserved amongst eukaryotes (Buurman *et al.*, 1998; Gow *et al.*, 1999). Major divergences in the pathway between evolutionary lineages seems to begin with the removal of mannose molecules by Class I α -1,2-mannosidases in the ER and Golgi. This region of the fungal glycosylation pathway may thus provide an ideal antifungal target. The diversity of the Class I α -1,2-mannosidase gene family in filamentous fungi is being mapped with the

sequence analysis of several novel genes from different species, including: *A. satoi* (Inoue *et al.*, 1995), *P. citrinum* (Yoshida and Ichishima, 1995), *T. reesei* (Maras *et al.*, 2000), and *A. nidulans* (Eades and Hintz, 2000b). It is expected that the diversity found in a particular species should be reflected with similar diversity in related species. Here, we have cloned the Class I α -1,2-mannosidase gene from *O. novo-ulmi*.

The *O. novo-ulmi* α -1,2-mannosidase gene encoded a 625 polypeptide which did not contain any introns in the reading frame of the gene. This was supported by the fact that the only possible insertion which could represent an intron, as determined by multiple sequence alignment, was also found in the mRNA and thus was not an intron sequence. Though the majority of genes in higher eukaryotes contain introns, only about two-thirds of genes sequenced in filamentous fungi contain introns, and even these tend to be quite short and sparse (Ballance, 1986). Indeed, of the small number of genes which have been sequenced in *O. novo-ulmi*, at least two do not contain any introns [*col1* (Pereira *et al.*, 2000), GenBank Accession #U35661); polygacturonase (GenBank Accession #AF052061)]. The fact that the *O. novo-ulmi* α -1,2-mannosidase gene does not contain any introns thus does not seem unusual. The *O. novo-ulmi* Class I α -1,2-mannosidase protein sequence contains a typical type-II transmembrane domain and, similar to other Class I α -1,2-mannosidases, is likely found in the ER and Golgi.

Sequence characterization of the Class I α -1,2-mannosidase isolated from *O. novo-ulmi* indicated that this gene showed relatively low overall similarity to the mammalian and insect Class I α -1,2-mannosidases, and much higher similarity to other fungal sequences. The dissimilarity of the fungal Class I α -1,2-mannosidases to other major lineages increases the potential of this region as a point of control. Obviously, any

drugs that could be developed as inhibitors of this region of the pathway need to be specific for the pathogen, without affecting the host pathway. By utilizing high-throughput screening of small molecule libraries, it may be possible to identify compounds which interact specifically with fungal Class I α -1,2-mannosidases which do not affect these enzymes in the host, allowing the development of antifungal drugs.

Assessment of possible functions of the Class I α -1,2-mannosidases were investigated by gene disruption. Loss of function of this gene product through gene disruption reduced the radial growth rate of the fungus and altered its morphology, although these effects were mild. Disruption mutants were fluffy with more aerial hyphae than wild type strains. This is in contrast with the result of *mns1* disruption in *S. cerevisiae*, which had no effect on the yeast growth or morphology (Camirand *et al.*, 1991). The distinct differences between the glycosylation pathways of *S. cerevisiae* and filamentous fungi may explain these contradictory results. The *mns1* gene removes a single mannose residue from $\text{Man}_9\text{GlcNAc}_2$ to produce $\text{Man}_8\text{GlcNAc}_2$, however, this step does not appear to be necessary for further elongation of the precursor to produce high mannose and hypermannosylated structures (Puccia *et al.*, 1993; Romero and Herscovics, 1989). This step of the yeast pathway appears to be involved in quality control to ensure proper protein folding in the ER, a function which cannot necessarily be attributed to the *O. novo-ulmi* Class I α -1,2-mannosidase. Synthesis of high mannose N-glycan structures in *S. cerevisiae* does not seem to rely on mannose removal, since the mannosyltransferase enzymes which add mannose to the N-glycan core can act equally on $\text{Man}_9\text{GlcNAc}_2$ or $\text{Man}_8\text{GlcNAc}_2$.

Similarity of the *O. novo-ulmi* α -1,2-mannosidase to the *A. satoi* and *P. citrinum* Class I α -1,2-mannosidases suggests that, similar to the *A. satoi* and *P. citrinum* enzymes, this enzyme is likely involved in removal of up to four mannose residues, which is also similar to the process found in higher eukaryotes. The presence in filamentous fungi of N-glycan structures which appear to have been modified by removal of mannose prior to outer chain elongation (Maras *et al.*, 1999) suggests that these enzymes are more crucial for production of N-glycan structures than in *S. cerevisiae*. Unlike *S. cerevisiae*, certain filamentous fungi contain multiple Class I α -1,2-mannosidases (Eades and Hintz, 2000b), suggesting diversification of this pathway in filamentous fungi. This increased diversification may be reflected in a diversification of function which illustrates the increased importance of the pathway in filamentous fungi for N-glycan function, which in turn increases the viability of this pathway as an antifungal target. Due to the relatedness and common ancestry of filamentous fungi, it is expected that the diversity of the Class I α -1,2-mannosidase gene family in *A. nidulans* would also be found in other filamentous fungi. That these genes have not yet been identified may be more a result of insufficient searching, rather than an absence.

That the disruption of the *O. novo-ulmi* Class I α -1,2-mannosidase gene did not radically impact the growth or morphology of the organism suggested that this may not present an optimal target for the development of antifungals. It must be considered, however, that slight alterations in plate morphology may actually be significant with respect to pathogenicity. Further analysis would be required to assess the effect of gene disruption on the structures of N-glycans produced. Also, the alteration of the morphology to a 'fluffy' phenotype may be a result of altered cell surface hydrophobicity, which may

affect the ability of the fungus to adhere. Such phenotypes have been associated with the ability of the fungus to form aerial hyphae, a trait which is highly dependent on cell surface hydrophobicity (Temple and Horgen, 2000; Temple *et al.*, 1997). The ultimate effects that altered N-glycan structures may have on the aforementioned pathogenicity factors which may be dependent on appropriate N-glycosylation would require pathogenicity trials on host elms.

A second consideration when considering the effects of disruption of the Class I α -1,2-mannosidase from *O. novo-ulmi* is the possibility of the existence of multiple genes with overlapping functions. Three such genes were found in *A. nidulans* (Eades and Hintz, 2000b), and multiple genes have been found in humans and mice (Bause *et al.*, 1993; Lal *et al.*, 1998; Tremblay *et al.*, 1998; Tremblay and Herscovics, 2000). Roberts *et al.* (1998) also found evidence for functional redundancy after disruption of the Class I α -1,2-mannosidase gene in *D. melanogaster*. It thus appears that multiple genes may be involved in mannose removal in filamentous fungi, with the effect of increasing the efficiency and fidelity of this pathway. This would explain the mild effects of disruption of this gene in *O. novo-ulmi*, and would suggest that multiple gene disruptions of all Class I α -1,2-mannosidases found in this organism may produce more significant effects. It will be necessary to fully characterize this gene family in *O. novo-ulmi*, possibly utilizing the information provided by the similar study in *A. nidulans* (Eades and Hintz, 2000b), in order to fully assess this region as a potential antifungal target.

We are undertaking a novel approach by investigating the general process of protein glycosylation, which can simultaneously affect a large number of pathogenicity determinants and thus the general ability of a pathogen to invade a host tissue (parasitic

fitness). Investigation of the glycosylation machinery in such host-pathogen systems will allow for a greater understanding of host-pathogen interactions and may provide targets for interrupting the flow of signals between the organisms, thus providing a method for control of fungal infectious diseases. Further characterization of the glycosylation machinery in the filamentous fungi may provide a better understanding of the molecular basis of the host-pathogen interaction and may illustrate key differences between the host and pathogen which may be exploited for control purposes.

CHAPTER 6 - General Discussion

The general objective of this research project was to characterize the α -mannosidase gene family in filamentous fungi. The α -mannosidase gene family is involved in the processing of N-glycans as they pass through the ER and Golgi by removal of mannose residues prior to further modification to produce mature N-glycan structures. Members of this gene family also play a role in the catabolism and recycling of oligosaccharides in the cytosol and lysosome. Although a great deal is known about these processes in mammals (Kornfeld and Kornfeld, 1985; Moremen *et al.*, 1994) and yeast (Herscovics, 1999), much less is known in filamentous fungi. This research was initiated in the filamentous fungus *A. nidulans* which is a useful host for the expression of heterologous proteins. To utilize this host for the production of human therapeutic proteins which require specific N-glycan structures for their efficacy, it would be necessary to manipulate the N-glycan processing pathways to produce the N-glycans found on the native proteins. Such manipulations require elucidation and modulation of the N-glycosylation machinery in filamentous fungi.

In order to recreate the human N-glycan processing pathways in fungi, it will be necessary to fully elucidate the similarities and differences between the pathways in these organisms. Although initial stages of the N-glycan pathways are very similar amongst all eukaryotes, the intermediate stages of N-glycan processing involving mannose removal appear to vary amongst major lineages. This divergence between major lineages is reflected in the types of N-glycans produced in different organisms, which can be quite different. Mannose removal in mammalian systems involves several α -mannosidases in the ER and Golgi apparatus. Although there is evidence that mannose removal in filamentous

fungi resembles that found in mammals (Maras *et al.*, 1999), the enzyme or enzymes involved in this process were not known. Characterization of the α -mannosidases in *A. nidulans* is a significant step in resolving this process in filamentous fungi.

When this project was initiated, little was known about the phylogenetic structure of this gene family. A reverse genetics approach was utilized to identify and fully sequence a novel α -mannosidase in *A. nidulans*, using the limited data available from published α -mannosidase sequences. As additional α -mannosidases were subsequently characterized in other organisms, it became clear that this gene family contained two major classes (Class I and Class II), which were evolutionarily distinct (Daniel *et al.*, 1994; Moremen *et al.*, 1994), and our own phylogenetic analysis of published α -mannosidase sequences verified this distinction (Figure 9; Chapter 2; Eades and Hintz, 1998). It was unclear at the time whether the somewhat anomalous rat ER/cytosolic α -mannosidase (Bischoff *et al.*, 1990) and yeast vacuolar α -mannosidase (Yoshihisa and Anraku, 1989) should be included in the Class II group, or should form a third distinct group. Phylogenetic analysis of the *A. nidulans* α -mannosidase revealed that it was a Class II α -mannosidase which was highly similar to the rat ER/cytosolic and yeast vacuolar α -mannosidases (Figure 9; Chapter 2; Eades and Hintz, 1998), thus solidifying the inclusion of these α -mannosidases as Class II enzymes. Our analysis further revealed that the Class II α -mannosidase genes could actually be classified as three distinct subgroups, Class IIA, IIB, and IIC, the latter containing the rat ER/cytosolic, yeast vacuolar and *A. nidulans* α -mannosidases.

The Class IIC α -mannosidases are an anomalous group of enzymes with diverse functions and cellular locations. The rat Class IIC α -mannosidase appears to be found in

both the ER and the cytosol and is likely involved in N-glycan catabolism (Bischoff *et al.*, 1990; Grard *et al.*, 1996; Haeuw *et al.*, 1991; Moore and Spiro, 1994; Tulsiani and Touster, 1987). The yeast Class IIC α -mannosidase, however, is localized to the vacuole (Yoshihisa and Anraku, 1989, 1990). The cellular localization of the *A. nidulans* Class IIC α -mannosidase was not conclusively identified, although it may be localized to a subcellular compartment. What is clear is that this enzyme, similar to the yeast Class IIC α -mannosidase lacks a typical signal sequence, indicating that if the enzyme is indeed localized to a subcellular compartment, it is not reaching its destination via normal secretion signal routing mechanisms. This has also been suggested for the targeting of the yeast Class IIC α -mannosidase vacuole.

Although it became apparent that the Class II α -mannosidase likely was not involved in N-glycan production, we were interested to determine the necessity of this gene for cellular function. Disruption of the gene did not appear to have any significant effects on the growth or morphology of *A. nidulans*, implying that this gene is not essential. Overexpression of the gene also did not have any significant cytotoxic effects. Thus, it will be possible to express this gene intracellularly or extracellularly for future work without harm to the organism.

Phylogenetic analysis of the α -mannosidase gene family (Eades *et al.*, 1998), combined with emerging research in the field led to the subsequent investigation of the Class I, rather than the Class II, α -mannosidases in *A. nidulans*. It became apparent that these genes were directly involved in the processing of N-glycans during their synthesis and would provide excellent candidates for the manipulation of this pathway for the production of specific N-glycan structures. A single Class I α -1,2-mannosidase, *mnsIA*,

was identified using degenerate PCR primers designed from multiple sequence alignments of newly available published Class 1 α -1,2-mannosidase amino acid sequences. Two additional Class 1 α -1,2-mannosidases, *mnsIB* and *mnsIC*, were subsequently identified in utilizing the *A. nidulans* Expressed Sequence Tag (EST) database and these three α -1,2-mannosidase genes were cloned and fully sequence characterized (Eades and Hintz, 2000b; Chapter 3).

Single members of the Class I α -mannosidase genes were found in *A. satoi* (Inoue *et al.*, 1995), *P. citrinum* (Yoshida and Ichishima, 1995), and *T. reesei* (Maras *et al.*, 2000). Purification and biochemical characterization of the *A. satoi* and *P. citrinum* α -1,2-mannosidase showed that these enzymes were able to catalyze the stepwise removal of mannose from $\text{Man}_9\text{GlcNAc}_2$, indicative of their putative role in N-glycan processing. This is consistent with the fact that filamentous fungi produce N-glycans as small as $\text{Man}_5\text{GlcNAc}_2$, indicating that four mannose are removed from N-glycans during processing in these organisms. The process of mannose removal in the N-glycan pathways of filamentous fungi thus appears to be more similar to mammalian systems than to yeast systems (Figure 20). Filamentous fungi produce mainly oligomannosidic N-glycans, likely due to the action of specific mannosyltransferase enzymes. The removal of mannose by α -1,2-mannosidases prior to the addition of mannose by mannosyltransferases could be explain by the necessity for $\text{Man}_{5-9}\text{GlcNAc}_2$ intermediates for the proper folding, targeting and secretion of glycoproteins in fungi. The similarity of the mannose removal process in filamentous fungi and mammals indicates that the associated processes of glycoprotein folding, targeting and secretion may also be similar in these organisms.

The identification of three Class I α -1,2-mannosidases in *A. nidulans* was the first report of multiple Class I α -1,2-mannosidases in a single fungal species. This reveals further similarity to mammalian N-glycan processing pathways, which utilize multiple Class I α -1,2-mannosidases to process $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$. The human IA, IB, and IC α -1,2-mannosidases are all capable of reducing $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$, however, each has slightly differing specificities (Tremblay *et al.*, 1998; Tremblay and Herscovics, 2000). Human α -1,2-mannosidase IC, for instance, rapidly converts $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_6\text{GlcNAc}_2$, but slowly converts that structure into $\text{Man}_5\text{GlcNAc}_2$. This contrasts with human α -1,2-mannosidase IB which rapidly converts $\text{Man}_8\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$. The diversification of the substrate specificities of these enzymes, and the products produced has also diversified the N-glycan processing pathway in mammals. This diversification may merely increase the efficiency and fidelity of the pathway, or alternatively, these intermediate structures might be utilized as lectin binding molecules for glycoprotein targeting or modified to produce novel N-glycan structures.

To determine the exact role of multiple Class I α -1,2-mannosidases in filamentous fungi requires purification and biochemical characterization of these enzymes. Multiple sequence alignments and phylogenetic analysis revealed that α -1,2-mannosidase IB amino acid sequence was highly related to the Class I α -1,2-mannosidases from *A. satoi* and *P. citrinum*, thus the *mnsIB* gene is likely an orthologue of the Class I α -1,2-mannosidase genes identified in these organisms (Eades and Hintz, 2000b). The *A. satoi* and *P. citrinum* Class I α -1,2-mannosidases have been previously characterized (Ichishima *et al.*, 1999; Yoshida *et al.*, 1998), and it is believed that the biochemical properties of the α -1,2-

mannosidase IB would be relatively similar. The α -1,2-mannosidase IA and α -1,2-mannosidase IC amino acid sequences were more distantly related. Biochemical characterization of the Class I α -1,2-mannosidases from *A. nidulans* was thus initiated with the α -1,2-mannosidase IC (Chapter 4). This member of the gene family was chosen since it is less similar to the *A. satoi* and *P. citrinum* Class I α -1,2-mannosidases which have been previously purified and characterized.

Characterization of the properties of the α -1,2-mannosidase IC was a natural first step into elucidation of the roles of these enzymes in N-glycan processing. The N-terminal domain of the α -1,2-mannosidase IC gene was replaced with a synthetic signal sequence and expressed in the extracellular media by an inducible promoter. Assays confirmed the presence of high levels of α -1,2-mannosidase activity which was absent from wild-type cells. Once this α -1,2-mannosidase is purified, it will be a relatively simple matter to characterize the substrate specificity, reaction optima, and kinetic parameters of this enzyme. The *A. satoi* and *P. citrinum* Class I α -1,2-mannosidases are both able to reduce the oligosaccharide precursor $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$, but have reduced activity on certain intermediate structures, leading to the accumulation of these structures in the reaction (Ichishima *et al.*, 1999; Yoshida *et al.*, 1998). This suggests a possible role for another Class I α -1,2-mannosidase, such as the α -1,2-mannosidase IC, with higher specificity for these intermediate structures.

Although the primary objective of this research project was the characterization of the α -mannosidase gene family in *A. nidulans* for the purposes of N-glycan remodeling, several avenues were explored for the utilization of this information for

studying the pathogenesis of the Dutch Elm Disease pathogen *O. novo-ulmi*. Impaired glycosylation may affect numerous processes which are necessary for growth and virulence of this organism, hence inhibition of the pathway would be detrimental to the 'pathogenic fitness' of the organism, thus the glycosylation machinery of this organism might provide an ideal antifungal target. We initiated this project with the identification of a Class I α -1,2-mannosidase gene from *O. novo-ulmi* (Chapter 5). This gene was identified utilizing the same degenerate PCR primers used for the identification of the *mnsIA* gene from *A. nidulans*, and was cloned and fully sequenced. The gene was subsequently disrupted to determine its potential as an antifungal target. Disruption of this gene reduced the radial growth rate and changed the morphology of the organism. The 'fluffy' phenotype of the mutant strains may reflect a change in the cell surface hydrophobicity due to alterations in N-glycan structures exhibited on the cell surface. The role of N-glycans in determining cell surface hydrophobicity has been well established in *C. albicans* (Hazen and Glee, 1994; Masuoka and Hazen, 1997, 1999). It has also been observed that changes in cell surface hydrophobicity due to alterations in the production of the hydrophobin cerato-ulmin (Temple and Horgen, 2000; Temple *et al.*, 1997) can change the fluffiness of the organism in a similar manner to the observed phenotype of the α -1,2-mannosidase disruption mutants (*unpublished results*). The disruption of the *O. novo-ulmi* α -1,2-mannosidase gene, although non-lethal, may have potential as a target for the inhibition of fungal pathogenesis. Alterations in the cell-surface and secreted N-glycans may exert numerous effects on fungal pathogens, reducing their 'pathogenic fitness'.

Three Class I α -1,2-mannosidase genes were thus identified in the filamentous fungus *A. nidulans* (Eades and Hintz, 2000b; Chapter 3), and one Class I α -mannosidase

has so far been identified in another filamentous fungus, *O. novo-ulmi* (Chapter 5). These two fungal species are phylogenetically divergent within the Ascomycetes - *A. nidulans* belongs to the class *Plectomycetes*, while *O. novo-ulmi* belongs to the other major class within the Ascomycetes, *Pyrenomycetes* (Berbee and Taylor, 1992a, 1992b, 1993). It is thus likely that other filamentous fungi also contain multiple members of this gene family and that this is the rule rather than the exception (Eades and Hintz, 2000a). The presence of multiple Class I α -mannosidases in filamentous fungi is a reflection of the similarities which exist between the glycosylation pathways of these organisms with the pathways of higher eukaryotes (Figure 20). Mannose processing during N-glycan processing in filamentous fungi is achieved by multiple Class I α -mannosidases which may have partially overlapping or redundant functions, and disruption of specific members may have little physiological effect. Such redundancies are found in other lineages, as well. Roberts *et al.* (1998) analyzed N-glycans produced by *mas-1* null mutants of *Drosophila melanogaster* and found that the N-glycans structures were not significantly changed, but that the production of these N-glycans was not as efficient as in the wild-type. A reduction in efficiency of protein glycosylation may reduce the overall fitness of the organism and may favor the maintenance of redundancy in this pathway.

Phylogenetic analyses of the Class I α -mannosidases have produced rather complicated results (Eades *et al.*, 1998; Eades and Hintz, 2000a, 2000b; Moremen *et al.*, 1994; Roberts *et al.*, 1998), but a pattern is beginning to emerge to explain the molecular evolution of this gene family (Figure 19). The mouse and human species each contain two Class I α -mannosidases, which are highly related and arose from a recent duplication event. This duplication occurred after the divergence of this lineage from insect and fungal

lineages. A third highly related Class I α -mannosidase is also found in humans, although the mouse orthologue has not been reported. The multiple Class I α -mannosidases found in *A. nidulans* and *O. novo-ulmi* also appear to have arisen from recent duplication events, in that these proteins are more similar to each other than to either of the mammalian Class I α -mannosidases. This is significant, because it suggests that the Class I α -mannosidase gene families in both lineages have independently undergone expansion and evolution through gene duplication and divergence. Characterization of the biochemical properties of the various α -mannosidases in *A. nidulans* and *O. novo-ulmi* will provide further insight into the purpose of such duplication and divergence.

The second family of mannosidases, the Class II α -mannosidases are more diverse in their biochemical properties and physiological functions (Daniel *et al.*, 1994). This group of enzymes consists of three subfamilies of genes (Classes IIA, IIB and IIC) with distinct cellular functions. The Class IIA subfamily is involved in N-glycan synthesis in the Golgi, while the other Class IIB and Class IIC are involved in N-glycan breakdown, removal and recycling in the cytoplasm, lysosome and vacuole. The first subfamily of Class II genes (Class IIA) are responsible for removal of α -1,3- and α -1,6-linked mannose residues from N-glycans during their synthesis, a process which occurs in the higher eukaryotes, but does not occur in lower eukaryotes, such as fungi. The second subfamily (Class IIB) are also found only in higher eukaryotes. These enzymes are involved in N-glycan degradation in the lysosome. The third subfamily (Class IIC) contains a more heterogeneous set of enzymes, with a diversity of functions and cellular localizations. Members of this family are found in higher and lower eukaryotes and are likely involved in many aspects of N-glycan degradation and recycling. Sequence analysis clearly resolves

the various inter-relationships of these proteins (Eades *et al.*, 1998; Chapter 2). The Class IIC subfamily has very low sequence similarity to the other two subfamilies. Phylogenetic analysis of the sequences shows that the Class IIA and Class IIB subfamilies have diverged more recently than the Class IIC subfamily (Figure 9). A likely scenario is that a single common ancestor was duplicated after the divergence of lower eukaryotes, such as fungi, from the higher eukaryotes. Thus, the lower eukaryotes only contain the orthologue of the common ancestor. Subsequent duplication in the higher eukaryotes led to the formation of the three subfamilies of Class II genes found in higher eukaryotes. These gene sequences diverged and evolved more specialized functions, such as the more complex N-glycan pathways (Class IIA), and more efficient degradation pathways (Class IIB).

It is intriguing that the two classes of α -mannosidases have such similar and overlapping functions. The Class I genes and the Class IIA genes have complementary functions in the N-glycan synthesis pathway of higher eukaryotes. The other Class II genes have broad substrate specificities and are able to cleave α -1,2 (as well as α -1,3, and α -1,6) mannose linkages, a property it shares with the Class I genes. The Class I and Class II genes show no sequence similarity and appear to have originated independently and represent a classic case of convergent evolution.

The present research represents a significant step in the characterization of the α -mannosidase gene family in filamentous fungi. The glycosylation pathways of these organisms appear to resemble mammalian systems, in which several mannose residues are removed from the $\text{Man}_3\text{GlcNAc}_2$ precursor molecule prior to further modification (Kornfeld and Kornfeld, 1985), rather than yeast, in which only a single mannose residue

is removed prior to modification (Herscovics, 1999). This suggests that the yeast glycosylation pathway serves a more primary function in the production of cell wall mannoproteins, and is indicative of a more ancestral state of this pathway. Filamentous fungi, through duplication and divergence of the α -mannosidase gene family, have a more complex pathway which may be utilized for more diverse functions. The presence of terminal sugar residues such as galactose and sialic acid in some filamentous fungi (Alviano *et al.*, 1999; Maras *et al.*, 1999) is also indicative of the evolution of this pathway towards more complex function. It is believed that the diversification of the terminal branches of N-glycans throughout evolution, by the addition of novel glycosyltransferase enzymes to the pathway, was a major driving force in the evolution of novel functions for the N-glycans (Drickamer and Taylor, 1998; Gagneux and Varki, 1999).

The eventual goal of this research is the development of expression systems in filamentous fungi that produce glycoproteins, containing native N-glycan structures, for human therapeutic use. The production of complex type N-glycans requires the complete removal of α -1,2-mannose from $\text{Man}_9\text{GlcNAc}_2$ to produce $\text{Man}_3\text{GlcNAc}_2$, the substrate for the addition of GlcNAc by Gnt I, the first committed step in the production of complex N-glycans (Hintz *et al.*, 1995; Kalsner *et al.*, 1995). The identification of three Class I α -1,2-mannosidases in *A. nidulans* represents a significant step towards this goal. Controlled overexpression of these genes in Gnt I expressing strains should produce a steady supply of $\text{Man}_3\text{GlcNAc}_2$ on which Gnt I can act. A second goal of this research is the development of novel antifungal targets with which to control fungal pathogens. The Class I α -1,2-mannosidases may provide such a target. It will be necessary to identify all the

genes in this family from *O. novo-ulmi* to determine if disruption of this region of the pathway will have significant detrimental effects on the growth or pathogenicity of the organism.

7. LITERATURE CITED

- Alviano, C. S., L. R. Travassos and R. Schauer (1999). Sialic acids in fungi: a minireview. *Glycoconjugate Journal* **16**: 545-54.
- Andersen, D. C. and C. F. Goochee (1994). The effect of cell-culture conditions on the oligosaccharide structures of secreted glycoproteins. *Current Opinion in Biotechnology* **5**: 546-9.
- Appenzeller, C., H. Andersson, F. Kappeler and H. P. Hauri (1999). The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nature Cell Biology* **1**: 330-4.
- Archer, D. B., D. J. Jeenes and D. A. Mackenzie (1994). Strategies for improving heterologous protein production from filamentous fungi. *Antonie Van Leeuwenhoek* **65**: 245-50.
- Archer, D. B. and J. F. Peberdy (1997). The molecular biology of secreted enzyme production by fungi. *Critical Reviews in Biotechnology* **17**: 273-306.
- Ballance, D. J. (1986). Sequences important for gene expression in filamentous fungi. *Yeast* **2**: 229-236.
- Ballance, D. J. (1991). Transformation systems for filamentous fungi and an overview of fungal gene structure. *in* Molecular Industrial Mycology: Systems and Applications for Filamentous Fungi. R. Berka and S. Leong. New York, Marcel Dekker, Inc.
- Basse, C. W. and T. Boller (1992). Glycopeptide elicitors of stress responses in tomato cells: N-linked glycans are essential for activity but act as suppressors of the same activity when released from the glycopeptides. *Plant Physiology* **98**: 1239-1247.
- Basse, C. W., A. Fath and T. Boller (1993). High affinity binding of a glycopeptide elicitor to tomato cells and microsomal membranes and displacement by specific glycan suppressors. *Journal of Biological Chemistry* **268**: 14724-31.
- Bause, E., E. Bieberich, A. Rolfs, C. Volker and B. Schmidt (1993). Molecular cloning and primary structure of Man9-mannosidase from human kidney. *European Journal of Biochemistry* **217**: 535-40.
- Berbee, M. L. and J. W. Taylor (1992a). 18S ribosomal RNA gene sequence characters place the human pathogen *Sporothrix schenckii* in the genus *Ophiostoma*. *Experimental Mycology* **16**: 87-91.
- Berbee, M. L. and J. W. Taylor (1992b). Two ascomycete classes based on fruiting-body characters and ribosomal DNA sequence. *Molecular Biology and Evolution* **9**: 278-84.

- Berbee, M. L. and J. W. Taylor (1993). Dating the evolutionary radiations of the true fungi. *Canadian Journal of Botany* **71**: 1114-1127.
- Bieberich, E., K. Tremel, C. Volker, A. Rolfs, B. Kalz-Fuller and E. Bause (1997). Man9-mannosidase from pig liver is a type-II membrane protein that resides in the endoplasmic reticulum. cDNA cloning and expression of the enzyme in COS 1 cells. *European Journal of Biochemistry* **246**: 681-9.
- Bircher, U. and H. R. Hohl (1997). Surface glycoproteins associated with appressorium formation and adhesion in *Phytophthora palmivora*. *Mycological Research* **101**: 769-775.
- Bischoff, J., K. Moremen and H. F. Lodish (1990). Isolation, characterization, and expression of cDNA encoding a rat liver endoplasmic reticulum alpha-mannosidase. *Journal of Biological Chemistry* **265**: 17110-7.
- Bowden, C. G., W. E. Hintz, R. Jeng, M. Hubbes and P. A. Horgen (1994). Isolation and characterization of the cerato-ulmin toxin gene of the Dutch elm disease pathogen, *Ophiostoma ulmi*. *Current Genetics* **25**: 323-9.
- Bradford, W. W. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**: 248-254.
- Brasier, C. M. (1991). *Ophiostoma novo-ulmi* sp. nov. causative agent of the current Dutch elm disease pandemics. *Mycopathologia* **115**: 151-161.
- Broekhuijsen, M. P., I. E. Mattern, R. Contreras, J. R. Kinghorn and C. A. van den Hondel (1993). Secretion of heterologous proteins by *Aspergillus niger*: production of active human interleukin-6 in a protease-deficient mutant by KEX2-like processing of a glucoamylase-hIL6 fusion protein. *Journal of Biotechnology* **31**: 135-45.
- Brul, S., A. King, J. M. van der Vaart, J. Chapman, F. Klis and C. T. Verrips (1997). The incorporation of mannoproteins in the cell wall of *S. cerevisiae* and filamentous Ascomycetes. *Antonie Van Leeuwenhoek* **72**: 229-37.
- Buckholz, R. G. and M. A. Gleeson (1991). Yeast systems for the commercial production of heterologous proteins. *BioTechnology* **9**: 1067-72.
- Buurman, E. T., C. Westwater, B. Hube, A. J. Brown, F. C. Odds and N. A. Gow (1998). Molecular analysis of CaMnt1p, a mannosyl transferase important for adhesion and virulence of *Candida albicans*. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 7670-7675.

- Calmels, T. P., F. Martin, H. Durand and G. Tiraby (1991). Proteolytic events in the processing of secreted proteins in fungi. *Journal of Biotechnology* **17**: 51-66.
- Camirand, A., A. Heysen, B. Grondin and A. Herscovics (1991). Glycoprotein biosynthesis in *Saccharomyces cerevisiae*. Isolation and characterization of the gene encoding a specific processing alpha- mannosidase. *Journal of Biological Chemistry* **266**: 15120-7.
- Campbell, R. M., M. Metzler, M. Granovsky, J. W. Dennis and J. D. Marth (1995). Complex asparagine-linked oligosaccharides in Mgat1-null embryos. *Glycobiology* **5**: 535-43.
- Caplan, S., R. Green, J. Rocco and J. Kurjan (1991). Glycosylation and structure of the yeast MF alpha 1 alpha-factor precursor is important for efficient transport through the secretory pathway. *Journal of Bacteriology* **173**: 627-35.
- Cenis, J. L. (1992). Rapid extraction of fungal DNA for PCR amplification. *Nucleic Acids Research* **20**: 2380.
- Cereghino, J. L. and J. M. Cregg (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology Reviews* **24**: 45-66.
- Chiba, Y., M. Suzuki, S. Yoshida, A. Yoshida, H. Ikenaga, M. Takeuchi, Y. Jigami and E. Ichishima (1998). Production of human compatible high mannose-type (Man5GlcNAc2) sugar chains in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **273**: 26298-304.
- Chiba, Y., Y. Yamagata, S. Iijima, T. Nakajima and E. Ichishima (1993). The carbohydrate moiety of the acid carboxypeptidase from *Aspergillus saitoi*. *Current Microbiology* **27**: 281-8.
- Chui, D., M. Oh-Eda, Y. F. Liao, K. Panneerselvam, A. Lal, K. W. Marek, H. H. Freeze, K. W. Moremen, M. N. Fukuda and J. D. Marth (1997). Alpha-mannosidase-II deficiency results in dyserythropoiesis and unveils an alternate pathway in oligosaccharide biosynthesis. *Cell* **90**: 157-67.
- Clark, A. G. (1994). Invasion and maintenance of a gene duplication. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 2950-4.
- Contreras, R., D. Carrez, J. R. Kinghorn, C. A. van den Hondel and W. Fiers (1991). Efficient KEX2-like processing of a glucoamylase-interleukin-6 fusion protein by *Aspergillus nidulans* and secretion of mature interleukin-6. *Bio/Technology* **9**: 378-81.

- Crout, J. E., N. S. Brewer and R. B. Tompkins (1977). Sporotrichosis arthritis: clinical features in seven patients. *Annals of Internal Medicine* **86**: 294-7.
- Cumming, D. A. (1991). Glycosylation of recombinant protein therapeutics: control and functional implications. *Glycobiology* **1**: 115-30.
- Daniel, P. F., B. Winchester and C. D. Warren (1994). Mammalian alpha-mannosidases--multiple forms but a common purpose? *Glycobiology* **4**: 551-66.
- Davies, R. W. (1991). Molecular biology of a high-level recombinant protein production system in *Aspergillus*. in *Molecular Industrial Mycology: Systems and Applications for Filamentous Fungi*. S. Leong and R. Berka. New York, Marcel Dekker, Inc.: 45-81.
- de Wit, P. J., M. H. Joosten, G. Honee, J. P. Wubben, G. F. van den Ackerveken and H. W. van den Broek (1994). Molecular communication between host plant and the fungal tomato pathogen *Cladosporium fulvum*. *Antonie Van Leeuwenhoek* **65**: 257-62.
- Dean, J. F. D. and J. D. Anderson (1991). Ethylene biosynthesis-inducing xylanase: purification and physical characterization of the enzyme produced by *Trichoderma viride*. *Plant Physiology* **95**: 316-323.
- Dean, N. (1995). Yeast glycosylation mutants are sensitive to aminoglycosides. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 1287-91.
- Dean, N. (1999). Asparagine-linked glycosylation in the yeast Golgi. *Biochimica et Biophysica Acta* **1426**: 309-22.
- Delorme, E., T. Lorenzini, J. Giffin, F. Martin, F. Jacobsen, T. Boone and S. Elliott (1992). Role of glycosylation on the secretion and biological activity of erythropoietin. *Biochemistry* **31**: 9871-6.
- Depiereux, E. and E. Feytmans (1992). MATCH-BOX: a fundamentally new algorithm for the simultaneous alignment of several protein sequences. *Computer Applications in the Biosciences* **8**: 501-9.
- Di Pietro, A. and M. I. G. Roncero (1996). Endopolygalacturonase from *Fusarium oxysporum f. sp. lycopersici*: Purification, characterization, and production during infection of tomato plants. *Phytopathology* **86**: 1324-30.
- Doolittle, R. F. (1995). The multiplicity of domains in proteins. *Annual Review of Biochemistry* **64**: 287-314.

- Dordal, M. S., F. F. Wang and E. Goldwasser (1985). The role of carbohydrate in erythropoietin action. *Endocrinology* **116**: 2293-9.
- Drickamer, K. (1991). Clearing up glycoprotein hormones. *Cell* **67**: 1029-32.
- Drickamer, K. and M. E. Taylor (1998). Evolving views of protein glycosylation. *Trends in Biochemical Sciences* **23**: 321-4.
- Dube, S., J. W. Fisher and J. S. Powell (1988). Glycosylation at specific sites of erythropoietin is essential for biosynthesis, secretion, and biological function. *Journal of Biological Chemistry* **263**: 17516-21.
- Dunn-Coleman, N. S., P. Bloebaum, R. Berka, E. Bodie, N. Robinson, G. Armstrong, M. Ward, M. Przetak, G. L. Carter, R. LaCost, L. J. Wilson, B. Felenbok, D. Sequeval, M. Mathieu, S. Sibley, D. I. Gwynne and R. W. Davies (1988). The ethanol regulon in *Aspergillus nidulans*: characterization and sequence of the positive regulatory gene *alcR*. *Gene* **73**: 385-396.
- Dwek, R. A. (1995). Glycobiology: more functions for oligosaccharides. *Science* **269**: 1234-5.
- Dwek, R. A. (1998). Biological importance of glycosylation. *Developments in Biological Standardization* **96**: 43-47.
- Eades, C. J., A. M. Gilbert, C. D. Goodman and W. E. Hintz (1998). Identification and analysis of a class 2 alpha-mannosidase from *Aspergillus nidulans*. *Glycobiology* **8**: 17-33.
- Eades, C. J. and W. E. Hintz (2000a). Characterization of the α -mannosidase gene family in filamentous fungi: N-glycan remodelling for the development of eukaryotic expression systems. *Biotechnology and Bioprocess Engineering* **5**: 227-233.
- Eades, C. J. and W. E. Hintz (2000b). Characterization of the class I α -mannosidase gene family in the filamentous fungus *Aspergillus nidulans*. *Gene* **255**: 25-34.
- Elbein, A. D. (1988). Glycoprotein processing and glycoprotein processing inhibitors. *Plant Physiology* **87**: 291-295.
- Fernandes, K. S., H. L. Mathews and L. M. Lopes Bezerra (1999). Differences in virulence of *Sporothrix schenckii* conidia related to culture conditions and cell-wall components. *Journal of Medical Microbiology* **48**: 195-203.
- Fiedler, K. and K. Simons (1995). The role of N-glycans in the secretory pathway. *Cell* **81**: 309-12.

- Fiedler, K. and K. Simons (1996). Characterization of VIP36, an animal lectin homologous to leguminous lectins. *Journal of Cell Science* **109**: 271-6.
- Fincham, J. R. S. (1989). Transformation in fungi. *Microbiological Reviews* **53**: 148-170.
- Foster, J. M., B. Yudkin, A. E. Lockyer and D. B. Roberts (1995). Cloning and sequence analysis of GmII, a *Drosophila melanogaster* homologue of the cDNA encoding murine Golgi alpha-mannosidase II. *Gene* **154**: 183-6.
- Fukuda, M. (1994). Cell surface carbohydrates: cell type specific expression. in *Molecular Glycobiology*. M. Fukuda and O. Hindsgaul. New York, Oxford University Press.
- Fussenegger, M., J. E. Bailey, H. Hauser and P. P. Mueller (1999). Genetic optimization of recombinant glycoprotein production by mammalian cells. *Trends in Biotechnology* **17**: 35-42.
- Gagneux, P. and A. Varki (1999). Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **9**: 747-55.
- Geisow, M. J. (1992). Glycoprotein glycans--roles and controls. *Trends in Biotechnology* **10**: 333-5.
- Gellissen, G., K. Melber, Z. A. Janowicz, U. M. Dahlems, U. Weydemann, M. Piontek, A. W. Strasser and C. P. Hollenberg (1992). Heterologous protein production in yeast. *Antonie Van Leeuwenhoek* **62**: 79-93.
- Goochee, C. F., M. J. Gramer, D. C. Andersen, J. B. Bahr and J. R. Rasmussen (1991). The oligosaccharides of glycoproteins: bioprocess factors affecting oligosaccharide structure and their effect on glycoprotein properties. *Bio/Technology* **9**: 1347-55.
- Gouka, R. J., P. J. Punt and C. A. van den Hondel (1997). Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Applied Microbiology and Biotechnology* **47**: 1-11.
- Gow, N. A., S. Bates, A. J. Brown, E. T. Buurman, L. M. Thomson and C. Westwater (1999). *Candida* cell wall mannosylation: importance in host-fungus interaction and potential as a target for the development of antifungal drugs. *Biochemical Society Transactions* **27**: 512-6.
- Grabenhorst, E., A. Hoffmann, M. Nimtz, G. Zettlmeissl and H. S. Conradt (1995). Construction of stable BHK-21 cells coexpressing human secretory glycoproteins and human Gal(beta 1-4)GlcNAc-R alpha 2,6- sialyltransferase alpha 2,6-linked NeuAc is preferentially attached to the Gal(beta 1-4)GlcNAc(beta 1-2)Man(alpha 1-3)-branch of diantennary oligosaccharides from secreted recombinant beta-trace protein. *European Journal of Biochemistry* **232**: 718-25.

- Grard, T., V. Herman, A. Saint-Pol, D. Kmiecik, O. Labiau, A. M. Mir, C. Alonso, A. Verbert, R. Cacan and J. C. Michalski (1996). Oligomannosides or oligosaccharide-lipids as potential substrates for rat liver cytosolic alpha-D-mannosidase. *Biochemical Journal* **316**: 787-92.
- Grinnell, B. W., J. D. Walls and B. Gerlitz (1991). Glycosylation of human protein C affects its secretion, processing, functional activities, and activation by thrombin. *Journal of Biological Chemistry* **266**: 9778-85.
- Gurr, S. J., S. E. Unkles and J. R. Kinghorn (1987). The structure and organization of nuclear genes of filamentous fungi. *in* Gene Structure in Eukaryotic Microbes. J. R. Kinghorn. Oxford, WA, IRL Press: 93-139.
- Gwynne, D. I., F. P. Buxton, S. A. Williams, S. Garvin and R. W. Davies (1987). Genetically engineered secretion of active human interferon and a bacterial endoglucanase from *Aspergillus nidulans*. *Bio/Technology* **5**: 713-717.
- Gwynne, D. I. and M. Devchand (1992). Expression of foreign proteins in the genus *Aspergillus*. *Biotechnology* **23**: 203-14.
- Haeuw, J. F., G. Strecker, J. M. Wieruszkeski, J. Montreuil and J. C. Michalski (1991). Substrate specificity of rat liver cytosolic alpha-D-mannosidase. Novel degradative pathway for oligomannosidic type glycans. *European Journal of Biochemistry* **202**: 1257-68.
- Hammond, C., I. Braakman and A. Helenius (1994). Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 913-7.
- Hammond, C. and A. Helenius (1994). Quality control in the secretory pathway: retention of a misfolded viral membrane glycoprotein involves cycling between the ER, intermediate compartment, and Golgi apparatus. *Journal of Cell Biology* **126**: 41-52.
- Hara-Kuge, S., T. Ohkura, A. Seko and K. Yamashita (1999). Vesicular-integral membrane protein, VIP36, recognizes high-mannose type glycans containing alpha1-->2 mannosyl residues in MDCK cells. *Glycobiology* **9**: 833-9.
- Hauri, H., C. Appenzeller, F. Kuhn and O. Nufer (2000). Lectins and traffic in the secretory pathway. *FEBS Letters* **476**: 32-7.
- Hazen, K. C. (1989). Participation of yeast cell surface hydrophobicity in adherence of *Candida albicans* to human epithelial cells. *Infection and Immunity* **57**: 1894-900.

- Hazen, K. C. and P. M. Glee (1994). Hydrophobic cell wall protein glycosylation by the pathogenic fungus *Candida albicans*. *Canadian Journal of Microbiology* **40**: 266-72.
- Hemming, F. W. (1995). Expression and secretion of glycoproteins by hyphal fungi. *Biochemical Society Transactions* **23**: 180-5.
- Henikoff, S. and J. G. Henikoff (1992). Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America* **89**: 10915-9.
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* **280**: 309-16.
- Henrissat, B. and A. Bairoch (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* **293**: 781-8.
- Henrissat, B. and A. Bairoch (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal* **316**: 695-6.
- Henrissat, B. and A. Romeu (1995). Families, superfamilies and subfamilies of glycosyl hydrolases. *Biochemical Journal* **311**: 350-1.
- Herscovics, A. (1999). Processing glycosidases of *Saccharomyces cerevisiae*. *Biochimica et Biophysica Acta* **1426**: 275-85.
- Herscovics, A. and P. Orlean (1993). Glycoprotein biosynthesis in yeast. *FASEB Journal* **7**: 540-50.
- Herscovics, A., J. Schneikert, A. Athanassiadis and K. W. Moremen (1994). Isolation of a mouse Golgi mannosidase cDNA, a member of a gene family conserved from yeast to mammals. *Journal of Biological Chemistry* **269**: 9864-71.
- Hickman, S. and S. Kornfeld (1978). Effect of tunicamycin on IgM, IgA, and IgG secretion by mouse plasmacytoma cells. *Journal of Immunology* **121**: 990-6.
- Hintz, W. E. (1999). Sequence analysis of the chitin synthase A gene of the Dutch elm pathogen *Ophiostoma novo-ulmi* indicates a close association with the human pathogen *Sporothrix schenckii*. *Gene* **237**: 215-21.
- Hintz, W. E., I. Kalsner, E. Plawinski, Z. Guo and P. A. Lagosky (1995). Improved gene expression in *Aspergillus nidulans*. *Canadian Journal of Botany* **73**(Suppl. 1): S876-S884.

- Hintz, W. E. and P. A. Lagosky (1993). A glucose-derepressed promoter for expression of heterologous products in the filamentous fungus *Aspergillus nidulans*. *Biotechnology* **11**: 815-8.
- Hollenstein, L., S. Balsiger and H. R. Hohl (1995). The influence of IgG, proteases, and glycosidases on adhesion to and infection of soybean plants by *Phytophthora megasperma f. sp. glycinea*. *Botanica Helvetica* **105**: 221-232.
- Huang, X. (1994). On global sequence alignment. *Computer Applications in the Biosciences* **10**: 227-35.
- Hughes, A. L. (1994). The evolution of functionally novel proteins after gene duplication. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **256**: 119-24.
- Hughes, M. K. and A. L. Hughes (1993). Evolution of duplicate genes in a tetraploid animal, *Xenopus laevis*. *Molecular Biology and Evolution* **10**: 1360-9.
- Huynen, M. A. and E. van Nimwegen (1998). The frequency distribution of gene family sizes in complete genomes. *Molecular Biology and Evolution* **15**: 583-9.
- Ichishima, E., M. Arai, Y. Shigematsu, H. Kumagai and R. Sumida-Tanaka (1981). Purification of an acidic α -D-mannosidase from *Aspergillus saitoi* and specific cleavage of 1,2- α -D-mannosidic linkage in yeast mannan. *Biochimica et Biophysica Acta* **658**: 45-53.
- Ichishima, E., N. Taya, M. Ikeguchi, Y. Chiba, M. Nakamura, C. Kawabata, T. Inoue, K. Takahashi, T. Minetoki, K. Ozeki, C. Kumagai, K. Gomi, T. Yoshida and T. Nakajima (1999). Molecular and enzymic properties of recombinant 1,2- α -mannosidase from *Aspergillus saitoi* overexpressed in *Aspergillus oryzae* cells. *Biochemical Journal* **339**: 589-97.
- Imperiali, B. and K. W. Rickert (1995). Conformational implications of asparagine-linked glycosylation. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 97-101.
- Inoue, T., T. Yoshida and E. Ichishima (1995). Molecular cloning and nucleotide sequence of the 1,2- α -D-mannosidase gene, *msdS*, from *Aspergillus saitoi* and expression of the gene in yeast cells. *Biochimica et Biophysica Acta* **1253**: 141-5.
- Jarvis, D. L. and E. E. Finn (1996). Modifying the insect cell N-glycosylation pathway with immediate early baculovirus expression vectors. *Nature Biotechnology* **14**: 1288-92.

- Jenkins, N., R. B. Parekh and D. C. James (1996). Getting the glycosylation right: implications for the biotechnology industry. *Nature Biotechnology* **14**: 975-81.
- Kalsner, I., W. Hintz, L. S. Reid and H. Schachter (1995). Insertion into *Aspergillus nidulans* of functional UDP-GlcNAc: α -3-D-mannoside- β -1,2-N-acetylglucosaminyl-transferase I, the enzyme catalysing the first committed step from oligomannose to hybrid and complex N-glycans. *Glycoconjugate Journal* **12**: 360-70.
- Karlsson, K. A. (1991). Glycobiology: a growing field for drug design. *Trends in Pharmacological Sciences* **12**: 265-72.
- Kawar, Z., A. Herscovics and D. L. Jarvis (1997). Isolation and characterization of an alpha 1,2-mannosidase cDNA from the lepidopteran insect cell line Sf9. *Glycobiology* **7**: 433-43.
- Keller, N. P. and T. M. Hohn (1997). Metabolic pathway gene clusters in filamentous fungi. *Fungal Genetics and Biology* **21**: 17-29.
- Kerscher, S., S. Albert, D. Wucherpfennig, M. Heisenberg and S. Schneuwly (1995). Molecular and genetic analysis of the *Drosophila mas-1* (mannosidase-1) gene which encodes a glycoprotein processing alpha 1,2-mannosidase. *Developmental Biology* **168**: 613-26.
- Klis, F. M., L. H. Caro, J. H. Vossen, J. C. Kapteyn, A. F. Ram, R. C. Montijn, M. A. Van Berkel and H. Van den Ende (1997). Identification and characterization of a major building block in the cell wall of *Saccharomyces cerevisiae*. *Biochemical Society Transactions* **25**: 856-60.
- Knight, P. (1989). The carbohydrate frontier. *Bio/Technology* **7**: 35-40.
- Kobata, A. (1992). Structures and functions of the sugar chains of glycoproteins. *European Journal of Biochemistry* **209**: 483-501.
- Kornfeld, R. and S. Kornfeld (1985). Assembly of asparagine-linked oligosaccharides. *Annual Review of Biochemistry* **54**: 631-64.
- Kornfeld, S. (1987). Trafficking of lysosomal enzymes. *FASEB Journal* **1**: 462-8.
- Kornfeld, S. (1998). Diseases of abnormal protein glycosylation: an emerging area. *Journal of Clinical Investigation* **101**: 1293-1295.
- Lal, A., P. Pang, S. Kalelkar, P. A. Romero, A. Herscovics and K. W. Moremen (1998). Substrate specificities of recombinant murine Golgi alpha 1, 2- mannosidases IA

and IB and comparison with endoplasmic reticulum and Golgi processing alpha1,2-mannosidases. *Glycobiology* **8**: 981-95.

- Lal, A., J. S. Schutzbach, W. T. Forsee, P. J. Neame and K. W. Moremen (1994). Isolation and expression of murine and rabbit cDNAs encoding an alpha 1,2-mannosidase involved in the processing of asparagine-linked oligosaccharides. *Journal of Biological Chemistry* **269**: 9872-81.
- Li, Q. X., P. Hasty and B. C. Shur (1997). Targeted mutation in β 1,4-galactosyltransferase leads to pituitary insufficiency and neonatal lethality. *Developmental Biology* **181**: 257-267.
- Li, W.-H. (1997). Molecular Evolution. Sunderland, MA, Sinauer Associates, Inc.
- Lipari, F. and A. Herscovics (1996). Role of the cysteine residues in the α -1,2-mannosidase involved in N-glycan biosynthesis in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **271**: 27615-27622.
- Lis, H. and N. Sharon (1993). Protein glycosylation. Structural and functional aspects. *European Journal of Biochemistry* **218**: 1-27.
- Liu, D. T. (1992). Glycoprotein pharmaceuticals: scientific and regulatory considerations, and the US Orphan Drug Act. *Trends in Biotechnology* **10**: 114-20.
- MacKenzie, D. A., D. J. Jeenes, N. J. Belshaw and D. B. Archer (1993). Regulation of secreted protein production by filamentous fungi: recent developments and perspectives. *Journal of General Microbiology* **139**: 2295-307.
- Maras, M., N. Callewaert, K. Piens, M. Claeysens, W. Martinet, S. Dewaele, H. Contreras, I. Dewerte, M. Penttila and R. Contreras (2000). Molecular cloning and enzymatic characterization of a *Trichoderma reesei* 1,2-alpha-D-mannosidase. *Journal of Biotechnology* **77**: 255-63.
- Maras, M., A. De Bruyn, J. Schraml, P. Herdewijn, M. Claeysens, W. Fiers and R. Contreras (1997a). Structural characterization of N-linked oligosaccharides from cellobiohydrolase I secreted by the filamentous fungus *Trichoderma reesei* RUTC 30. *European Journal of Biochemistry* **245**: 617-25.
- Maras, M., X. Saelens, W. Laroy, K. Piens, M. Claeysens, W. Fiers and R. Contreras (1997b). In vitro conversion of the carbohydrate moiety of fungal glycoproteins to mammalian-type oligosaccharides--evidence for N- acetylglucosaminyltransferase-I-accepting glycans from *Trichoderma reesei*. *European Journal of Biochemistry* **249**: 701-7.

- Maras, M., I. van Die, R. Contreras and C. A. van den Hondel (1999). Filamentous fungi as production organisms for glycoproteins of bio- medical interest. *Glycoconjugate Journal* **16**: 99-107.
- Marth, J. D. (1994). Will the transgenic mouse serve as a Rosetta Stone to glycoconjugate function? *Glycoconjugate Journal* **11**: 3-8.
- Masuoka, J. and K. C. Hazen (1997). Cell wall protein mannosylation determines *Candida albicans* cell surface hydrophobicity. *Microbiology* **143**: 3015-21.
- Masuoka, J. and K. C. Hazen (1999). Differences in the acid-labile component of *Candida albicans* mannan from hydrophobic and hydrophilic yeast cells. *Glycobiology* **9**: 1281-6.
- Matta, K. L. and O. P. Bahl (1972). Glycosidases of *Aspergillus niger*. IV. Purification and characterization of α -mannosidase. *Journal of Biological Chemistry* **247**: 1780-7.
- Merkle, R. K. and K. Moremen (1993). Cloning of a murine homolog of the *Dictyostelium discoideum* lysosomal α -mannosidase. *Glycoconjugate Journal* **10**: 240.
- Meynial-Salles, I. and D. Combes (1996). In vitro glycosylation of proteins: an enzymatic approach. *Journal of Biotechnology* **46**: 1-14.
- Misago, M., Y. F. Liao, S. Kudo, S. Eto, M. G. Mattei, K. W. Moreman and M. N. Fukuda (1995). Molecular cloning and expression of cDNAs encoding human α -mannosidase II and a previously unrecognized α -mannosidase IIx isoenzyme. *Proceedings of the National Academy of Sciences of the United States of America* **92**.
- Moore, S. E. and R. G. Spiro (1994). Intracellular compartmentalization and degradation of free polymannose oligosaccharides released during glycoprotein biosynthesis. *Journal of Biological Chemistry* **269**: 12715-21.
- Moremen, K. W. and P. W. Robbins (1991). Isolation, characterization, and expression of cDNAs encoding murine α -mannosidase II, a Golgi enzyme that controls conversion of high mannose to complex N-glycans. *Journal of Cell Biology* **115**: 1521-34.
- Moremen, K. W., R. B. Trimble and A. Herscovics (1994). Glycosidases of the asparagine-linked oligosaccharide processing pathway. *Glycobiology* **4**: 113-25.
- Nadeau, J. H. and D. Sankoff (1997). Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. *Genetics* **147**: 1259-66.

- Nakei, K. and M. Kanehisa (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14**: 897-911.
- Nebes, V. L. and M. C. Schmidt (1994). Human lysosomal alpha-mannosidase: isolation and nucleotide sequence of the full-length cDNA. *Biochemical and Biophysical Research Communications* **200**: 239-45.
- Ohno, S. (1970). *Evolution by Gene Duplication*. New York, Springer Verlag.
- Ohta, T. (1989). Role of gene duplication in evolution. *Genome* **31**: 304-10.
- Ohta, T. (1990). How gene families evolve. *Theoretical Population Biology* **37**: 213-9.
- Ohta, T. (1994). Further examples of evolution by gene duplication revealed through DNA sequence comparisons. *Genetics* **138**: 1331-7.
- Opendakker, G., P. M. Rudd, C. P. Ponting and R. A. Dwek (1993). Concepts and principles of glycobiology. *FASEB Journal* **7**: 1330-7.
- Orlean, P. (1992). Enzymes that recognize dolichols participate in three glycosylation pathways and are required for protein secretion. *Biochemistry and Cell Biology* **70**: 438-47.
- Parekh, R. B. and T. P. Patel (1992). Comparing the glycosylation patterns of recombinant glycoproteins. *Trends in Biotechnology* **10**: 276-80.
- Parodi, A. J. (2000). Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. *Biochemical Journal* **348 Pt 1**: 1-13.
- Paulson, J. C. (1989). Glycoproteins: what are the sugar chains for? *Trends in Biochemical Sciences* **14**: 272-6.
- Peberdy, J. F. (1994). Protein secretion in filamentous fungi--trying to understand a highly productive black box. *Trends in Biotechnology* **12**: 50-7.
- Pereira, V., J. C. Royer, W. E. Hintz, D. Field, C. Bowden, K. Kokurewicz, M. Hubbes and P. A. Horgen (2000). A gene associated with filamentous growth in *Ophiostoma novo-ulmi* has RNA-binding motifs and is similar to a yeast gene involved in mRNA splicing. *Current Genetics* **37**: 94-103.
- Porwoll, S., N. Loch, C. Kannicht, R. Nuck, D. Grunow, W. Reutter and R. Tauber (1998). Cell surface glycoproteins undergo postbiosynthetic modification of their

- N-glycans by stepwise demannosylation. *Journal of Biological Chemistry* **273**: 1075-85.
- Puccia, R., B. Grondin and A. Herscovics (1993). Disruption of the processing α -mannosidase gene does not prevent outer chain synthesis in *Saccharomyces cerevisiae*. *Biochemical Journal* **290**: 21-6.
- Punt, P. J., R. P. Oliver, M. A. Dingemans, P. H. Pouwels and C. A. van den Hondel (1987). Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. *Gene* **56**: 117-24.
- Punt, P. J., G. Veldhuisen and C. A. van den Hondel (1994). Protein targeting and secretion in filamentous fungi. A progress report. *Antonie Van Leeuwenhoek* **65**: 211-6.
- Rademacher, T. W., R. B. Parekh and R. A. Dwek (1988). Glycobiology. *Annual Review of Biochemistry* **57**: 785-838.
- Rasmussen, J. R. (1992). Effect of glycosylation on protein function. *Current Opinion in Structural Biology* **2**: 682-686.
- Riederer, M. A. and A. Hinnen (1991). Removal of N-glycosylation sites of the yeast acid phosphatase severely affects protein folding. *Journal of Bacteriology* **173**: 3539-46.
- Roberts, D. B., W. J. Mulvany, R. A. Dwek and P. M. Rudd (1998). Mutant analysis reveals an alternative pathway for N-linked glycosylation in *Drosophila melanogaster*. *European Journal of Biochemistry* **253**: 494-8.
- Rodrigues, M. L., S. Rozental, J. N. Couceiro, J. Angluster, C. S. Alviano and L. R. Travassos (1997). Identification of N-acetylneuraminic acid and its 9-O-acetylated derivative on the cell surface of *Cryptococcus neoformans*: influence on fungal phagocytosis. *Infection and Immunity* **65**: 4937-42.
- Romero, P. A. and A. Herscovics (1989). Glycoprotein biosynthesis in *Saccharomyces cerevisiae*: characterization of the α -1,6-mannosyltransferase which initiates outer chain formation. *Journal of Biological Chemistry* **264**: 1946-1950.
- Roy, S. K., Y. Chiba and Y. Jigami (2000). Production of therapeutic glycoproteins through the engineering of glycosylation pathway in yeast. *Biotechnology and Bioprocess Engineering* **5**: 219-226.
- Royer, J. C., K. Dewar, M. Hubbes and P. A. Horgen (1991). Analysis of a high frequency transformation system for *Ophiostoma ulmi*, the causal agent of Dutch elm disease. *Molecular and General Genetics* **225**: 168-76.

- Rudd, P. M., R. J. Woods, M. R. Wormald, G. Opdenakker, A. K. Downing, I. D. Campbell and R. A. Dwek (1995). The effects of variable glycosylation on the functional activities of ribonuclease, plasminogen and tissue plasminogen activator. *Biochimica et Biophysica Acta* **1248**: 1-10.
- Ruiz-Herrera, J. (1992). Fungal Cell Wall: Structure, Synthesis, and Assembly. Boca Raton, CRC Press.
- Scaman, C. H., F. Lipari and A. Herscovics (1996). A spectrophotometric assay for alpha-mannosidase activity. *Glycobiology* **6**: 265-70.
- Schatzle, J., J. Bush and J. Cardelli (1992). Molecular cloning and characterization of the structural gene coding for the developmentally regulated lysosomal enzyme, alpha-mannosidase, in *Dictyostelium discoideum*. *Journal of Biological Chemistry* **267**: 4000-7.
- Sharon, N. and H. Lis (1993). Carbohydrates in cell recognition. *Scientific American* **268**: 82-9.
- Soares, R. M., C. S. Alviano, J. Angluster and L. R. Travassos (1993). Identification of sialic acids on the cell surface of hyphae and yeast forms of the human pathogen *Paracoccidioides brasiliensis*. *FEMS Microbiology Letters* **108**: 31-4.
- Soares, R. M., F. Costa e Silva-Filho, S. Rozental, J. Angluster, W. de Souza, C. S. Alviano and L. R. Travassos (1998). Anionogenic groups and surface sialoglycoconjugate structures of yeast forms of the human pathogen *Paracoccidioides brasiliensis*. *Microbiology* **144**: 309-14.
- Soares, R. M., A. S. R. M. de, D. S. Alviano, J. Angluster, C. S. Alviano and L. R. Travassos (2000). Identification of sialic acids on the cell surface of *Candida albicans*. *Biochimica et Biophysica Acta* **1474**: 262-8.
- Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**: 503-17.
- Souza, E. T., F. C. Silva-Filho, W. De Souza, C. S. Alviano, J. Angluster and L. R. Travassos (1986). Identification of sialic acids on the cell surface of hyphae and conidia of the human pathogen *Fonsecaea pedrosoi*. *Journal of Medical and Veterinary Mycology* **24**: 145-54.
- Swofford, D. L. (1998). PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Sunderland, MA, Sinauer.

- Takeuchi, M., N. Inoue, T. W. Strickland, M. Kubota, M. Wada, R. Shimizu, S. Hoshi, H. Kozutsumi, S. Takasaki and A. Kobata (1989). Relationship between sugar chain structure and biological activity of recombinant human erythropoietin produced in Chinese hamster ovary cells. *Proceedings of the National Academy of Sciences of the United States of America* **86**: 7819-22.
- Takeuchi, M., S. Takasaki, M. Shimada and A. Kobata (1990). Role of sugar chains in the in vitro biological activity of human erythropoietin produced in recombinant Chinese hamster ovary cells. *Journal of Biological Chemistry* **265**: 12127-30.
- Tanner, W., M. Gentsch, T. Immervoll, A. Scheinost and S. Strahl-Bolsinger (1995). Fungal glycoproteins and their biosynthetic pathway as potential targets for antifungal agents. *Acta Biochimica Polonica* **42**: 505-8.
- Tautz, D. (1992). Redundancies, development and the flow of information. *Bioessays* **14**: 263-6.
- Taylor, A. K. and R. Wall (1988). Selective removal of alpha heavy-chain glycosylation sites causes immunoglobulin A degradation and reduced secretion. *Molecular and Cellular Biology* **8**: 4197-203.
- Temple, B. and P. A. Horgen (2000). Biological roles for cerato-ulmin, a hydrophobin secreted by the elm pathogens, *Ophiostoma ulmi* and *O. novo-ulmi*. *Mycopathologia* **92**: 1-9.
- Temple, B., P. A. Horgen, L. Bernier and W. E. Hintz (1997). Cerato-ulmin, a hydrophobin secreted by the causal agents of Dutch elm disease, is a parasitic fitness factor. *Fungal Genetics and Biology* **22**: 39-53.
- Thompson, J. D., D. G. Higgins and T. J. Gibson (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-80.
- Tremblay, L. O., N. Campbell Dyke and A. Herscovics (1998). Molecular cloning, chromosomal mapping and tissue-specific expression of a novel human α -1,2-mannosidase gene involved in N-glycan maturation. *Glycobiology* **8**: 585-95.
- Tremblay, L. O. and A. Herscovics (1999). Cloning and expression of a specific human α -1,2-mannosidase that trims Man₉GlcNAc₂ to Man₈GlcNAc₂ isomer B during N-glycan biosynthesis. *Glycobiology* **9**: 1073-1078.
- Tremblay, L. O. and A. Herscovics (2000). Characterization of a cDNA encoding a novel human golgi α -1,2- mannosidase (IC) involved in N-glycan biosynthesis. *Journal of Biological Chemistry* **275**: 31655-60.

- Tulsiani, D. R. and O. Touster (1987). Substrate specificities of rat kidney lysosomal and cytosolic alpha-D- mannosidases and effects of swainsonine suggest a role of the cytosolic enzyme in glycoprotein catabolism. *Journal of Biological Chemistry* **262**: 6506-14.
- van den Hombergh, J. P., P. J. van de Vondervoort, L. Fraissinet-Tachet and J. Visser (1997). *Aspergillus* as a host for heterologous protein production: the problem of proteases. *Trends in Biotechnology* **15**: 256-63.
- Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**: 97-130.
- Wagner, A. (1994). Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 4387-91.
- Walsh, J. B. (1995). How often do duplicated genes evolve new functions? *Genetics* **139**: 421-8.
- Wernars, K., T. Goosen, B. M. Wenekes, K. Swart, C. A. van den Hondel and H. W. van den Broek (1987). Cotransformation of *Aspergillus nidulans*: a tool for replacing fungal genes. *Molecular and General Genetics* **209**: 71-7.
- West, C. A. (1981). Fungal elicitors of the phytoalexin response in higher plants. *Naturwissenschaften* **68**: 447-457.
- Wilson, R., R. Ainscough, K. Anderson, C. Baynes, M. Berks, J. Bonfield, J. Burton, M. Connell, T. Copsey, J. Cooper, A. Coulson, M. Craxton, S. Deer, Z. Du, R. Durbin, A. Favello, L. Fulton, A. Gardner, P. Green, T. Hawkins, L. Hillier, M. Jier, L. Johnston, M. Jones, J. Kershaw, J. Kirsten, N. Laister, P. Latreille, J. Lightning, C. Lloyd, A. McMurray, B. Mortimore, M. O'Callaghan, J. Parsons, C. Percy, L. Rifken, A. Roopra, D. Saunders, R. Schownkeen, N. Smaldon, A. Smith, E. Sonnhammer, R. Staden, J. Sulston, J. Thierry-Mieg, K. Thomas, M. Vaudin, K. Vaughan, R. Watterston, A. Watson, L. Weinstock, J. Wilkenson-Sproat and P. Wohldman (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of *Caenorhabditis elegans*. *Nature* **368**: 32-38.
- Wolfe, K. H. and D. C. Shields (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**: 708-13.
- Yamashita, K., S. Hara-Kuge and T. Ohkura (1999). Intracellular lectins associated with N-linked glycoprotein traffic. *Biochimica et Biophysica Acta* **1473**: 147-60.

- Yamashita, K., E. Ichishima, M. Arai and A. Kobata (1980). An alpha-mannosidase purified from *Aspergillus saitoi* is specific for alpha 1,2 linkages. *Biochemical and Biophysical Research Communications* **96**: 1335-42.
- Yoshida, T. and E. Ichishima (1995). Molecular cloning and nucleotide sequence of the genomic DNA for 1,2- alpha-D-mannosidase gene, msdC from *Penicillium citrinum*. *Biochimica et Biophysica Acta* **1263**: 159-62.
- Yoshida, T., T. Inoue and E. Ichishima (1993). 1,2-alpha-D-mannosidase from *Penicillium citrinum*: molecular and enzymic properties of two isoenzymes. *Biochemical Journal* **290**: 349-54.
- Yoshida, T., T. Nakajima and E. Ichishima (1998). Overproduction of 1,2-alpha-mannosidase, a glycochain processing enzyme, by *Aspergillus oryzae*. *Bioscience, Biotechnology, and Biochemistry* **62**: 309-15.
- Yoshihisa, T. and Y. Anraku (1989). Nucleotide sequence of AMS1, the structure gene of vacuolar alpha- mannosidase of *Saccharomyces cerevisiae*. *Biochemical and Biophysical Research Communications* **163**: 908-15.
- Yoshihisa, T. and Y. Anraku (1990). A novel pathway of import of alpha-mannosidase, a marker enzyme of vacuolar membrane, in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **265**: 22418-25.