UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Identification of South African fungi with high protein secreting capabilities

By

MUBANGA HELEN KABWE

Submitted in partial fulfillment of the requirements for the degree

MAGISTER SCIENTIAE (MSc.) MICROBIOLOGY

In the

Faculty of Natural & Agricultural Sciences

Department of Microbiology and Plant Pathology

University of Pretoria

Pretoria

SUPERVISOR: PROFFESOR DON COWAN

CO-SUPERVISOR: DR ELDIE BERGER

18 September 2015

DECLARATION

I, Mubanga Helen Kabwe, declare that the dissertation, which I hereby submit for the degree *Magister Scientiae* (MSc.) in Microbiology to the Department of Microbiology and Plant Pathology at the University of Pretoria, Hatfield campus, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution. Every source used or quoted has been indicated and acknowledged.

Mubanga Helen Kabwe

Signature: ……

Date: ……

ABSTRACT

Recombinant protein technology has facilitated high level production of proteins to try and meet the demand from industry. One major factor that has been reported to influence the scales of production in industrial settings is the relationship between host cell growth rate and protein production rate. Therefore, expression systems with fast growth and high protein production rates are favorable for high volume recombinant protein production. Filamentous fungi have over the years been used to produce a range of valuable metabolites including biologically active proteins and organic acids. Their ability to synthesize and secrete extracellular enzymes in large amounts made them the preferred host systems for heterologous protein production. In this study the CMW: Culture Collection of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa, was screened using SDS-PAGE gels to identify a fungal isolate with the natural ability to secrete a highly expressed protein (HEP). The HEP was identified through liquid chromatography-mass spectrometry (LC-MS/MS) and the 330 bp gene encoding the HEP was isolated, cloned and transformed into *E. coli* cells. The gene sequence was further investigated using a modified SiteFinding-PCR chromosome walking technique, to acquire the full *man* gene sequence and to identify the upstream promoter and regulatory sequences flanking the gene in the 5' to 3' direction. *In silico* analysis of acquired nucleotide sequences in the 5' direction promoter region of the gene, predicted transcription initiation sites, transcription factor binding sites as well as TATA boxes that have been previously identified in the promoter region of other *Ophiostoma* proteins. The growth conditions of *O. phasma* 20676 were also optimized for high level protein production.

ACKNOWLEDGEMENTS

This research project is the collective effort of many people. To these people, I extend my deepest and sincere gratitude for making this research possible.

My **family** for their support and faith in me, for always encouraging me to do more.

My **colleagues** from the Centre for Microbial Ecology and Genomics **(CMEG)**, for their assistance and friendship.

The Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, for allowing access and use of the CMW fungal culture collection.

The **University of Pretoria** and the Department of Microbiology and Plant Pathology, for taking me on as an MSc student in their program.

Mostly I would like to thank my supervisors, **Professor Don Cowan**, for seeing potential in me and taking a chance on me. **Dr Eldie Berger**, who exhibited extraordinary enthusiasm and intellect. Your high spirits helped me get up during those very tough mornings.

The financial assistance of the **National Research Foundation (NRF)** and the University of Pretoria **Institutional Research Theme (IRT) Genomics**, towards this research is hereby acknowledged.

I mostly thank God for giving me such an opportunity.

DEDICATION

This research project is dedicated to the Kabwe family, my wonderful parents, Martin and Catherine, my two lovely and delightful sisters, Rachel and Chongo.

“Per aspera ad astra”

THESIS COMPOSITION

Chapter one of this thesis introduces recombinant protein production by giving a brief history on protein production and the importance thereof. The chapter outlines knowledge regarding different available recombinant protein expression systems and emphasizes some of their advantages and disadvantages. Filamentous fungal expression systems are elaborately discussed and the requirements for developing an expression system are outlined. Lastly transformation methods that are commonly used such as *Agrobacterium tumefactions*-mediated transformation (ATMT) are discussed and the aims and objectives of this study are listed in this chapter.

Chapter two comprises the materials and methods used in this study. Screening of the CMW culture collection for a high protein secreting isolate through SDS-PAGE electrophoresis is described, as well as the optimization of growth parameters of *O. phasma* 20676 for high level protein production. Molecular biology techniques such as ITS identification of the fungal isolate and isolation of the *man* gene through PCR amplification, cloning and sequencing are also described. The last section of this chapter describes SiteFinding PCR as a chromosome walking technique used to identify the up- and down-stream regulatory sequences flanking the *man* gene.

Chapter three outlines the results obtained during screening of the CMW fungal culture collection, ITS isolate identification, LC-MS/MS peptide identification and optimization of growth conditions of *O. phasma* 20676 for high level protein production. Finally this chapter reports on the *in silico* identification of the up-stream promoter region and other transcriptional elements flanking the *man* gene in the up- and down-stream regions.

Chapter four discusses the results obtained in this study and key results which have contributed to a better understanding of the effects of different growth parameters on mannanase production by *O. phasma* 20676 are highlighted.

The **References** chapter lists all the literature used and cited in this study.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

APS	Ammonium persulphate
ATMT	*Agrobacterium tumefaciens*-mediated transformation
BLAST	Basic local alignment tool
bp	Base pair
BSA	Bovine serum albumin
CMW	Culture collection by Michael Wingfield
dH_2O	Sterile distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
g	Gram
g	gravity-force
GFP	Green fluorescent protein
GRAS	Generally regarded as safe
HEP	Highly expressed protein
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITS	Internal transcribed spacer
kb	kilo base
kDa	kilo Dalton
LBG	Locust bean gum
LC-MS/MS	Liquid chromatography-mass spectrometry
ME	Malt extract
MYE	Malt yeast extract
ml	Milliliter

mM	Millimolar
mRNA	Messenger ribonucleic acid
NCBI	National center for biotechnology information
OD	Optical density
ORF	Open reading frame
PEG	Poly-ethylene-glycol
PCR	Polymerase chain reaction
pH	Potential of hydrogen
RNA	Ribonucleic acid
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SOD-PI	Superoxide dismutase-human pro-insulin
TEMMED	N,N,N-tetramethyl ethylenediamine
TAE	Tris-acetate EDTA
TRIS	Tris-(hydroxymethyl) aminomethane
U	Units of enzyme
US-FDA	United States-Food and Drug Administration
UV	Ultra violet
µl	Microliter
V	Voltage
v/v	Volume per volume
w/v	Weight per volume

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

1. Introduction

In the 1970s traditional methods for protein production involved extractions from plant, human and animal tissue and/or body fluid (Josephson & Bishop, 1988). However, this resulted in very low yields that were costly and limited in availability; e.g. 3000 kg of bovine intestine could only yield 80 µg of pure biologically active secretin (Carlquist et al., 1984). In 1954 chemical synthesis of peptides and proteins was introduced by production of the synthetic oligopeptide hormone, oxytocin (du Vigneaud et al., 1954). Compared to conventional methods, chemical synthesis increased the amount and availability of some important peptides and proteins. For example, the artificial dipeptide sweetener, aspartame was produced in high yields of 4000 tonnes per year (Josephson and Bishop 1988). In addition, chemical synthesis was suitable for large-scale production of shorter peptides that were free from contamination by prions and viral material. However, according to Josephson and Bishop (1988) the downside to chemical synthesis was limited production of longer peptides (batches of 1-10kg), which accounted for only 50-60% of total protein produced and recovered by costly High performance liquid chromatography (HPLC).

Preparation of the first commercial enzyme Taka-diastase, was patented in 1984 by Jokichi Takamine (Bennet, 1998). This was a crude amylase extracted from *Aspergillus oryzae* and mainly used for digestion of starch. This development marked the beginning of large-scale microbial protein and metabolite production in the late 1900s (Demain & Vaishnav, 2009). Microbial production of enzymes in the late 1980s and early 1990s resulted in a thriving industry due to advances in genetic engineering and DNA technology. This allowed for routine cloning and expression of foreign genes outside of their

natural expression systems, providing a possible solution for previously low levels of protein or peptide production. In 1980, the gene encoding human insulin was successfully expressed in *Escherichia coli strain* K-12 (Goeddel et al. 1979; Keen et al. 1980). To increase the yield of insulin, the fusion protein superoxide dismutase-human pro-insulin (SOD-PI) was expressed in an alternative host, the yeast *Saccharomyces cerevisiae* (Tøttrup & Carlsen, 1990). This led to the production of 1500 mg SOD-PI from 1 L of *S. cerevisiae* culture (Mahmound, 2007). The heterologously expressed recombinant insulin was found to be identical to human insulin, both functionally and immunologically. This development provides a reliable and safe source of insulin for diabetics, with potential for scale-up in production (Rotenstein et al., 2012). Subsequently, advances in biotechnology and protein engineering have facilitated the commercial scale production of native and recombinant proteins used in an array of industries. Recombinant and native proteins are used in the enzyme industry, pharmaceutical industry and in agriculture. They are used in manufacturing of medicine, food, textile, paper, leather, polymers and detergents (Demain & Vaishnav, 2009). In academic settings these proteins are used for exploratory research in human and veterinary medicine (Saraswat et al., 2013). Recombinant protein use also extends to include environmental bioremediation and development of bio control agents for pests and pathogens (Magaña-Ortíz et al., 2013).

Novozymes is an industrial biotechnology company which focuses on the commercial scale production of enzymes. In 2012 the company held approximately 47% of industrial enzyme sales on the global market. According to the Novozymes sales and marketing report for 2012, enzyme business sales were US$ 1.49 million in 2012, an increase of 7% compared to the previous

year. According to the report, enzymes used in household detergents and feed were major contributors to growth in 2012 (Novozymes 2012). At present recombinant protein production is a lucrative market holding approximately US$ 100 billion of recombinant therapeutic protein sales, whereas the industrial enzyme market is around US$ 6 billion (Delic, Göngrich et al., 2014).

Even though a significant number of recombinant proteins are produced at commercial scale, recombinant protein production remains a challenge (Palomares, 2004). The success of any commercial recombinant protein production industry is influenced by the ability to produce proteins at a sufficiently high economically competitive level (Gasser & Mattanovich, 2007). Protein yields are expressed as quantities in grams, milligrams, or micrograms of protein produced in a defined volume (usually a liter) of a culture. As a standard, commodity proteins are expressed at the multi-gram per fermentation-liter level, specialized enzymes at multi-milligram per liter and therapeutic proteins at microgram per liter scales (Gasser & Mattanovich, 2007). High protein yields are in the grams per liter range usually between 25-100 g/L (Zhang, 2010). However, Dyadic International Inc., a global biotechnology company responsible for discovering, developing, manufacturing and selling enzymes and other proteins for the bioenergy, bio-based chemical, biopharmaceutical and industrial enzyme sector, have reported protein production levels of up to 100 g/L of protein (Demain et al., 2009). Protein yields in the milligrams per liter range are considered intermediate levels, with low yields in the micrograms per liter range. Intermediate and low protein yields are usually sufficient for most biochemical analysis and for limited biochemical studies, respectively (Chen, 2012). The required level of expression is closely linked to the commercial nature of the

protein. Therefore, various expression technologies may be used to increase protein yield.

With the number of applications requiring large volumes of high-quality proteins increasing globally, the demand for efficient production strategies are increasing. High production efficiencies, low production time and costs, protein quality and functionality are all major determining factors for a commercially viable process (Demain & Vaishnav, 2009). Therefore, large-scale recombinant protein production requires expression systems with high product formation capabilities at low growth rates (Hensing et al., 1995).

1.1 Protein production systems

Production of recombinant proteins involves transcription of the gene encoding the highly expressed protein, followed by protein translation in a suitable system (Demain & Vaishnav, 2009). A variation of *in vivo* and *in vitro* protein production systems are available, with the former being the most preferred. This is because chemical synthesis of proteins cannot compare to the accuracy of native hosts for synthesis of peptides and polymers (Agbo, 2012). Two classes of *in vivo* expression systems are available: prokaryotic and eukaryotic systems. Host organisms used in the production of heterologous proteins include plants, mammals, insects, fungi and bacteria. Table 1.1 summarizes some characteristics, pros and cons of available expression systems.

Table 1.1 Advantages and disadvantages of available expression systems (Desai et al. 2010; Su et al., 2012).

System	Production cost	Production time scale	Scale-up capacity	Expression level	Glycosylation	Contamination risk	Storage cost
Bacteria	Low	Short	High	High	Absent	Endotoxins	Moderate
Yeast	Medium	Medium	High	Low-high	Higher monosylation	Low risk	Moderate
Filamentous fungi	Low	Medium	High	High	Present	Low risk	Moderate
Insect cell culture	High	Medium	Medium	Low-high	Higher monosylation	High risk	Expensive
Mammalian cell culture	High	Long	Very low	Low-moderate	Similar to human	Viruses, Prions	Expensive
Transgenic animal	High	Very long	Low	Moderate-high	Similar to human	Viruses, Prions	Expensive
Plant cell	Low	Short	High	Moderate-high	Minor differences	Low risk	Inexpensive
Transgenic plants	Very low	Long	Very high	Moderate-high	Minor differences	Low risk	Inexpensive

1.1.1 Microbial expression systems

Microorganisms are characterized by the ability to grow to high biomass with secretion of a variety of proteins and enzymes in very limited time. They are capable of exponential growth in a variety of environmental conditions and on a number of different substrates (Agbo, 2012). Therefore, microbial systems are more economical than animal and plant systems as cultivation is simpler and faster. Advances in genetic engineering have allowed for high quality and high volume production of proteins and/or enzymes. Currently complete protein secretion pathways have been altered and foreign genes added for production of recombinant proteins (Delic et al., 2014). Genomics and genetic engineering have provided molecular tools that allow for easy genetic manipulation of microbial systems, to produce high amounts of good quality recombinant proteins (Chen, 2012; Demain and Vaishnav, 2009). These are some of the reasons that make microbial systems a preferred choice for the production of high volume proteins and peptides.

1.1.1.1 Prokaryotic systems

The Gram-negative bacterium *E. coli* is one of the first and vastly utilized prokaryotic expression system (Terpe, 2006). The first heterologous peptide (mammalian somatostatin), was expressed in *E. coli* in 1977 (Itakure et al., 1977). Thereafter, *E. coli* has become the work horse for molecular biology as a protein production host, leading to industrial scale production of many commercial proteins such as, lipases, polymarases, phosphotases , endonucleases (Huang et al., 2012) and manganese superoxide dismutase (Newman et al., 2008). Some reasons why *E. coli* is the favored choice of expression system can be attributed to its rapid and inexpensive culturing, well-characterized genetics and high product yields (Huang et al., 2012). Progress in understanding processes in *E. coli* such as transcription, translation and protein folding accompanied by availability of improved genetic tools, makes this bacterium favorable for expression of non-glycosylated proteins (Demain & Vaishnav, 2009). The pros and cons of this expression system are outlined in Table 1.2.

Table 1.2 Advantages and disadvantages of *E.coli* as an expression system (Demain & Vaishnav, 2009).

Advantages	Disadvantages
Rapid expression	Proteins with disulphide bonds are difficult to express
High yields	Produce un-glycosylated proteins
Ease of culture and genome modifications	Proteins produced with endotoxins
Inexpensive	Acetate formation resulting in cell toxicity
Mass production fast and cost effective	Some proteins produced as inclusion bodies, are inactive; require refolding

Though *E. coli* contains many qualities of a good expression system, challenges remain in production of certain proteins (Demain and Vaishnav,

2009). *E. coli* is a prokaryotic based expression system; therefore, heterologous proteins from eukaryotes often lack post translational modifications when expressed in *E. coli* (Khow & Suntrarachun, 2012). Moreover, the secretion of proteins expressed in large amounts is made difficult due to precipitation which forms insoluble, often inactive inclusion bodies which are not completely folded. Therefore, complicated downstream processes are needed for purification (Demain & Vaishnav, 2009).

Gram-positive *Bacillus* is another industrially important expression system. It is mainly used for expression of bacterial enzymes such as amylases and proteases and protein titres of 20-25 g/L have been reported in this organism (Deb et al., 2013). Well-known industrial recombinant protein producers are *B. megaterium*, *B. brevis*, *B. subtilis* and *B. licheniformis* (Terpe, 2006).

According to Demain et al. (2009), some advantages of using bacilli as expression systems include their capability to extracellularly produce proteins making recovery efficient and cost effective. The genome of bacilli has been studied in great detail and information and tools are available for genetic manipulation, transformation and gene replacement technologies (Dong & Zhang, 2014). Bacilli are capable of growth to high biomass in a short time and are metabolically robust. The US-FDA has endorsed some *Bacillus* species as Generally Recognized as Safe (GRAS) for application in consumer goods.

1.1.1.2 Eukaryotic systems

Most eukaryotic proteins require post translational modifications to amino acid structure for the production of bio active proteins. These modifications include glycosylation, phosphorylation, acetylation and proteolytic cleavage (Demain & Vaishnav, 2009). Glycosylation involves the covalent attachment of carbohydrate residues through nitrogen (N-linked) or oxygen (O-linked) bonds

to specific amino acids (Lang and Looman 1995; Wang et al. 1996). N-linked bonds have the sugar covalently attached to the nitrogen on the asparagine acid, whereas the O-linked is attached through the oxygen on the lysine, serine, threonine or proline acids (Lang and Looman, 1995; Wang et al., 1996). The problem with prokaryotic systems is that they lack the mechanism to efficiently add post translational modifications to eukaryotic proteins (Kamionka, 2011).

Available eukaryotic expression systems are sourced from; yeast cells, filamentous fungal cells, cells of mammals, insects, transgenic animals and plant cells. However, in this study focus will be on filamentous fungi and yeast microbial systems.

a) Yeasts

These are unicellular fungal organisms that have increasingly become preferred hosts for expressing heterologous genes. They are often used as alternative hosts to bacterial systems for production of recombinant proteins due to their ability to add post translational modifications (Demain & Vaishnav, 2009). Compared to mammalian and insect host cells, yeasts utilize glycerol and a diverse range of other sugars as carbon source, and therefore grow rapidly on simple low cost heterogeneous biomass. Filamentous fungi and yeasts have been cultured on yeast extract, malt extract and dextrose based media (Santamauro et al., 2014).

Molecular tools have been developed to engineer strains successfully for protein production (Padh, Rai, & Padh, 2001). The two most widely used yeast strains are *S. cerevisiae* and *Pichia pastoris* (Demain & Vaishnav, 2009).

A considerable 20 % of all recombinant pharmaceutical and diagnostic proteins are produced in *S. cerevisiae* (Glick, Barth, & Macleod, 2010), including insulin precursor, human transferrin, hepatitis surface antigen used in production of

vaccines and glucagon (Nielsen, 2013). Table 1.3 Shows some biopharmaceutical proteins produced by *S. cerevisiae*

Table 1.3 Some biopharmaceuticals produced by *S. cerevisiae* (Nielsen 2013).

Type	Protein	Therapeutic Application	Leader sequence	Titre
Blood related	Human serum albumin	Surgery (plasma expander)	Native	3 g/L
	Hirudin	Blood coagulation disorders	α-factor	460 mg/L
	Human transferin	Anemia	Native	1.8 g/L
Hormones	Insulin precursor	Diabetes	Synthetic	80 mg/L
	Glucagon	Diabetes	α-factor	17.5 mg/L
Antigene	Hepatitis surface antigene	Hepatitis vaccination	Native	19.4 mg/L

In the biofuel industry *S. cerevisiae* is used in large scale production of fuels and chemicals (Wilde et al., 2012). Important characteristics that are exploited include the ability to grow well under industrial anaerobic fermentation conditions, at low pH, and the general tolerance to elevated ethanol concentrations (Almeida et al., 2007; Wilde et al. 2012). These characteristics make *S. cerevisiae* suitable for large scale fermentation.

However, one major limitation in using *S. cerevisiae* is that it adds rich mannose type post translational modifications (N-glycosylation) (Nielsen, 2013), which is unfavorable for most mammalian proteins due to lack of sialylated O-linked chains (Demain & Vaishnav, 2009).

b) Filamentous fungi

The natural capabilities of filamentous fungi to secrete large amounts of enzymes makes them a favored system for the expression of heterologous genes (Lubertozzi & Keasling, 2009; Punt et al., 2002; Ward, 2011). Uses of filamentous fungi as sources of enzymes and metabolites dates back centuries ago (Ward, 2012). In traditional fermentation processes, filamentous fungi were used for the production of food and beverages such as bread, cheese, sake, wine and beer (Punt et al., 2002). Over the past century their use in industry have increased to include synthesis of organic acids, secondary metabolites, heterologous proteins and enzymes valuable to the pharmaceutical, chemical, textile and food industries (Moore, 2007). In the enzyme industry, *A. niger* is used for high level production of glucoamylases. Protein yields as high as 25 g/L have been reported (Ward et al., 2006). Lamsa and Bloebaum, (1990) improved titers of a genetically-engineered bovine chymosin-producing strain of *A. awamori* by 500% through conventional mutagenesis and screening. The biopharmaceutical protein, human lactoferrin, was also produced in titres of 2 g/L by *A. awamori* using rDNA technology and classical strain improvement (Demain et al., 2009). Large-scale fungal fermentation technologies are well established and utilized for high level recombinant protein production (Nevalainen, Te'o, & Bergquist, 2005).

Characteristics that make filamentous fungi preferred hosts for heterologous protein production include:

1. Their natural ability to produce proteins in large quantities: e.g., *A. niger*.
2. Proteins are often secreted outside the cell, reducing labor and the cost of protein recovery.

3. They have post translational protein processing machinery which adds modifications such as glycosylation or formation of multiple disulfide bonds often needed for biological activity of eukaryotic proteins (Nevalainen et al, 2014). Filamentous fungi have high mannose N- and the highly conserved O-Linked mannosyltransferase glycosylation (Deshpande et al., 2008).
4. They grow to high biomass on relatively inexpensive culturing medium, in reasonable time; approximately 2 days to a week.
5. The food and drug administration (FDA) has endorsed some filamentous fungi: e.g., *A. niger*, *A. oryzae* (Zhao et al., 2014) and *Trichoderma viride* (Fadel et al., 2015), as having GRAS (generally regarded as safe) status.
6. Advances in fungal genomics and biotechnology have provided molecular biology tools used for manipulating and genetically engineering filamentous fungi.

Available filamentous fungal expression systems include *Aspergillus* and *Trichoderma* species, which were principal organisms, involved in food fermentation processes (Nevalainen & Peterson, 2014; Ward, 2012). The highest yields of recombinant proteins have been obtained from expression of fungal genes in fungal hosts. Tamayo-Ramos and Orejas, (2014) successfully improved protein production and reduced production time of a heterologous α-L-rhamnosidase from *A. culeatus* in *A. nidulans*. α-L-rhamnosidase activity in transformed *A. nidulans strains* was approximately 588 mU/mL; 9300 mU/mg after 24 hours and 8.5-fold more than that of the wild type (Tamayo-Ramos & Orejas, 2014). Nevalainen et al., (2005) compared expression systems (yeast, insect cells, mammalian cells, plants and bacteria) for heterologous protein

production in liquid culture, and filamentous fungi produced the highest yields at minimum production costs.

Although heterologous fungal genes are efficiently expressed and high to moderate protein yields in the g/L and mg/L range are obtained, the expression of genes from bacteria, mammals, insects, and plants results in low yields ranging in the µg/L range (Nevalainen et al, 2005). Several strain and yield improvement techniques have been successfully applied to increase production of foreign proteins in filamentous fungi. Improvements comprise increasing the gene copy number, development of protease deficient host strains, fusion of a highly-expressed endogenous protein acting as a carrier protein, and the use of strong promoters to drive expression (Li et al., 2012; Sharma, 2009).

1.2 Development of fungal expression systems

Transformation systems that have been developed for some industrial fungal isolates, stimulated research towards sourcing new filamentous fungal recombinant expression systems (Lubertozzi & Keasling, 2009). This technology has facilitated the improvement of existing fungal strains for production of industrially important homologous and heterologous proteins. The general requirements for a successful fungal expression system include:

1. A recombinant expression vector constituting transcriptional, translational and at times secretory signals flanking the DNA sequences encoding the protein of choice.

2. A fungal host organism that can efficiently express the recombinant gene.
3. An efficient method for introduction of the expression vector into the host organism (Rivera et al., 2014).

1.2.1 Construction of expression vectors

Expression vectors (mostly plasmids) are used to introduce foreign DNA into host cells for expression. Highly active fungal promoters are crucial for efficient expression vectors as they initiate transcription and therefore control expression (Fleißner & Dersch, 2010). Basic components of an expression vector for fungal protein production include:

1. Strong promoter for efficient transcription.
2. Transcription start site.
3. Nucleotide sequence encoding a signal peptide for protein secretion.
4. Transcription terminator sequence.
5. Selection marker.
6. Multiple cloning site (MCS) for insertion of target gene.
7. Nucleotide sequence encoding gene of interest

1.2.2 Fungal promoters

A promoter is a stretch of DNA that drives and regulates gene expression. Elevated transcription levels are necessary for efficient secretion of heterologous proteins (Sharma et al., 2009). The identification of strong promoters to drive efficient transcription is needed (Su et al., 2012). Two types of promoters need to be considered: constitutive or inducible promoters. Constitutive promoters allow continuous transcription of a gene, whereas activity of inducible promoters is activated by physical or chemical factors

(Sharma et al., 2009; Su et al., 2012). Inducible promoters are switched on by factors such as temperature, pH, a deficiency in required amino acids or ion concentrations (Weinhandl et al., 2014). An ideal inducible promoter is described as one that is highly selective, tightly regulated and induction of gene expression is inexpensive but results in high product yields (Moore, 2007). The decision on which type of promoter is used for transcription is dependent on requirements of the bioprocess as well as the properties of heterologous protein (Weinhandl et al., 2014). Several promoters have been evaluated and used successfully in filamentous fungi to express and/or produce foreign proteins (Meyer et al., 2011). Table 1.4 lists commonly used constitutive and inducible promoters used in expression of foreign genes in filamentous fungi.

Table 1.4 Constitutive and inducible promoters commonly used to express recombinant genes in filamentous fungi (Moore, 2007).

Constitutive promoter	Gene Function	Source DNA	Reference
Pna2/tpi hybrid promoter	Neutral amylase + triose phosphate isomerase	*A. nidulans* *A. niger*	(Olempska-Beer et al. 2006)
gpdA	Glyceraldehyde-3-phosphate dehydrogenase	*A. nidulans*	(Kourtoglou et al., 2011)
enol	Bovine RNase A	*Neurospora crassa*	(Allgaier et al. 2009)
pki1	Pyruvate kinase	*Trichooderma virens*	(Moran-Diez et al. 2010)

Inducible promoters	Gene Fucntion	Source DNA	Reference
Taka-A amaylase	Amylase hydrolysis	*A. oryzae*	(Moore 2007)
glaA	Glucoamylase	*A. niger var. awamori*	(Punt, van Biezen, et al. 2002)
alcA/AlcR	Alcohol dehydrogenase	*A. nidulans*	(Liao et al. 2009)
niiA	Nitrite reductase	*A. fumigatus*	(Hu et al., 2007)
cbh1	Cellobiohydrolase	*T. reesei*	(Li et al. 2012)
thiA	Thiamine biosynthesis	*A. oryzae*	(Shoji et al., 2005)
ctr4	High affinity copper transporter	*Schizosaccharomyces pombe*	(Bellemare et al., 2001)

a) Constitutive promoters

Constitutive promoters are generally sourced from housekeeping genes. They are active during most processes within a cell and produce high level of expression (Sharma et al., 2009). Several constitutive promoters are used for recombinant protein production; e.g., the *alcA* from *A. nidulans* was used for intracellular expression of carotenoid biosynthesis-related enzymes in *Neuspora crassa* (Wang & Keasling, 2002), the *trpC* promoter from *A. nidulans* was used to express the hygromycin resistance gene (*hph*) in *N. crassa*, (Bardiya & Shiu, 2007) and the *A. niger glaA* promoter was used for expression

of hemagglutinin and neuraminidase in *N. crassa* (Allgaier et al., 2009). One noted disadvantage in using a constitutive promoter is that they retain normal function during growth; therefore if a toxic heterologous protein is produced during the growth phase, it may cause cell death (Sharma et al., 2009). Therefore using an inducible promoter to first increase biomass before the expression of heterologous protein will be desirable.

The *gpdA* promoter

The constitutive promoter that is most often used is the *gpdA* promoter from *A. nidulans* (Kourtoglou et al., 2011; Sharma et al., 2009) which drives strong transcription of heterologous genes in different hosts (Van den Hondel et al.,1991) such as *A. niger*, *A. vadensis* and *Fusarium oxysporum* (Culleton, 2014). The *gpdA* gene encodes the Glyceraldehyde-3-phosphate dehydrogenase (GPDA) enzyme that catalyzes the conversion of dihydroxyacetone phosphate to sn-glycerol 3-phosphate (Su et al., 2012; Xianjin et al., 2006). The *gpdA* promoter was utilized in the expression of the β-galactosidase (*LacZ*) and the β-glucoronidase (*uiaA*) genes in *A. niger* (Punt et al., 1991), and for expressing the phosphoglucomutase gene from *Fusarium oxysporum* in *A. nidulans* (Kourtoglou et al., 2011). Phosphoglucomutase activity in the transformants was higher than in the wild type. Other important heterologous genes that have been successfully expressed by the *gpdA* promoter include; the thaumatin protein II gene in *A. awamori* (Moralejo et al., 1999), the human interleukin-6 protein gene and the manganese peroxidase gene in *A. niger* (Broekhuijsen, 1993; Punt et al., 2002). Due to this promoter's ability to drive constitutive transcription, it is functional in expressing selection markers and reporter genes of heterologous origin, under variable culture conditions (Su et al., 2012).

A *gpdA* promoter isolated from an *A*. *nidulans* strain adapted to high levels of sodium chloride, was used to direct transcription of the β-D-glucuronidase (*uidA)* gene to higher levels with gradual increments of the sodium chloride concentration (Redkar, Herzog, & Singh, 1998). The study proved that, the *gpdA* promoter could be used for high level expression of heterologous genes under culture conditions that are considered non-favorable.

b) Inducible promoters

Inducible promoters facilitate controlled gene expression at specific stages during host organism development (Kück & Hoff, 2010; Weinhandl et al., 2014). They allow for maximum growth of the host organism cell before induction of a potentially stressful expression phase (Weinhandl et al., 2014). Examples of commonly used inducible promoters are the *A. niger* glucoamylase promoter (P*glaA*) and the *T. reesei* cellobiohydrolase promoter (P*cbh1)* (Larrondo et al., 2009; Punt et al. 2002).

The *glaA* promoter

Glucoamylases (GlaA) are exo enzymes that are highly expressed when maltose is used as a carbon source by *A. niger* (Kwon et al., 2012). These enzymes are suitable for use in starch hydrolysis during food and beverage processing making them valuable industrial enzymes. Glucoamylases that are most extensively studied are those secreted by *A*. *niger* and *A. awamori* (Cornett & Reilly, 2003). The induction of extracellular enzymes by *A. niger* is controlled at the transcriptional level (Yuan et al., 2006). The activity of the *glaA* promoter is activated by starch, maltose and low concentrations of glucose, but is repressed by xylose (Fowler, Berka, & Ward, 1990; Su et al. 2012). Transcription initiation of *glaA* requires a DNA sequence approximately 214 bp

upstream of the start codon whereas high level transcription requires sequences located between 318 and 800 bp upstream within the *glaA* promoter region (Fowler et al. 1990; Verdoes, Punt, & Hondel, 1995). The CCAAT protein binding motif which plays a role in regulating *glaA* transcription was determined in *A. niger* (Qui et al., 2002). Subsequently Liu et al. (2003) showed that introducing multiple copies of the CCAAT-motif in *A. niger* considerably increased expression levels of the heterologous protein vitreoscilla haemoglobin (Liu et al., 2003). Zhu et al., (2004), proceeded to confirm that transcription of the *glaA* gene in *A. niger* was regulated by CCAAT-binding proteins, which were later purified, characterized and named as AngCP1 and AngCP2. The *glaA* promoter is also highly compatible and is therefore widely used for expression of heterologous genes across a number of *Aspergillus* species; *A. oryzae* (Bleichrodt et al., 2012), *A. awamori* (Kosalková et al., 2012) and *A. niger* (Su et al., 2012).

1.2.3 Selectable markers

Selectable marker genes confer a trait suitable for facilitating artificial selection of recombinant transformants carrying the vector containing the marker gene (Chuan-Wei & Terry, 2013). A large number of selectable markers that are effective across many fungal species have been developed. Table 1.5 lists some of the most widely used markers. Two types of selection markers used in fungal transformation include auxotrophic markers and dominant markers. Auxotrophic markers compliment a nutritional deficiency in an auxotrophic strain while dominant markers are genes that confer antibiotic resistance (Moore, 2007). For the well-studied model filamentous fungi *N. crassa*, *A. nidulans* and *T. reesei,* some auxotrophic markers have been developed. They include *pyrG* from *A. nidulans* (Gruber et al., 1990) and *argB* from *A. fumigatus*

(Penttila et al., 1987). However, a major problem when using auxotrophy as a selectable marker is the need for an auxotrophic mutant strains (Ruiz-Díez, 2002). For this reason dominant selection markers that impart drug resistance are most widely used. The hygromycin B resistance gene *hph* (Cullen et al., 1987) from *E. coli* was successfully used to transform filamentous fungi; *A. niger, T. reesei*, *Ophiostoma piceae* and *Ophiostoma ulmi* (Hoffman & Breuil, 2004; Loppnau et al., 2004; Mach, Schindler, & Kubicek, 1994; Punt et al., 1987; Royer et al., 1991).

Table 1.5 Selectable markers used in filamentous fungi (Moore, 2007; Su et al., 2012).

	Function of resistance/nutritional gene	Source organism	Reference
Auxotrophic markers*			
niaD		*A. nidulans*	(Daboussi et al., 1989)
pyrG	Oritidine-5'-phosphate decarboxylase	*A. niger*	(Balance and turner, 1985)
Dominant markers			
Benomyl-*tub*	Benomyl-resistant tubulin mutants	*N. crassa*	(Orbach et al., 1986)
Bialophos/phosphinothricin-*bar*	Phosphinothricin acetyltransferase	*Streptomyces hygroscopicous*	(Avalos et al., 1989)
Blasticidin *S-bsr* *bsd*	Blasticidin S deaminase	*B. cereus* *A. terreus*	(Kimura et al., 1994)
Carboxin-*cbx*^{R}	Carboxin-resistant succinate Dehydrogenase mutants	*Ustilago maydis*	(Kojic and Holloman,, 2000)
amds	Acetamidase	*A. nidulans*	(Tilburn et al., 1983)
Phleomycin/zeocin/bleomycin-*ble*	Bleomycin binding protein	*Streptoalloteichus hindustanas*	(Mattern and Punt, 1988)
Pyrithiamine-*ptrA*	Mutated allele of thiamine biosynthesis gene	*A. oryzae*	(Kubodera et al.,2000)
Nourseothricin-*nat*	Nourseothricin	*Streptomyces noursei*	(Krugel et al., 1993)
Hygromycin-*hph*	Hygromycin phosphotransferase	*E. coli*	(Punt et al., 1987)

***Both are positive-negative selection systems

1.2.4 Signal peptides

Signal peptides are short peptide sequences which usually constitute 15 to 60 amino acids (Su et al., 2012). They are necessary for directing the expressed protein into the secretory pathway (Bzymek et al., 2004; Jalving et al., 2000) Signal sequences are located at the N-terminus or the C-terminus of a protein and are cleaved in the endoplasmic reticulum (ER) during secretion (Bzymek et al., 2004; Jalving et al., 2000; Ruiz-Díez, 2002). To facilitate the secretion of

proteins by filamentous fungi, N-terminal signal peptides are most commonly used (Su et al., 2012).

Signal peptides are divided into three regions namely: the N-region located where the amino acid starts (N-terminus) which is five to six amino acids long and positively charged. The H-region has a hydrophobic core consisting of 7-16 amino acids, and is uncharged. Lastly the C-region located at the end of an amino acid chain (C-terminus) and is 4-6 amino acids long (Izard & Kendall, 1994; Nothwehr & Gordon, 1990; Su et al. 2012). Though amino acid sequences of signal peptides are diverse, the signal peptides from different organisms functions in a similar manner (Su et al., 2012). Consequently, even some prokaryotic signal peptides are recognized by the translocation system of eukaryotic organisms, and vice versa (Watts et al., 1983). The secretion of human interleukin-6 (hIL-6) into the culture medium of *A. nidulans* directed by the *A. niger* glaA signal peptide significantly improved extracellular levels from 0.4 x 10^3 U/mL to 11.8 x 10^3 U/mL (Su et al., 2012; Watts et al., 1983). Brandhorst and Kenealy (1995) proved that the extracellular production of heterologous restrictocin with its own signal peptide in *A. nidulans* was higher than secretion of restrictocin when fused to the signal peptide from an endogenous glucoamylase. The examples prove that even though different signal peptides are capable of directing the translocation of the same heterologous proteins, they do not all have the same efficiency (Su et al., 2012). Therefore, the analysis of how different signal peptides (homologous and heterologous) influence the secretion of heterologous protein, may be essential for experiments aimed at increasing secreted protein levels of filamentous fungi (Su et al., 2012).

The amino acid sequences of different signal peptides have been shown to affect their efficiency. Therefore systematic selection and engineering of

available signal peptides based on their amino acid sequence, may lead to increased yields of produced heterologous proteins in filamentous fungi (Chang & Sheu, 2010; Su et al., 2012).

1.2.5 Reporter proteins

Reporter genes and encoded proteins are used to investigate a number of biological processes, especially to monitor the regulation of transcription. They are defined as genes that are joined to a regulatory sequence, which when introduced into a biological system, produces an easily measurable signal output (usually a chromogenic product), upon modulation of expression (Ghim et al., 2010; Wood, 1995). This can be used to compare expression of different promoters or to compare efficiency of different signal peptides fused to the reporter protein. Some widely used reporter proteins are listed in Table 1.6.

Table 1.6 Commonly used reporter genes and their encoded proteins (Craney et al., 2007; Ghim et al., 2010; Rollins et al., 2001).

Reporter gene	Protein encoded
lacZ (E. coli)	β-galactosidase
gfp (Aequorea victoria)	Green fluorescent protein
gus (E. coli)	β-glucuronidase
luc (Photinus pyralis)	Luciferase
DsRed (Discosoma spp.)	DsRed protein variants
luxCDABE (Photorhabdus luminescens)	Bacterial luciferase

In *E. coli*, β-galactosidase hydrolyzes lactose to produce galactose and glucose. However, this enzyme can hydrolyze other substrates, including chromogens *O*-nitrophenyl β-galactopyranoside (ONPG), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and 3,4-cyclohexenoesculetin-β-D-galactopyranoside (S-gal), which produce yellow, blue and black products,

respectively (James et al., 1996). This reporter protein is applied in the X-gal dependent blue/white colony assay, commonly used for the screening of transformants (Ghim et al., 2010). The two major limitations in using this reporter protein are the use of costly and potentially toxic chemicals needed for the assay (Ghim et al., 2010).

GFP is a fluorescent protein which was originally isolated from marine invertebrates *Aequorea victoria* and *Renilla renifonnis* (Ward, 1979). Studies over the years demonstrated that GFP could be used to study the transcriptional activities within a wide range of hosts (Rollins et al., 2001). One of the traits that make the GFP of the jellyfish *A. victoria* a favorable reporter protein is that formation of the fluorophore is cell-autonomous (Cubitt et al., 1995; Ward, 1979). This means that gene expression in individual cells can be visualized directly therefore, processes such as cell lysis, biochemical analysis, tissue distortion and staining can be avoided (Chiu et al., 1996). Over the years several amino acid substitutions in the chromophore region of the GFP have been done and has led to improved fluorescence yields (Rollins et al., 2001). The most commonly used GFP mutants carry the S65T substitution which leads to strongly increased fluorescence (Chiu et al., 1996; Heim et al., 1995). In filamentous fungi, the GFP was first expressed as a vital marker in *Ustilago maydis* (Spellig et al., 1996), followed by expression for the study of protein localization and mitosis in *A. nidulans* (Fernandez-Abalos et al., 1998). Leroch et al. (2011) developed and used a codon-optimized version of the gray mold fungus *Botrytis cinerea* enhanced GFP (eGFP)-encoding gene (Bcgfp) to improve GFP and therefore gene expression in *B. cinerea*. GFP was used as a reporter protein in a study to identify the effect of functional nuclear localization signals (NLSs) of the transcription factor *AreA* in *A. nidulans,* on

utilization of nitrogen sources or *AreA*-dependent gene expression (Hunter et al., 2014). The examples show that GFP is successfully expressed and used as a reporter protein in filamentous fungi.

1.2.6 Transformation of filamentous fungi

Transformation systems are used to introduce foreign DNA into the cell. They make it possible for molecular manipulation of gene expression and function in different organisms (Rivera et al., 2014). The development of transformation techniques for filamentous fungi has lagged behind that of the model organisms *E.coli* and *S. cerevisiae* used for heterologous gene expression (Lubertozzi & Keasling, 2009). This was due to the complex multicellular morphology of filamentous fungi, their thick chitinous cell walls and lack of self-propagating plasmids (Lubertozzi & Keasling 2009). Since the first reports of gene cloning in 1982 (Kinghorn & Hawkins, 1982) and transformation of *Aspergillus* in 1983 (Tilburn et al., 1983), several transformation methods have been explored and developed for filamentous fungi (Su et al., 2012). Table 1.7 lists the transformation methods that have been developed with their advantages and disadvantages. However the most commonly used method of transformation is the PEG-mediated transformation whereas, the most efficient and successful method of transformation is *Agrobacterium tumefaciens*-mediated transformation.

Table 1.7 Advantages and disadvantages of transformation methods used in filamentous fungi (Rivera et al., 2014; Su et al., 2012).

Method	Advantage	Disadvantage
PEG-mediated protoplast fusion	No requirement for special equipment.	Optimization of making and regenerating the protoplasts is time consuming.

Electroporation	Can be applied to any fungi *in vivo* or frozen. Efficient protocols: simple fast and easily optimized.	Depends on the electrophysiological characteristics of the fungus. Low transformation efficiency. Medium cost. Need electroporation.
***Agrobacterium tumefaciens*-mediated transformation (ATMT)**	Possibility of recovering T-DNA flanking sequences by PCR-based techniques. Alleviates protoplast preparation.	The transformation process is time consuming.
Biolistic transformation	Simple. No pre-treatment of the cell wall required. Independent of the physiological properties of the fungi. Transformation with multiple transgenes is possible.	DNA can be damaged. Produces multiple copies of introduced genes. Complex protocols due to projectile preparation efficiency. The equipment is expensive.

1.2.6.1 Protoplast transformation

Poly-ethylene-glycol (PEG)-mediated protoplast fusion was discovered in 1979 and has become one of the most commonly used method for fungal transformation (Colacino et al., 2012). The protocol for this method of transformation includes three steps: firstly the preparation of protoplasts, secondly mixing of the DNA with the protoplasts and thirdly, the regeneration of cell walls of the protoplasts in an appropriate regeneration medium (Su et al. 2012).

Protoplasts are cells from which the walls have been removed either mechanically or enzymatically. Lysing enzymes (Prabavathy et al., 2006; Wang & Xia, 2011) are used for the digestion of the cell walls to form protoplasts.

Protoplasts can be prepared from different fungal structures that include intact conidia, germinating conidia, young hyphae or yeast-like cells. The choice of cell type used for preparation of protoplast depends on the fungus (Ruiz-Díez, 2002; Tanguay and Breuil, 2003). A PEG based buffer is used to perforate the cell membrane allowing insertion of exogenous DNA into the cell. The

protoplasts and DNA are then mixed to allow DNA adsorption and uptake. In the final step, the protoplast/DNA mixture is spread onto regeneration media for regeneration of the cell wall (Masani et al., 2014).

Some factors influencing the successful preparation of protoplasts include the choice of lysing enzyme for degradation of the cell walls, the quality and concentration of PEG and the incubation time for conidia germination (Masani et al., 2014).

1.2.6.2 Agrobacterium tumefaciens mediated transformation (ATMT)

This method was originally used in the transformation of plants (Abuodeh et al., 2000). The Gram-negative soil bacterium *A. tumefaciens* is extensively used to transfer part of its Ti tumor-inducing plasmid transfer DNA (T-DNA), into plant cells. Subsequently, the T-DNA is integrated into the genome of the target cells by random induction of a set of virulence (*vir*) genes located on the Ti plasmid (Su et al. 2012). The accuracy of T-DNA integration has extended the use of this method to facilitate gene targeted integration in filamentous fungi making ATMT the most successful method for transforming filamentous fungi (de Boer et al., 2013). ATMT has been applied successfully to transform fungi such as *Penicillium chrysogenum, A. awamori, A. niger, F. venenatum, T. reesei, Colletotrichum gloeosporioides* and *N. crassa* (de Boer et al., 2013; de Groot et al., 1998). A major advantage of ATMT is flexibility that allows use of protoplasts, hyphae or spores as target material for transformation (Mullins et al., 2001).

1.3 The effect of growth conditions on protein production

Developments in genetic engineering and molecular biology have adavanced the use of filamentous fungi as hosts for the production of heterologous proteins. However, low production levels in the range of milligrams per liter of culture medium (Gouka et al., 1997) have been reported for most non-fungal proteins (of mammalian, bacterial or plant origin). Various efforts have been made in strain and fermentation process improvements (Punt et al., 2002) to increase product yields. The optimization of bioprocessing parameters has significantly enhanced the production of heterologous proteins in filamentous fungi (Wang et al., 2003). Bioprocess parameters such as the culture medium composition, temperature, pH, agitation and aeration have been reported to have an influence on the yield and quality of product (Çelik, & Çalık, 2012).

In submerged cultures, two forms of fungal morphology dominate, these are pellets and freely dispersed mycelia. Different fungal morphologies are required for the production of specific fungal products. For example, free mycelia are required for the production of penicillin from *Penicillium chrysogenum*, whereas pellets are required for the production of citric acid from *A. niger* (Vecht-Lifshitz et al., 1990). The bioprocessing parameters such as inoculum level, initial pH, agitation, medium composition and temperature also influence fungal morphology (Metz and Kossen, 1977). Due to the complex morphology of filamentous fungi, bioprocessing parameters cannot be generalized, therefore process optimization is crucial for improved protein production titres.

1.4 Aims and objectives

Filamentous fungi have a natural ability to produce and secrete high amounts of metabolites and proteins. For this reason they have become a favored host

for the expression of heterologous proteins. The regulation of protein expression in filamentous fungi mainly occurs on the transcriptional level. Therefore, the use of strong fungal transcriptional regulatory regions is important for the development of efficient expression systems for heterologous protein production. In this study an indigenous fungal isolate with the natural ability to produce and secrete high levels of protein was identified. The aim is to identify the isolate's strong transcriptional elements for future use in developing a host-vector system for high level protein production. To achieve this aim the specific objectives of this study were to:

1. Screen the CMW culture collection of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, to identify an indigenous isolate with the ability to produce and secrete high levels of proteins.
2. Identify the extracellular protein.
3. Identify and isolate the gene encoding the highly expressed protein (HEP).
4. Adjust the growth conditions to further improve protein production.
5. Identify (*in silico*) the up-stream and down-stream transcriptional elements flanking the gene encoding the HEP.

CHAPTER TWO

2.1 Materials and Methods

2.1.1 Chemical reagents

All the chemicals used in this study such as salts, sugars, organic solvents, and alcohols, were of analytical grade, and were obtained from different suppliers. All buffers and solutions not described within the text are listed in Appendix B.

2.1.2 Antibiotics

All antibiotics were obtained from Sigma-Aldrich, USA. Stock solutions were prepared by completely dissolving the antibiotics in the appropriate solvent and filter sterilizing using 0.22 µm filters (Table 2.1). Sterilized antibiotics were added to media as needed once media was cooled (~55°C). The antibiotics were stored in 500 µL aliquots at -20ºC.

Table 2.1 Antibiotics used in this study.

Antibiotic	Solvent	Stock concentration
Chloramphenicol (Chlr)	Ethanol	10 mg/mL
Ampicillin (Amp)	Milli-Q H2O	100 mg/mL
Streptomycin (Strep)	Milli-Q H2O	100 mg/mL
Cefotaxime (Cef)	Milli-Q H2O	250 mg/mL
Tetracycline (Tet)	Milli-Q H2O	12.5 mg/mL
Kanamycin (Kay)	Milli-Q H2O	30 mg/mL

2.1.3 Enzymes

Enzymes used in this study, their function and source are listed in Table 2.2.

Table 2.2 List of the enzymes used in the study, their function and source.

Enzyme	Function	Supplier

EcoR1	Restriction enzyme digestion	Fermentas Life Sciences, Lithuania
Dream*Taq*™ DNA polymerase	PCR	Fermentas Life Sciences, Lithuania
T4 DNA ligase	Ligation	Fermentas Life Science, Lithuania
One ™ *Taq* DNA polymerase	Site-Finding PCR	New England BioLabs, UK
Q5™ High-fidelity polymerase	Site-Finding PCR	New England BioLabs, UK
Fast link ligase	Ligation	Fermentas Life Sciences, Lithuania

2.1.4 Strains, vectors and kits

The bacterial strain, vector and kits used in this study are provided in Table 2.3.

Table 2.3 Bacterial strain, vector and kits used in this study.

	Characteristics	Source
Bacterial Strain	**Genotype**	
E.coli JM109	*endA1, recA1, gyrA96, thi, hsdR17 (rk^-, mk^+), relA1, supE44, Δ(lac-proAB), [F´, traD36, proAB, laqIqZΔM15]*	Promega, USA

Vectors/plasmid	**Source**
pGEM ®T-easy system	Promega, USA

Kit	**Source**
BigDye Terminator v3.1 Cycle Sequencing Kit	Life Tchnologies
Zymo Research Fungal/Bacterial DNA MiniPrep™ kit	Zymo Research, USA
PureLink ® Quick Gel Extraction Kit	Life Technologies, Invitrogen, USA
Fast-Link QIAprep® Spin Miniprep Kit ™ DNA	Epicentre, USA
Ligation Kit	QIAGEN, USA

2.2 Media, fungal growth and maintenance

The indigenous fungal isolates screened in this study were obtained from the CMW fungal culture collection of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa and are listed in Appendix A. The isolates were cultured at 28°C on 2% (w/v) malt extract agar (MEA),

sometimes supplemented with antibiotics to abate and treat contamination (Zipfel et al., 2006). Unless otherwise stated liquid cultures were routinely grown at 28°C on a rotary shaker 160 rpm. For long term storage, agar blocks (5 mm^2) of actively growing mycelia were stored in 1 mL of sterile H_2O at 4°C. Liquid cultures of harvested spore suspensions were stored in 500 µL aliquots of 20% (v/v) glycerol at -80ºC for long term storage. All media used in this study are listed in Appendix B.

2.3 Screening for high level extracellular protein titres

2.3.1 TCA protein concentration

Fungal isolates were screened for extracellular protein production by aseptically inoculating 250 mL of sterile malt extract (ME) or malt yeast extract (MYE) liquid media (Chatzifragkou et al., 2011; Santamauro et al., 2014)., with small agar blocks of mycelia actively growing on MEA. Five days post inoculation, the cultures (2 mL) were centrifuged at 11,000 x *g* for 5 minutes in an Eppendorf 5424/R centrifuge (Eppendorf, Germany) and the supernatants were stored at -20ºC for further analysis.

Extracellular protein from the culture supernatant was routinely concentrated with Trichloroacetic acid (TCA). To 1 mL of culture supernatant, 333 µL of 20% (w/v) TCA was added, followed by incubation on ice for 1 hour. Protein was precipitated by centrifugation at 13,000 x *g* for 10 minutes at 4ºC and supernatants were carefully discarded without dislodging the pellet. The pellets were then washed with 500 µL ice cold acetone and centrifuged at 13,000 x *g* for 5 minutes at 4ºC. Acetone was discarded carefully and the pellets were air

dried at room temperature to remove residual acetone. Protein was resuspended in 25 µL of 1X SDS sample buffer (Appendix B.6.2) and incubated for 5 minutes at 100ºC. The protein samples were centrifuged at 11,000 x *g* for 2 minutes to precipitate any remaining impurities and 20 µL of each sample was analysed on SDS-PAGE gels.

2.3.2 SDS-PAGE gel electrophoresis

SDS-PAGE gel electrophoresis was carried out according to the method of Laemmli (1970) with a BioRad Mini-PROTEAN® Tetra Cell system. The SDS-PAGE gels were cast in glass cassettes using a casting stand. Gels were cast by first loading 3.5 mL of 12% (v/v) resolving gel into the glass cassette. Distilled sterilized H_2O (1 mL) was immediately loaded on top of the stacking gel to create a smooth and even gel interface on top of the resolving gel. The gel was allowed to polymerize for 30-45 minutes at room temperature before discarding the H_2O. After discarding the H_2O 1.5 mL of 4% (v/v) stacking gel was poured on top of the resolving gel until the top of the short plate was reached. A comb was inserted between the spacers and the stacking gel was allowed to polymerize for 30 minutes. The comb was gently removed and 20 µL of each protein sample was loaded in each lane of the gel.

SDS-PAGE gel electrophoresis was carried out at room temperature for 35 minutes at 200 Volts (V) in 1X running buffer (Appendix B). A protein ladder (Page Ruler Prestained Protein Ladder, Thermo Scientific, USA) was used to analyze the size of the protein. The gel was stained with Coomassie stain by heating in a microwave oven for 30 seconds followed by slow shaking (60 rpm) at room temperature for 1.5 hours. The gel was destained with destaining solution by heating in a microwave oven for 30 seconds followed by slow

shaking at room temperature until the gel was properly destained. Gel analysis and imaging was done using a Gel Doc system (BioRad Gel Doc ™ XR, BioRad, USA).

2.4 Identification of fungal isolate using internal transcribed spacer (ITS) region sequencing

The fungal isolate that produced the most prominent bands on the SDS-PAGE gels was subjected to ITS sequencing for confirmation of identity.

2.4.1 Chromosomal DNA isolation and quantification

Chromosomal DNA was isolated from isolate CMW 20676 using the PrepMan® Ultra Sample Preparation Reagent (Applied Biosystems, USA) according to the manufacturer's specifications. DNA was extracted from mycelia actively growing on MEA (5 days). An equivalent of a loopful of mycelia was transferred to a 2 mL tube containing 100 µL of reagent. The sample was vortexed for 10-30 seconds and incubated at 100°C for 10 minutes. The sample was cooled down to room temperature (2 minutes), centrifuged at 13,000 x *g* for 2 minutes to sediment any impurities and the supernatant (50 µL) was transferred to a new 1.5 mL eppendorf tube.

The amount of DNA in the sample was determined using a Nano drop ND-1000 (Fermentas, Canada) by first blanking the spectrophotometer with the PrepMan® Ultra Sample Preparation Reagent. DNA (2 µL) and the concentration of the sample was determined.

2.4.2 Agarose gel electrophoresis

The quality of the extracted DNA was analysed visually by agarose gel electrophoresis.

For preparation of 1% (w/v) agarose gel, 0.5 g of agarose (SeaKem® LE Agarose, Lonza, USA) was dissolved in 50 m L of 1X TAE buffer (Appendix B) by boiling in a microwave oven for 3-5 minutes with swirling at 1 minute intervals. After allowing the agarose to cool (~ 60°C) the agarose gel slab was cast using the Ultra-Violet transparent plastic (UVTP) tray set on the gel stage. Isolated chromosomal DNA was mixed with 6X GelRed™ (1000X , Biotium, USA) prepared in loading buffer (Appendix B) in the ratio 3:1 of sample to gel red . The agarose gel was electrophoresed at 90 Volts for 45 minutes in 1X TAE buffer at room temperature using the BioRad Sub-Cell® GT gel electrophoresis system. The gel was visualized using the BioRad Gel Doc system and viewed with a UV Trans-illuminator.

2.4.3 ITS gene amplification and sequencing

The isolated chromosomal DNA was used as template in a 50 µL PCR reaction, to amplify the ITS region using ITS lF and ITS 4 primers (Table 2.4) according to (Alvarado et al., 2010).

The PCR reaction mix was a follows:

PCR reaction mix

ITS1 F (10 µM)	2.5 µL
ITS4 (10 µM)	2.5 µL
Dream Taq DNA polymerase (5 U/µL)	0.05 µL
1X Taq buffer	5 µL
dNTP's (10 mM)	5 µL
4% (w/v) BSA	2.5 µL
6% (w/v) DMSO	3 µL

DNA Template (116.9 ng/µL) 2 µL

Volume was made up to 50 µL with nuclease free Milli-Q H_2O.

Table 2.4 Primers used for ITS amplification.

Primers	Sequence
ITS 1F	5'-CTTGGTCATTTAGAGGAAGTAA-3'
1TS 4	5'-TCCTCCGCTTATTGATATGC-3'

Amplification of DNA was done using a T100™ Thermal Cycler (BioRad, USA). The thermocycling conditions for the PCR reaction are outlined in Table 2.5 .The PCR product was visualized on a 1% agarose gel to confirm reaction had taken place and to verify DNA bands of the expected band size.

Table 2.5 PCR thermo cycling conditions for ITS gene amplification.

Step	Temperature (ºC)	Time	Number of cycles
Initial denaturing	95	5 minutes	1
Denaturing	94	1 minute	
Annealing	55	50 seconds	28
Extension	72	50 seconds	
Final extension	72	5 minutes	1

The amplicon was sequenced in a forward and reverse direction using ITS lF and ITS 4 primers respectively. Sequencing reactions were done using the Big Dye ™ Terminator v3.1 cycle sequencing premix kit (Applied Biosystems, USA)

with protocols supplied by the manufacturer. The sequencing PCR reaction mix contained:

Sequencing reaction mix

5x Sequencing buffer	2.4 µL
Big Dye premix	0.5 µL
DNA template (116.9 ng/µL)	4 µL
Forward/Reverse primer (10 µM)	1 µL

The reaction mixture was made up to 12 µL with nuclease free Milli-Q H_2O.

DNA sequences were generated using an ABI PRISM 3100 capillary sequencer (Perkin Elmer, USA). The cycling parameters are outlined in Table 2.6.

Table 2.6 Thermo cycling conditions for the sequencing PCR.

Step	Temperature (ºC)	Time	Number of cycles
Initial denaturing	94	5 seconds	1
Denaturing Annealing Extension	94 54 60	10 seconds 10 seconds 4 seconds	25

The PCR product (12 µL) was mixed with 16 µL of 100% ethanol and 2 µL of 3M NaOAC in a 600 µL PCR tube. The reaction was centrifuged at maximum speed (15,000 x *g*, Eppendorf 5424/R centrifuge) at 4ºC. The supernatant was discarded carefully without touching the walls of the tube and the precipitate was washed twice with 70% (v/v) ethanol by centrifugation at 21,130 x *g* (4ºC) for 5 minutes. The supernatant was discarded and the tube was incubated for 2-3 minutes at 90ºC with the lid open to remove residual ethanol. The samples were stored at -20ºC until sequencing.

2.4.4 Sequence analysis

Sequences generated were screened and trimmed to remove poor quality sequence using BioEdit sequence alignment editor Version 7.2.3.0(http://www.mbio.ncsu.edu/bioedit/bioedit.html). Consensus sequence searches was performed using the National Centre for Biotechnology Information (NCBI) BLAST standard nucleotide-nucleotide basic local alignment search tool (http://www.ncbi.nlm.nih.gov/BLAST/)]. Sequence-based identification was determined from the entry with the highest bit score listed in the BLAST search with E value closest to zero.

2.5 LC-MS/MS peptide identification

To identify the highly expressed protein, bands were excised from the SDS-PAGE gel, and sent for liquid chromatography-mass spectrometry (LC-MS/MS) at the Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa. Spectral peaks generated from LC-MS/MS were subjected to PEAKS proteomics mass spectrometry data analysis software (http://www.bioinfor.com/peaks/features/overview.html) for effective peptide and protein identification. Acquired peptide sequences were further analysed through NCBI BLAST.

2.6 Optimization of fungal growth conditions for improved protein production

O. phasma 20676 was cultured under different growth conditions to determine the optimum conditions required for high level protein production. Protein production levels were determined semi-quantitatively by SDS-PAGE gel analysis.

2.6.1 Inoculum preparation

Mycelium inoculation was through transferring an agar block (5 mm^2) of actively growing mycelia from a plate (MEA, 28ºC, 5 days post inoculation) into a shake flask (250 mL) containing MYE (50 mL).

Fungal spores were cultivated by aseptically inoculating a 5 mm^2 agar block of actively growing mycelia into a 250 mL flask (250 mL) containing 50 mL sterile liquid MYE medium. The culture was grown at 28°C with shaking at 160 rpm

for 5 days. Spores were then harvested by aseptically filtering the culture through three layers of sterile cheese cloth. Spores were quantified using a haemocytometer under a light microscope (UB 203i biological microscope) at 40X magnification. Spore suspension (10 µL) was transferred to a haemocytometer and spores were counted in at least four of the smallest squares. The total number of spores per mL was obtained using the equation:

Number of spores x 400 x 10^4 x dilution factor = number of spores per mL

The standardized spore inoculum concentration required for investigating the different growth parameters for high level protein production by *O. phasma,* was determined by inoculating shake flasks containing sterile MYE (50 mL) with 10^3 to10^8 spores/mL. Extracellular protein from culture supernatants 5 days post inoculation was concentrated with TCA (Section 2.3.1) and separated on SDS-PAGE gels (Section 2.3.2). The protein profiles were compared to determine the best inoculum concentration for high level extracellular protein production by *O. phasma* 20676.

The identified optimal concentration (10^7spores/mL) was then used as inoculum for all future experiments to investigate the growth conditions for improved protein production by *O. phasma* 20676.

2.6.2 Growth medium for improved protein production by *O. phasma* 20676

Five different liquid media; ME, MYE, potato dextrose broth (PDB), complete medium (CM) and nutrient broth (NB) (Appendix B) were tested for improved protein production by *O. phasma* 20676. Proteins in the culture supernatants were concentrated and compared on SDS-PAGE gels to determine the best medium for high level extracellular protein production by *O. phasma* 20676.

2.6.3 Effect of different carbon sources on protein production by *O. phasma* 20676

The effect of different carbon substrates [locust bean gum (LBG), glucose and acacia guar gum (AGG)] on protein production by *O. phasma* was determined by culturing *O. phasma* 20676 in MYE (50 mL) adjusted with 1% (w/v) LBG, glucose or AGG as main carbon sources. Extracellular protein from culture supernatants five days post inoculation was concentrated and compared on SDS-PAGE gels.

After identification of LBG as the preferred substrate, the optimum concentration of LBG for high level protein production was determined. *O. phasma* 20676 was cultured in shake flasks containing MYE (50 m L) adjusted with different concentrations of LBG [0.2%, 0.4%, 0.6%, 0.8% or 1% (w/v)]. Five days after inoculation, extracellular protein from culture supernatants was concentrated and protein samples (20 µL) compared on SDS-PAGE gels to determine the optimal LBG concentration for high level extracellular protein production by *O. phasma* 20676.

2.6.4 Effect of pH on protein production by *O. phasma* 20676

O. phasma 20676 was grown in MYE liquid media adjusted to pH 5, pH 6.5, pH 7, pH 8, pH 9 and pH 10. The pH of the cultures was monitored aseptically with pH test strips (pH range 0-14, Sigma- Aldrich, St Louis, USA) during the five days of incubation. Extracellular protein from day five culture supernatants was concentrated with TCA and compared on SDS-PAGE gels to determine the

optimal culture medium pH for high level extracellular protein production by *O. phasma* 20676.

2.6.5 Effect of temperature on protein production by *O. phasma* 20676

O. phasma 20676 was cultured in MYE medium at different temperatures (25°C, 28°C, 30°C, 32°C and 34°C). Extracellular protein from day five culture supernatants was concentrated with TCA and the different protein profiles compared on SDS-PAGE gels to determine the preferred growth temperature for high level extracellular protein production by *O. phasma* 20676.

2.6.6 Optimal conditions

After investigating and determining the optimum culture conditions for improved protein production, *O. phasma* 20676 was cultured under the optimized growth conditions (spore inoculum concentration 1 x 10^7 spores/mL, MYE medium adjusted with 1% LBG, pH 6 at 28ºC). Protein in the culture supernatant five days post inoculation was concentrated and separated on SDS-PAGE.

2.7 PCR amplification of manannase (*man*) gene from *O. phasma* 20676

2.7.1 Genomic DNA isolation

Genomic DNA from *O. phasma* 20676 was isolated from 200 mg of wet fungal biomass harvested five days post inoculation using a commercial kit (Zymo Research Fungal/Bacterial DNA MiniPrep™ kit, Zymo Research, USA). Fungal biomass was re-suspended in 750 µl of Lysis solution in a ZRBashingBead™

lysis tube. The tube was secured in a bead beater (PowerLyzer® 24 Bench Top Bead-Based Homogenizer, MO BIO Laboratories, Inc, USA) and processed at maximum speed for 5 minutes and then centrifuged at 10,000 x *g* for 1 minute. The supernatant (400 µL) was transferred to a Zymo-Spin™ IV Spin Filter placed in a collection tube and centrifuge at 7,000 x *g* for 1 minute. Fungal DNA binding buffer (1200 µL) was added to the filtrate in the collection tube and 800 µL of the mixture was transferred to a Zymo-Spin™ IIC column in a collection tube. The tube was centrifuged at 10,000 x *g* for 1 minute and the flow through from the collection tube was discarded. The remaining mixture was transferred to the same Zymo-Spin™ IIC column in a collection tube. The tube was centrifuged at 10,000 x g for 1 minute and the flow through discarded. DNA pre-wash buffer (200 µL) was added to the Zymo-Spin™ IIC column in a new collection tube and centrifuged at 10,000 x *g* for 1 minute. Fungal DNA wash buffer (500 µL) was added to the Zymo-Spin™ IIC column and centrifuged at 10,000 x *g* for 1 minute. The Zymo-Spin™ IIC column was transferred to a clean 1.5 mL eppendorf tube and 50 µL of Milli-Q H_2O was directly added to the column matrix. The tube was centrifuged at 10,000 x *g* for 30 seconds to elute the DNA.

The quantity of DNA in the sample was determined using the Nano drop spectrophotometer (Section 2.4.1) and the quality of DNA was verified visually by agarose gel electrophoresis (Section 2.4.2).

2.7.2 PCR amplification of internal *man* gene fragment

Isolated genomic DNA was screened by PCR to obtain an internal gene fragment using the MANF and MANRev primers (Table 2.7) and the cycling parameters outlined in Table 2.8.

Table 2.7 Primers used in this study.

Primers	Sequence	Source
MANF	5'-CAACGACGTCACGACCAAGCC-3'	This study
MANRev	5'-CTCGTTGGCCAGCTCCCAGGC-3'	This study

The PCR reaction mix was asfollows:

PCR reaction mix

MANF (10 µM)	1.25 µL
MANRev (10 µM)	1.25 µL
10x Taq buffer	2.5 µL
dNTP's (10 mM)	2.5 µL
DNA template (209 ng/µL)	1 µL
Dream Taq polymerase	0.125 units

The mixture was made up to 25 µL with nuclease free Milli-Q H_2O.

Table 2.8 Thermo cycling conditions for PCR amplification of internal gene fragment.

Step	Temperature (ºC)	Time	Number of cycles
Initial denaturing	95	5 minutes	1
Denaturing	95	30 seconds	28
Annealing	58	50 seconds	
Extension	72	50 seconds	
Final extension	72	5 minutes	1

After analysis of DNA quality on agarose gel, the PCR product was purified and concentrated using the PureLink® Quick Gel Extraction and PCR Purification Combo kit (Life technologies, Invitrogen, USA) according to manufacturer's specifications. PCR product was mixed with binding buffer (B2) in the ratio 1:4 of PCR product (50-100 µL) to buffer. The sample was mixed gently by inverting the tube 3-6 times. The sample was transferred to a PureLink® Clean-up

Column in a wash tube and centrifuged at 11,000 x *g* for 1 minute. The flow-through was discarded and the column was placed back into the wash tube. The silica membrane in the column was washed by adding 650 µL of wash buffer (W1) with added ethanol to the column and centrifuging for 1 minute at 11,000 x *g*. The flow-through was discarded; the column placed back into the wash tube and the silica membrane was dried by centrifugation for 2 minutes at 14,000 x *g* to completely remove residual wash buffer. DNA bound to the silica membrane on the column was eluted by placing the column into a new 1.5 mL tube and adding 30 µL of nuclease free Milli-Q H_2O. The sample was incubated at room temperature (18–25 °C) for 1 min and centrifuged for 1 min at 14,000 x *g*. The purified DNA was stored at 4°C for immediate use or at -20°C for long term storage.

2.8 Cloning and sequencing of internal *man* gene fragment

2.8.1 Preparation of competent *E. coli* cells

Electro-competent *E. coli JM 109* cells were prepared according to methods of Nickoloff & Miller (1995) and Ausubel et al. (1987). *E. coli* cells were aseptically transferred from a frozen stock to a fresh Luria Bertani (LB) agar plate using an inoculation loop. The plate was incubated at 37ºC overnight. A single colony was inoculated into 5 mL LB liquid medium and incubated overnight at 37ºC in an orbital shaker (300 rpm). LB broth (200 mL) was inoculated to an OD_{600} of 0.05 from the overnight *E. coli* culture. The culture was incubated at 37ºC in an orbital shaker (300 rpm) until OD_{600} 0.4-0.5 was reached. The flask was rapidly transferred to an ice water bath and incubated on ice for 30 minutes with

constant swirling to ensure even cooling. The culture was transferred to pre-cooled centrifuge bottles and centrifuged (4,000 x *g* at 4ºC) (Eppendorf Centrifuge 5810 R) for 15 minutes. The supernatant was carefully discarded and the pellet was washed by gently re-suspending in 20 mL sterile ice cold 10% (v/v) glycerol. The cell suspension was centrifuged (4,000 x *g* at 4ºC) for 20 minutes and the supernatant was carefully discarded. The washing step was repeated for the second time before the pellet was gently re-suspended in a final volume (1 mL) of ice cold 10% (v/v) glycerol. The competent cells were stored as 50 µL aliquots at -80ºC.

2.8.2 DNA ligation

The pGEM ®-T Easy system (Promega, USA) was used for cloning of the amplified internal *man* gene fragment (Figure 2.1).

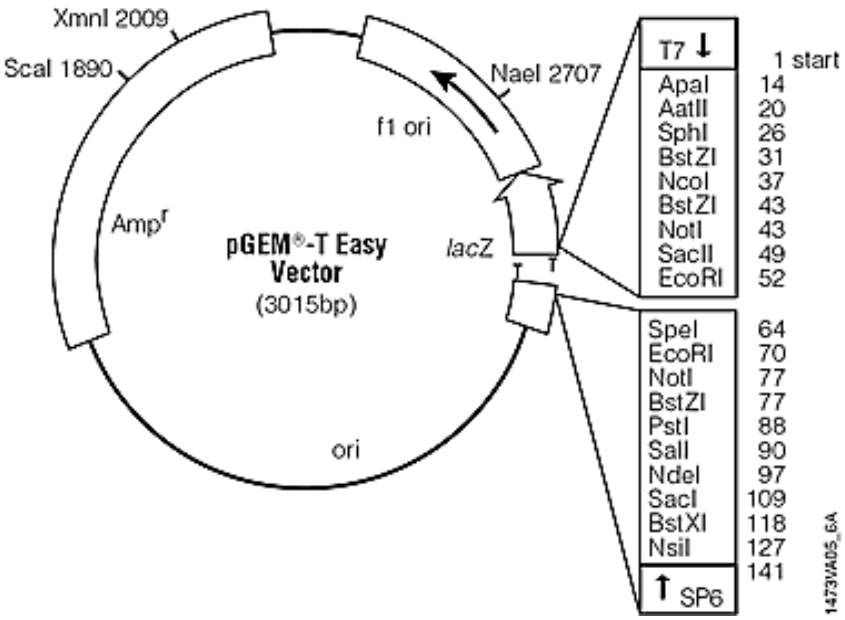


Figure 2.1 pGEM ®-T Easy Vector map and sequence reference points

The amount of vector to insert ratio for ligation was calculated using the formula:

$$\frac{\text{ng of vector} \times \text{kb}}{\text{kb of vector}} \text{insert: vector molar ratio} = \text{ng of insert}$$

Ligations were performed in sterile microfuge tubes with insert to vector: ratio of 1: 10.

The ligation reaction was set up as follows:

Ligation mix

2x ligation buffer	5 µL
pGEM®-T Easy vector (50 ng/µL)	1 µL
Insert (5 ng/µL)	1.5 µL
T$_4$ DNA Ligase (5 U/µL)	1 µL

Volume was made up to 10 µL will nuclease free Milli-Q H_2O.

The reaction was incubated at 4°C overnight.

2.8.3 Transformation of *E. coli* cells

E. coli cells were transformed by electroporation according to Dower et al. (1988). Electroporation was conducted using a BIORAD Gene Pulser™, and a BIORAD Pulse Controller. The machine was set to 200 Ohms (Ω), 25 micro-Faraday (µF) and 1.8 Volts (V) for the 0.1 cm cuvette used. Ligation reactions (1 µL) were transferred to 50 µL of electro competent cells that had been slowly thawed on ice, then placed in a 0.1 cm Gene Pulser® Cuvette (BIORAD) that had been chilled on ice. The cuvette was dried and the electrical pulse applied. Super optimal broth with catabolite repression (SOC) (950 µL) at room temperature was added to the transformed cells. The transformed cells were then transferred to a sterile 50 mL Falcon tube and incubated in an orbital shaker (175 rpm) at 37ºC for 1 hour. A 50-200 µL volume of the transformed

cells were then plated onto LB medium plates containing ampicillin (100 mg/mL), 40 µg/mL of isopropyl β-D-1-thiogalactopyranoside (IPTG) and 20 mg/mL of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) for blue/white colony selection.

2.8.4 Colony screening

Addition of X-gal and IPTG to the LB agar plates allowed for blue/white screening of colonies. Recombinant (white) colonies were screened by colony PCR according to the method by Güssow and Clackson (1989). Transformed single colonies were aseptically transferred into 50 µL of Milli-Q H_2O and incubated at 100ºC for 5 minutes. The cells were precipitated at 11,000 x *g* for 2 minutes. The supernatant (5 µL) was used in the PCR reaction as described in Section 2.7.2 using primers MANF and MANRev (Table 2.6).

2.8.5 Plasmid DNA extraction

DNA containing the specific gene fragment from transformed *E.coli JM109* cells was isolated using the QIAprep ® Spin Miniprep Kit (QIAGEN, USA), according to the manufacturer's specifications. Overnight cultures were prepared by aseptically inoculating white colonies into 5 mL of sterile LB supplemented with 5 µL of Ampicillin (100 mg/mL). The bacterial cells were harvested by centrifugation at 5,000 x *g* for 10 minutes at 25°C. The pellet was re-suspended in 250 µL of lysis buffer (P1) and transferred to a microfuge tube. Binding buffer (P2) (250 µL) was added and the tube was gently inverted 4-6 times for less than 5 minutes to mix the sample. Neutralizing buffer (N3) (350 µL) was added and the tube was gently inverted 4-6 times for mixing. The sample was centrifuged for 10 minutes at 10,000 x *g*. The supernatant was applied to a

QIAprep spin column in a 2 mL collection tube and centrifuged at 10,000 x *g* for 30-60 seconds. The flow through was discarded and the QIAprep spin column was washed with 500 µL of pre-wash buffer (PB). The flow-through from the washing step was discarded and the column was washed with 750 µL of wash buffer (PE) with centrifugation at 10,000 x *g* for 30-60 seconds. The flow-through was discarded and the column was centrifuged for an additional 1 minute to remove residual wash buffer. The column was placed in a clean 1.5 mL microfuge tube and DNA was eluted by adding 30 µL of nuclease free Milli-Q H_2O to the centre of the column. The column was incubated at room temperature (18°C- 25°C) for 1 minute, and centrifuged at 10,000 x *g* for 1 minute.

2.8.6 Restriction endonuclease digestion

Plasmid DNA was digested using a restriction enzyme in the following reaction:

Digestion reaction

Plasmid DNA (105 ng/µL)	3 µL
10 x Restriction buffer	2 µL
Restriction enzyme (10 U/µL)	1 µL

The mixture was made up to 20 µL with nuclease free Milli-Q H_2O and incubated at 37ºC for 1 hour. The digestion was analyzed on a 1% (w/v) agarose gel.

The plasmid DNA containing the *man* gene fragment was sequenced in the forward and reverse directions using the M13 forward and the M13 reverse primers (Table 2.9), respectively. The sequencing reaction was according to Section 2.4.3 and the cycling parameters for the reaction were as outlined in Table 2.7. Sequences generated were screened and trimmed to remove poor

quality sequences using BioEdit sequence alignment editor Version 7.2.3.0 (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Consensus sequence searches were performed using the National Centre for Biotechnology Information (NCBI) BLAST standard nucleotide-nucleotide basic local alignment search tool (http://www.ncbi.nlm.nih.gov/BLAST/)].

Table 2.9 Primers used in the PCR amplification reaction for sequencing the plasmid DNA.

Primers	Sequence	Source
M13 forward	5'-GTTTTCCCAGTCACGAC-3'	Promega, USA
M13 reverse	5'-CAGGAAACAGCTATGAC-3'	Promega, USA

2.9 Chromosome walking for obtaining complete *man* gene and regulatory regions

SiteFinding PCR was chosen as the technique for chromosome walking according to Tan et al. (2005).

2.9.1 Primer and Oligonucleotide design

Gene specific primers (Table 2.10) were designed within the sequenced gene fragment by using the acquired gene nucleotide sequence. Parameters considered when designing the gene-specific primers included:

- Primer length of between 18 and 23 base pairs
- Annealing temperature between 55°C and 76°C
- GC content of between 40% and 60%
- Avoiding di nucleotide repeats

The SiteFinder-1 oligonucleotide (Table 2.10) was designed according to Tan et al. (2005) with a *Not*I restriction enzyme site to facilitate cloning with commonly used vectors and four gene nucleotide sequences at the 3' end.

Table 2.10 Primers used for chromosome walking to obtain the complete *man* gene and its regulatory regions.

Primers	Sequence	Source
SiteFinder-1	5'CACGACACGCTACTCAACACACCACCTCGCACAGCGTC CTCAAGCGGCCGCNNNNNNGCTC-3'	This study
SFP 1	5'-CACGACACGCTACTCAACAC-3'	This study
SFP 2	5'-ACTCAACACACCACCTCGCACAGC-3'	This study
GSP1	5'-CTCGTTGGCCAGCTCCCAGGC-3'	
GSP 2	5'-CTTGGCCGCCGCTTGCACG-3'	This study
GSP 3	5'-CTCCGAGCCCGAGGCAGAC-3'	This study
GSP 4	5'-GCTCGACTTGGCCGCTCTTTATACG-3'	This study
GSP 4Fwd	5'-GGCCTACGTAACGGCGTTTGGC-3'	This study
GSP 5	5'-CGGAGGAGCAAGCGACCGC-3'	This study
GSP 5Fwd	5'-GTACATCAAGGCTGTGGTGACGCG-3'	This study

2.9.2 SiteFinding PCR amplification

The SiteFinding PCR amplification reaction involved three steps (Figure 2.2). In the first step (SiteFinding reaction), the double stranded DNA template was denatured allowing for the annealing and extension of the SiteFinder-1 (Table 2.10) at low temperature (Figure 2.2, diagram 2). The PCR product of the SiteFinding reaction was then used in the second step (Primary reaction) of the SiteFinding PCR amplification reaction. In the primary reaction the One Taq polymerase annealed to one strand of the target gene, and double stranded target molecules of varying sizes were generated by making use of the *SFP 1* and *GSP1* gene specific primers (Table 2.10) (Figure 2.2, diagram 3). The product of the primary PCR reaction was then used in the third and final step (Secondary reaction) of the SiteFinding PCR amplification reaction. In this reaction the target DNA was exponentially amplified with the use of the SFP 2,

GSP 2, GSP 3, GSP 4 and GSP 5 gene specific primers for upstream chromosome walking, and the SFP 2, GSP 4Fwd and GSP 5Fwd for downstream chromosome walking (Table 2.10) (Figure 2.2, diagram 3). Non-target gene amplification was suppressed by the formation of stem-loop structures (Tan et al. 2005). The PCR products generated by SFP 2 and the gene specific primers were purified by agarose gel electrophoresis and cloned into the pGEM®-T Easy vector (Figure 2.2, diagram 4). Clones were screened by colony-PCR and sequenced subsequently.

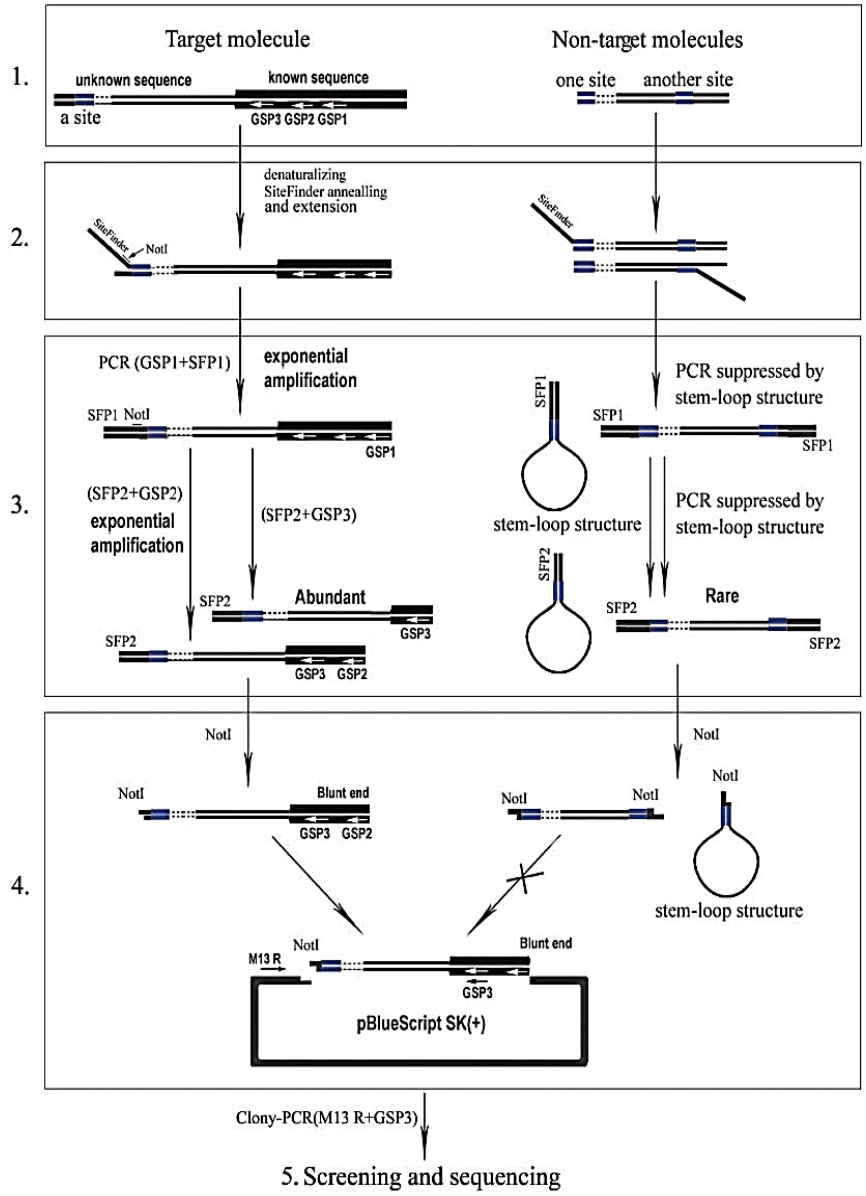


Figure 2.2 Schematic presentation of the SiteFinding PCR reaction for chromosome walking (Tan et al. 2005).

The constituents of the different SiteFinding PCR reactions are outlined in Table 2.11.

Table 2.11 Constituents of the SiteFiding PCR amplification reactions for chromosome walking.

Step 1 (SiteFinding reaction)		**Step 2 (Primary reaction)**		**Step3 (Secondary reaction)**	
5x One *Taq* standard buffer	12 µL	1x Taq buffer	1 µL	5x One Taq buffer	10 µL
dNTP's (10 mM)	6 µL	dNTP's (10 mM)	1 µL	dNTP's (10 mM)	5 µL
SiteFinder 1 (10 µM)	3 µL	GSP 1 (10 µM)	0.5 µL	GSP 2 (10 µM)	1 µL
OneTaq DNA polymerase	1.5 U	SFP 1 (10 µM)	1 µL	SFP 2 (10 µM)	1 µL
DNA template (209 ng/µL)	3 µL	SiteFinding PCR product (10 µM)	20 µL	Primary PCR product	1 µL
				One *Taq* DNA polymerase	1.25 U
Made up to 60 µL final volume with nuclease free Milli-Q H_2O		Made up to 25 µL final volume with nuclease free Milli-Q H_2O		Made up to 50 µL final volume with nuclease free Milli-Q H_2O	

The cycling parameters for the site finding PCR are outlined in Table 2.12. The PCR product from the secondary PCR reaction was analysed on a 1% agarose gel.

Table 2.12 Cycling parameters used for SiteFinding PCR.

Reaction	**Cycles**	**Thermal conditions**
Step 1 (SiteFinding)	1	92°C (2 min), 95°C (1 min), 25°C (1 min), ramp to 68°C over 3 min, 68°C (10 min)
Step 2 (Primary)	1	94°C (1 min)
	30	95°C (10 sec), 68°C (6 min)
	1	72°C (5 min)
Step 3 (Secondary)	1	94°C (1 min)
	30	95°C (10 sec), 68°C (6 min)
	1	72°C (5 min)

The DNA fragments were purified from the gel using the PureLink® Quick Gel Extraction and PCR Purification Combo kit (Life technologies, Invitrogen, USA) according to manufacturer's specifications. The purified DNA fragments were analysed on a 1% (w/v) agarose gel.

2.10 Cloning and sequencing of DNA fragments

The purified DNA fragments were cloned into the pGEM®-T Easy vector system (Figure 2.2). The vector to insert ratio was determined using the equation in Section 2.8. A commercial Fast-Link™ DNA Ligation Kit (Epicentre, USA) was used and ligations were performed in sterile microfuge tubes as follows:

Ligation mix

pGEM®-T Easy vector (50 ng/ µL)	0.5 µL
Insert (55 ng/µL)	5 µL
10x ligase buffer	0.75 µL
ATP (10 mM)	0.75 µL
Fast Link-DNA Ligase (2 U/µL)	0.5 µL
Final volume	7.5 µL

The reaction was incubated at room temperature for one hour or at 4ºC overnight. The ligation reaction (1 µL) was transformed into *E. coli* (JM109) cells and plated onto LB (Amp), X-gal and IPTG plates (Section 2.8.3).

2.10.1 Colony screening and sequencing

Recombinant (white) colonies were screened by colony PCR (Section 2.8.3) by aseptically transferring transformed single colonies into Milli-Q H_2O (50 µL) and incubating at 100ºC for 5 minutes. The cells were precipitated at 11,000 x *g* for 2 minutes and the supernatant (5 µL) was used in colony screening PCR reaction.

The reaction mixture contained:

PCR mixture

10x Taq buffer	2.5 µL
dNTP's (10 mM)	2.5 µL
Gene specific primer (10 µM)	1.25 µL
SFP 2 (10 µM)	1.25 µL

Plasmid DNA (209 ng/µL)	5 µL
Dream Taq DNA polymerase	0.125 U

Final volume was made up to 25 µL with nuclease free H_2O.

The sequencing reaction was done using SFP 2 and the gene specific primers (Table 2.10) and the cycling parameters for DNA amplification were as in Section 2.7.2 (Table 2.8). The PCR product was analysed on a 1 % agarose gel and positive clones were selected for sequencing (Section 2.8.4).

Plasmid DNA containing the *man* gene fragment was isolated from transformed clones using the QIAprep ® Spin Miniprep Kit (QIAGEN, USA) (Section 2.8.5). The plasmid DNA was digested (Section 2.8.6) and the digestion was analysed on a 1% agarose gel. The plasmid DNA containing the *man* gene fragment was sequenced in a forward and reverse direction using the SFP 2 and the gene specific primers (Table 2.10) respectively. The sequencing reactions were according to Section 2.4.3 and the cycling parameters for the reactions are outlined in Table 2.6. Sequences generated were edited using the BioEdit sequence editing software and further analyses to identify the 5' and 3' regulatory regions was carried out using online programs Match-1.0 Public and Patch 1.0 (http://www.gene-regulation.com/pub/programs.html).

CHAPTER THREE

RESULTS

3.1 Screening for high extracellular protein titres

For the identification of an indigenous fungal isolate with natural ability to secrete high levels of extracellular protein, a total of 140 fungal isolates (Appendix A) from the CMW fungal culture collection were screened for the production of extracellular protein in liquid ME and MYE media. Protein secreted into the culture supernatant was concentrated and separated on SDS-PAGE (Chapter 2.3.1). From the 140 fungal isolates screened, CMW 20676 (Figure 3.1) was chosen as the isolate of interest, based on the prominent bands observed on 12% SDS-PAGE which were indicative of high protein production capabilities.

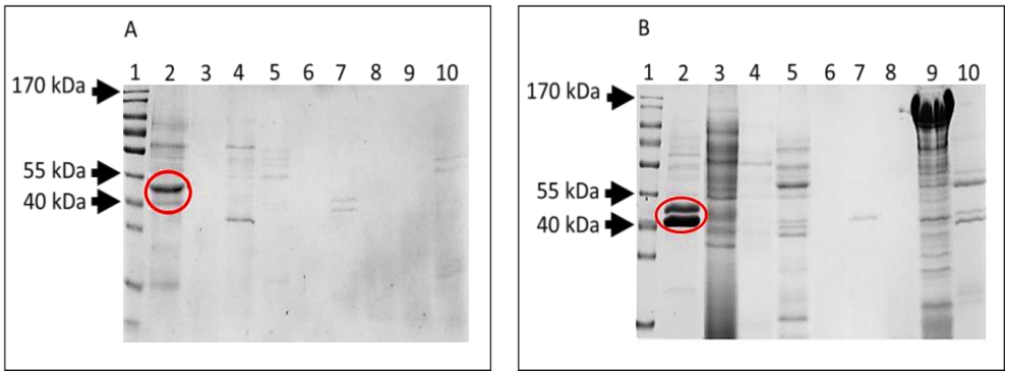


Figure 3.1 Comparison of extracellular protein production of some CMW fungal isolates cultured in (A) ME medium and (B) MYE medium at 28°C with shaking at 160 rpm five days after inoculation. Lane 1, pre-stained molecular weight marker; Lane 2, CMW 20676; Lane 3, CMW 2060; Lane 4, CMW 8401; Lane 5, CMW 30455; Lane 6, CMW 25878; Lane 7, CMW 29136; Lane 8, CMW 37322; Lane 9, CMW 24547 and Lane 10, CMW 18061. Each lane contains protein from 1 mL of culture supernatant.

3.2 ITS sequencing for strain confirmation

The identity of isolate CMW 20676 was confirmed by ITS sequencing (Chapter 2.4). Chromosomal DNA was isolated and used as template in the PCR amplification of the ITS region with primers ITS 1F and ITS 4 (Table 2.4, Chapter 2.4.3). PCR amplification generated an amplicon approximately 800 bp long (Figure 3.2).

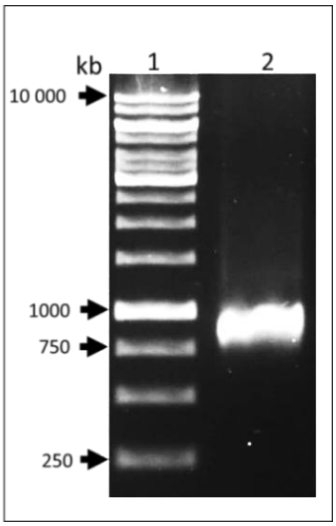
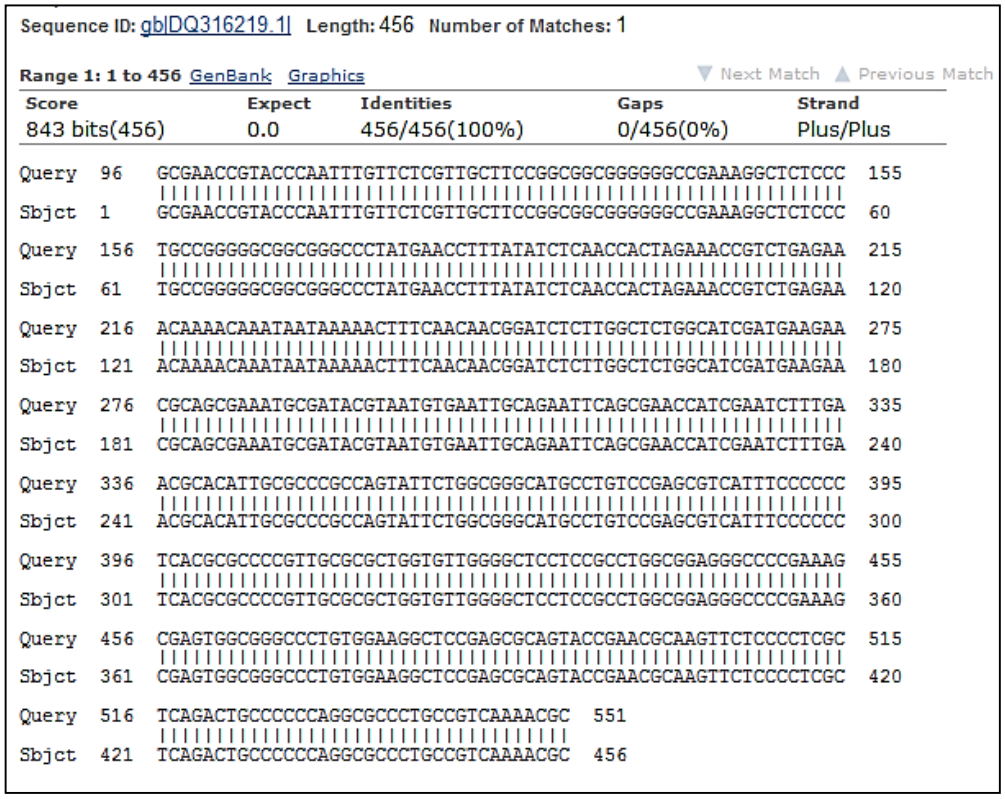


Figure 3.2 PCR analysis of the ITS region of isolate CMW 20676. Lane 1, 1kb DNA ladder and Lane 2, ITS amplicon of size 800 bp (116.9 ng/ µL).

The amplicon sequence generated was edited and trimmed to remove poor quality sequence, resulting in a consensus sequence 590 bp long. NCBI BLAST (blastn) search of the ITS sequence resulted in 100 high scoring hits with alignment score ≥ 200. The sequence-based identification confirmed the isolate as *Ophiostoma phasma* with 100% sequence identity and E value of zero. Sequence alignment resulted in a perfect alignment at all the nucleotide positions with no gaps and a 100% identity score (Figure 3.3).



Sequence ID: gb|DQ316219.1| Length: 456 Number of Matches: 1

Range 1: 1 to 456 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
843 bits(456)	0.0	456/456(100%)	0/456(0%)	Plus/Plus

```
Query  96   GCGAACCGTACCCAATTTGTTCTCGTTGCTTCCGGCGGCGGGGGGCCGAAAGGCTCTCCC  155
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    GCGAACCGTACCCAATTTGTTCTCGTTGCTTCCGGCGGCGGGGGGCCGAAAGGCTCTCCC  60

Query  156  TGCCGGGGGCGGCGGGCCCTATGAACCTTTATATCTCAACCACTAGAAACCGTCTGAGAA  215
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   TGCCGGGGGCGGCGGGCCCTATGAACCTTTATATCTCAACCACTAGAAACCGTCTGAGAA  120

Query  216  ACAAAACAAATAATAAAAACTTTCAACAACGGATCTCTTGGCTCTGGCATCGATGAAGAA  275
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  ACAAAACAAATAATAAAAACTTTCAACAACGGATCTCTTGGCTCTGGCATCGATGAAGAA  180

Query  276  CGCAGCGAAATGCGATACGTAATGTGAATTGCAGAATTCAGCGAACCATCGAATCTTTGA  335
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  CGCAGCGAAATGCGATACGTAATGTGAATTGCAGAATTCAGCGAACCATCGAATCTTTGA  240

Query  336  ACGCACATTGCGCCCGCCAGTATTCTGGCGGGCATGCCTGTCCGAGCGTCATTTCCCCCC  395
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  ACGCACATTGCGCCCGCCAGTATTCTGGCGGGCATGCCTGTCCGAGCGTCATTTCCCCCC  300

Query  396  TCACGCGCCCCGTTGCGCGCTGGTGTTGGGGCTCCTCCGCCTGGCGGAGGGCCCCGAAAG  455
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  TCACGCGCCCCGTTGCGCGCTGGTGTTGGGGCTCCTCCGCCTGGCGGAGGGCCCCGAAAG  360

Query  456  CGAGTGGCGGGCCCTGTGGAAGGCTCCGAGCGCAGTACCGAACGCAAGTTCTCCCCTCGC  515
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  CGAGTGGCGGGCCCTGTGGAAGGCTCCGAGCGCAGTACCGAACGCAAGTTCTCCCCTCGC  420

Query  516  TCAGACTGCCCCCCAGGCGCCCTGCCGTCAAAACGC  551
            ||||||||||||||||||||||||||||||||||||
Sbjct  421  TCAGACTGCCCCCCAGGCGCCCTGCCGTCAAAACGC  456
```

Figure 3.3 Sequence alignment produced by NCBI BLAST (bl2seq) of the consensus CMW 20676 ITS region and the *O. phasma* ITS region.

3.3 LC-MS/MS peptide identification

O. phasma 20676 produced two prominent bands on SDS-PAGE when grown in MYE medium (Figure 3.1 B). Both bands were excised from the gel and sent for LC-MS/MS peptide identification. Raw mass spectrometry data generated from LC-MS/MS was analysed using PEAKS software [software package which performs peptide and protein identification from raw spectrometry data (www.bioinfor.com/peaks/features/overview.html) for the identification of peptides and subsequently putative proteins. Peptide amino acid sequences from both the top and bottom bands showed significant similarity with the Man5 protein from a *Penicillium* species (Figure 3.4).

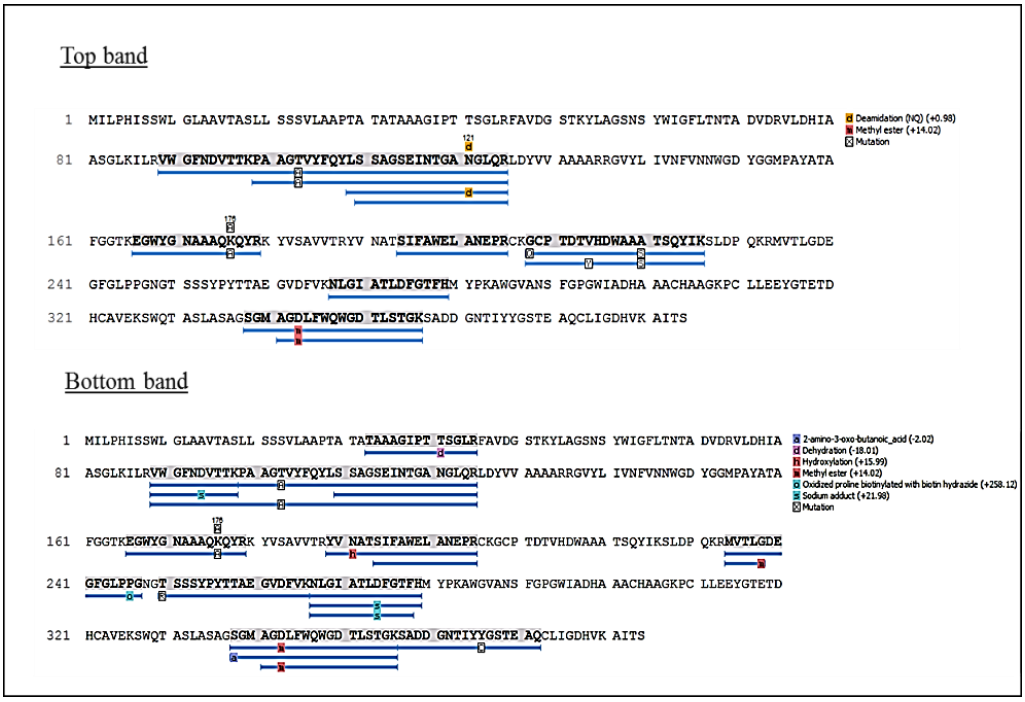


Figure 3.4 *Penicillium* Man5 protein coverage by identified peptides (highlighted and underlined in blue) generated from the top and bottom bands through LC-MS/MS.

BLAST search of the NCBI protein database showed similarity between the Man5 protein from the *Penicillium* species and the glycoside hydrolase family 5 protein from *O. piceae*. The sequences aligned with 68% sequence identity and 272/336 amino acids aligned perfectly (Figure 3.5).

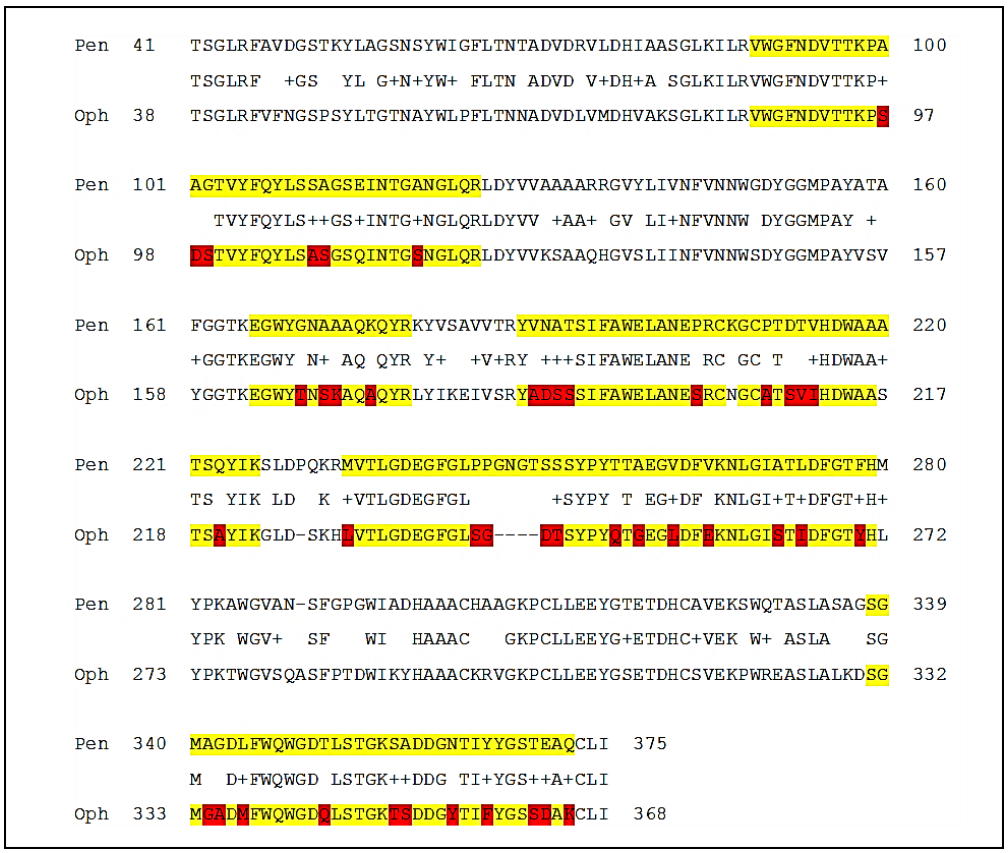


Figure 3.5 Amino acid sequence alignment of Man5 protein from a *Penicillium* species (Pen) and the amino acid sequence of the glycoside hydrolase family 5 protein from *O. piceae* UAMH 11346 (Oph). Peptides highlighted in yellow were generated from LC-MS/MS and are 100% similar to the Man5 peptides. Differences in amino acids of the *Ophiostoma* peptides are highlighted in red.

The 40 kDa protein produced by *O. phasma* 20676 was identified as a putative mannanase. The nucleotide sequences encoding both proteins were then obtained by reverse translation using the EXPASY reverse translation tool (http://www.bioinformatics.org/sms2/rev_trans.html) and the two nucleotide sequences were aligned (Fig 3.6). The nucleotide sequences of two of the identified peptides were then used to design the forward (MANF) and reverse (MANRev) mannanase (*man*) gene specific primers (Table 2.7).

Figure 3.6 Nucleotide sequence alignment of the Man5 protein from the *Penicillium* sp. enrichment culture clone C6 (Pen) and the glycoside hydrolase family 5 protein from *O. piceae* UAMH 11346 (Oph). Nucleotides highlighted in yellow were identified from reverse translated peptides generated through LC-MS/MS present in both the *Penicillium* and *Ophiostoma* proteins. The underlined nucleotides were used to design the *man* gene specific primers and bases highlighted in red are different between the two sequences.

3.4 Effect of different growth parameters on protein production

3.4.1 Effect of inoculum on protein production

There was no difference observed in *O. phasma* 20676 growth morphology when mycelia or spores were used as inoculum (MYE, 28ºC, 160 rpm), however, protein production levels were observed to be higher when spores were used as inoculum instead of mycelia. Therefore, spores were used to generate a standardised inoculum that facilitated the measurement of the effect of different growth parameters on protein production by *O. phasma* 20676. Different spore concentrations were used to inoculate MYE and five days after inoculation, extracellular protein produced was compared by 12% SDS-PAGE. The optimum spore inoculum concentration was determined to be 1 x 10^7spores/mL (Figure 3.7).

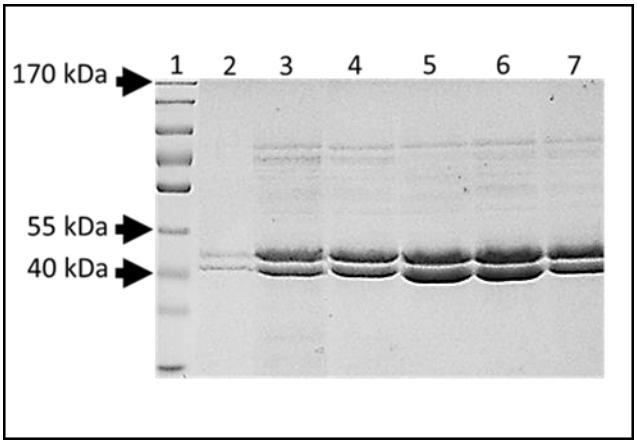


Figure 3.7 SDS-PAGE comparing protein production by *O. phasma* 20676 (MYE, 28°C, 160 rpm, five days after inoculation) using different spore concentrations as inoculum. Lane 1, pre-stained molecular weight marker; Lane 2, 1 x 10^3 Spores/mL; Lane 3, 1 x 10^4 spores/mL; Lane 4, 1 x 10^5 spores/mL; Lane 5, 1 x 10^6 spores/mL; Lane 6, 1 x 10^7 spores/mL and Lane 7, 1 x 10^8 spores/mL. Each lane contains protein from 500 µL of culture supernatant.

3.4.2 Effect of different growth media on protein production

To evaluate the effect of five growth media on protein production by *O. phasma* 20676, the isolate was cultured in ME, PDB, MYE, CM and NB (inoculum 1 x 10^7spores/mL, 28°C, 160 rpm). Five days after inoculation, extracellular protein secreted into the culture supernatant was compared between the different media. MYE proved to be the optimal medium from those tested for high level protein production by *O. phasma* 20676 (Figure 3.8).

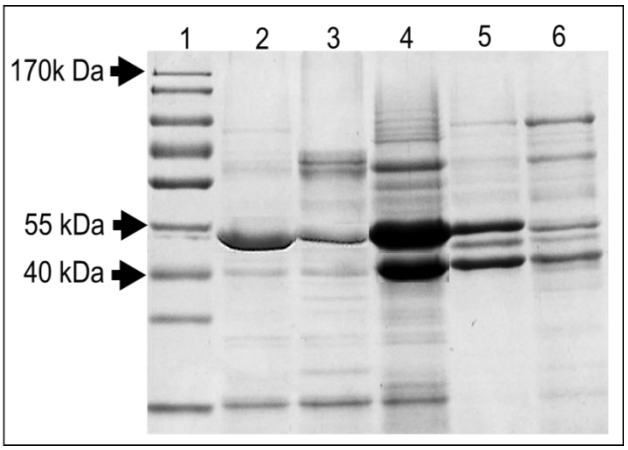


Figure 3.8 10% SDS-PAGE analysis of the effect of different culture media on production of extracellular protein by *O. phasma* 20676. Lane 1, pre-stained molecular weight marker. The different media were: Lane 2, ME; Lane 3, PDB; Lane 4, MYE; Lane 5, CM and Lane 6, NB. Each lane contains protein from 1 µL of culture supernatant.

3.4.3 Effect of different carbon sources on protein production by *O. phasma* 20676

The effect of three different carbon sources (1% glucose, 1% LBG and 1% AGG) added to MYE was evaluated for the production of extracellular protein by *O. phasma* 20676. Analysis of the SDS-PAGE protein profile of the culture supernatants identified LBG as the optimal carbon source for improved protein production (Figure 3.9 A). To determine the optimal concentration of LBG, *O.*

phasma 20676 was cultured in MYE medium containing different concentrations of LBG (0.2%, 0.4%, 0.6%, 0.8% and 1%) (Figure 3.9 B). Highest levels of protein produced by *O. phasma* 20676 were with 0.6% to 1% LBG. Therefore 1% LBG was chosen as the optimal substrate concentration.

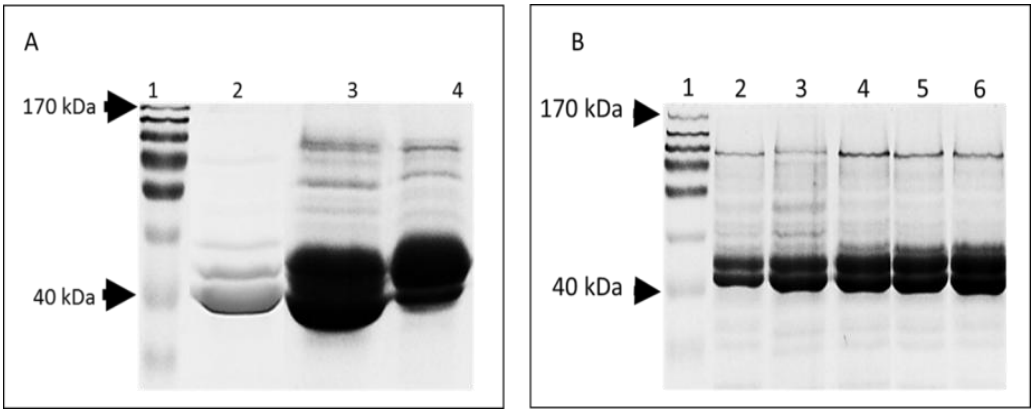


Figure 3.9 12% SDS-PAGE analysis of (A) the effect of carbon source and (B) LBG concentration on protein production by *O. phasma* 20676. (A) Lane 1, pre-stained molecular weight marker; Lane 2, MYE with 1% glucose; Lane 3, MYE with 1% LBG and Lane 4 , MYE with 1% AGG. (B) Lane 1, pre-stained molecular weight marker; Lane 2, 0.2% LBG; Lane 3, 0.4% LBG; Lane 4, 0.6% LBG; Lane 5, 0.8% LBG and Lane 6,1% LBG. Each lane in figure A contains protein from 1 mL of culture supernatant and each lane in figure B contains protein from 500 µL of culture supernatant.

3.4.4 Effect of pH on protein production by *O. phasma* 20676

Six culture pH conditions (MYE, 28°C, 160 rpm; pH 5, pH 6.4, pH 7, pH 8, pH 9 and pH 10) were evaluated for protein production by *O. phasma* 20676. Culture pH was monitored throughout the five day growth period (Table 3.1).

Table 3.1 The pH of *O. phasma* 20676 cultures at different days after inoculation.

pH of sterile medium	pH Day 1 Cultures	pH Day 2 Cultures	pH Day 3 Cultures	pH Day 4 Cultures	pH Day 5 Cultures
pH 5	5	5	6	8	8
pH 6. 4	6	5	7	8	8

pH 7	7	6	7	8	8
pH 8	7.5	6	7	8	8
pH 9	8.5	7.5	7.5	8	8
pH 10	9	8.5	8.5	9	7

Five days after inoculation the initial pH of the different cultures had either increased or decreased moving towards pH 8. A comparison of the amount of extracellular protein in the culture supernatant (Figure 3.10) indicated high levels of protein produced by *O. phasma* 20676 over a wide pH range (Figure 3.10). The optimal pH for high level protein production by *O. phasma* 20676 was therefore determined to be below pH 8.

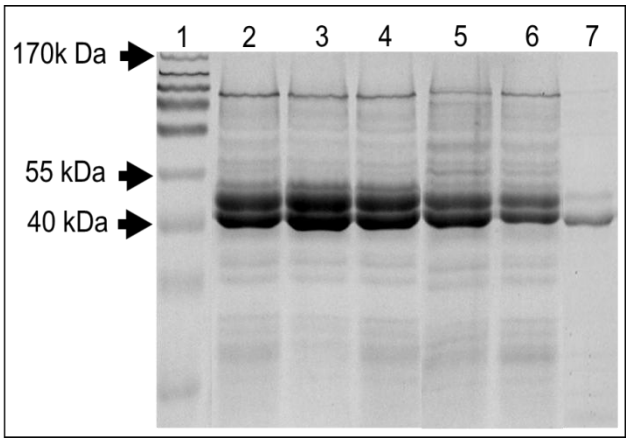


Figure 3.10 Comparison of the effect of five pH conditions on the production of extracellular protein by *O. phasma* (MYE, 28°C, 160 rpm and five days after inoculation). Lane 1 represents the pre stained molecular weight marker. The starting pH of the MYE medium was: Lanes 2, pH 5; Lane 3, pH 6.4; Lane 4, pH 7; Lane 5, pH 8, Lane 6, pH 9 and Lane 7, pH 10. Each lane contains protein from 500 µL of culture supernatant.

3.4.5 Effect of temperature on protein production

Temperature effect on protein production by *O. phasma* 20676 was determined by culturing *O. phasma* 20676 at different temperatures (25°C, 28°C, 30°C, 32ºC and 34°C in MYE, 160 rpm). The results from SDS-PAGE analysis of the extracellular protein five days after inoculation indicate the optimum

temperature for protein production by *O. phasma* 20676 to be between 25ºC and 28ºC (Figure 3.11).

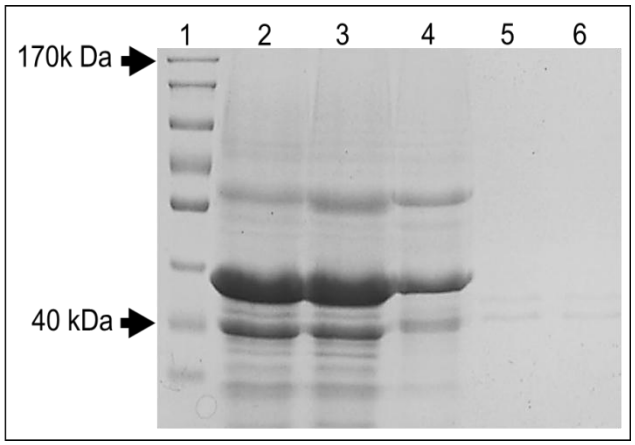


Figure 3.11 10% SDS-PAGE analysis of the effect of temperature on the production of extracellular protein by *O. phasma* 20676 (MYE, 160 rpm, five days after inoculation). Lane 1, pre-stained molecular weight marker; Lane 2, culture at 25°C; Lane 3, culture at 28ºC; Lane 4, culture at 30°C; Lane 5, culture at 32°C and Lane 6, culture at 34°C. Each lane contains protein from 1 mL of culture supernatant.

3.4.6 Optimal conditions

Extracellular protein produced by *O. phasma* 20676 grown under different conditions was compared on SDS-PAGE gel (Figure 3.12) and showed that the optimized growth conditions are; inoculum 1 x 10^7 spores/ mL in MYE containing 1% LBG ,pH between5-8 and temperature between 25°C-28ºC with shaking at 160 rpm.

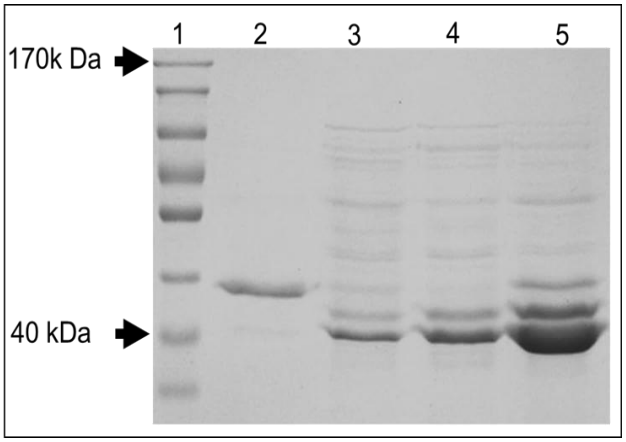


Figure 3.12 10% SDS-PAGE showing improved protein production by *O. phasma* 20676 through optimization of growth conditions. Lane 1, Pre stained molecular weight marker; Lane 2, ME (mycelium inoculum; 28°C); Lane 3, MYE (mycelium inoculum; pH 6.4; 28°C); Lane 4, MYE (1 x 10^7 spores/mL inoculum; pH 6.4; 28°C) and Lane 5, MYE (1 x 10^7 spores/mL inoculum; pH 6.4; 28°C; 1% LBG). Each Lane contains protein from 250 µL of culture supernatant.

3.5 PCR amplification, cloning and sequencing of internal (*man)* gene fragment

MANF and MANRev primers (Table 2.7) were designed based on the *Penicillium* Man5 protein nucleotide sequence (Figure 3.6). The primers were then used for amplification of an internal *man* gene fragment, using the isolated *O. phasma* 20676 genomic DNA as template. PCR amplification (Chapter 2.7) generated an amplicon of approximate 330 bp (Figure 3.13).

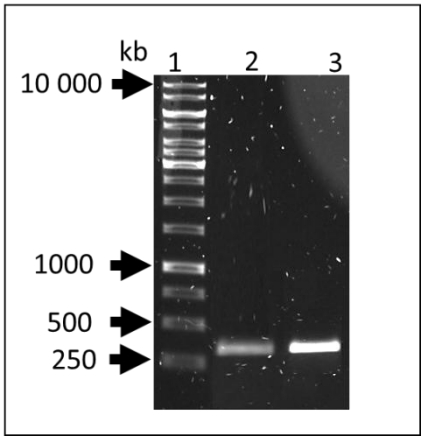


Figure 3.13 PCR amplification of internal *man* gene fragment. Lane 1, 1kb DNA ladder, Lane 2, internal *man* gene amplicon, 1 µL DNA template and Lane 3, internal *man* gene amplicon, 2 µL DNA template.

Purified DNA fragments were cloned into the pGEM®-T Easy vector system (Chapter 2.10) and transformants were screened by colony PCR (Chapter 2.10.1) using primers MANF and MANRev. From the colonies screened three were identified containing the 330 bp fragment (Figure 3.14).

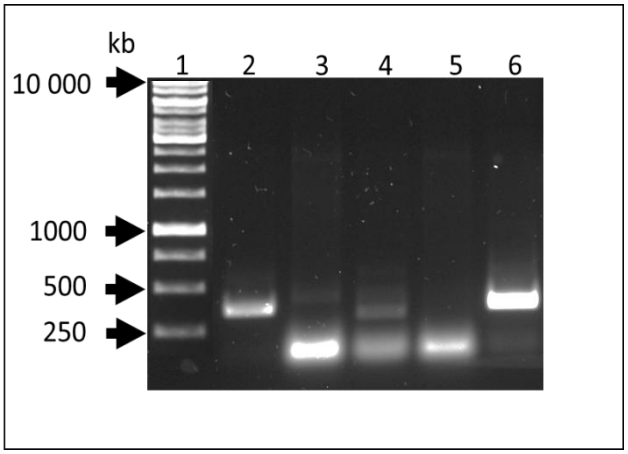


Figure 3.14 PCR Screening of colonies for presence of the 330 pb *man* gene fragment cloned into the pGEM®-T Easy vector. Lane 1, 1kb DNA ladder; Lane 2, 330 pb *man* gene fragment (positive control); Lane 3–6, clones screened for the presence of the 330 pb *man* gene fragment.

To confirm presence of the *man* gene fragment plasmid DNA was isolated (Chapter 2.8.5) from two positive clones (Figure 3.14; lanes 4 and 6). The *man* gene fragment ligated into the pGEM®T-Easy vector was flanked on the 3' and 5' end by an *Eco*RI restriction site. Therefore, plasmid DNA was digested with *Eco*RI restriction enzyme to release the 330 pb *man* gene fragment from the vector (Figure 3.15).

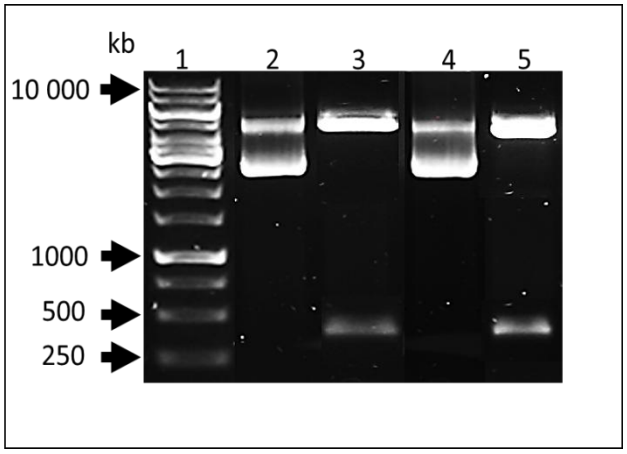
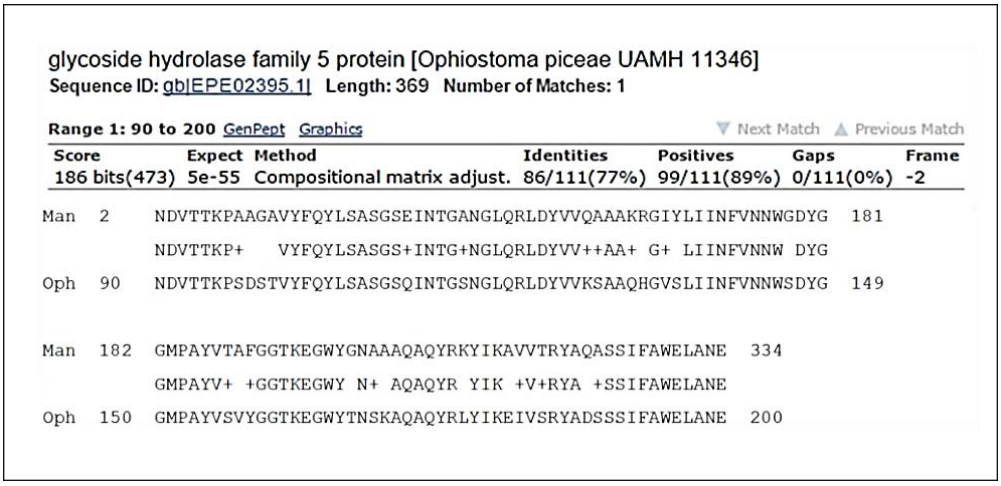


Figure 3.15 *Eco*RI restriction analyses to confirm presence of insert cloned into pGEMT®-T Easy vector. Lane 1, 1kb DNA Ladder; Lane 2, Clone 4 Uncut; Lane 3, Clone 4 *Eco*RI digest*;* Lane 4, Clone 6 Uncut and Lane 5, Clone 6 *Eco*RI digest.

Plasmid DNA from clone 4 and 6 was sequenced using the M13 forward and reverse primers and resulted in a 334 bp long nucleotide sequence. BLAST search of the translated nucleotide sequence againts the NCBI protein (blastx) database confirmed homology with the glycoside hydrolase family 5 protein [*O. piceae* UAMH 11346]. Amino acid sequence alignment resulted in a 77% identity and 99/111 amino acids aligned perfectly (Figure 3.16).


glycoside hydrolase family 5 protein [Ophiostoma piceae UAMH 11346]
Sequence ID: gb|EPE02395.1| Length: 369 Number of Matches: 1

Range 1: 90 to 200 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
186 bits(473)	5e-55	Compositional matrix adjust.	86/111(77%)	99/111(89%)	0/111(0%)	-2

```
Man  2    NDVTTKPAAGAVYFQYLSASGSEINTGANGLQRLDYVVQAAAKRGIYLIINFVNNWGDYG  181
          NDVTTKP+   VYFQYLSASGS+INTG+NGLQRLDYVV++AA+ G+ LIINFVNNW DYG
Oph  90   NDVTTKPSDSTVYFQYLSASGSQINTGSNGLQRLDYVVKSAAQHGVSLIINFVNNWSDYG  149

Man  182  GMPAYVTAFGGTKEGWYGNAAAQAQYRKYIKAVVTRYAQASSIFAWELANE  334
          GMPAYV+ +GGTKEGWY N+ AQAQYR YIK +V+RYA +SSIFAWELANE
Oph  150  GMPAYVSVYGGTKEGWYTNSKAQAQYRLYIKEIVSRYADSSSIFAWELANE  200
```

Figure 3.16 Amino acid sequence alignment generated from the BLAST search (blastx) of the *man* gene fragment translated nucleotide sequence against the NCBI protein database.

3.6 Chromosome walking by SiteFinding PCR

SiteFinding PCR (Tan et al., 2005) was used as a chromosome walking technique to obtain the complete *O. phasma* 20676 *man* gene sequence as well as the flanking regulatory regions. SiteFinding PCR involved two reaction steps; SiteFinding reaction and Nested PCR (primary and secondary) reactions. In the SiteFinding reaction, DNA was denatured and the SiteFinder oligonucleotide annealed to the single stranded DNA. One *Taq* DNA polymerase was used to generate double stranded target molecules of different lengths. The product of the SiteFinding PCR reaction was then used as template in the primary step of the Nested PCR reaction. In the primary reaction, the SFP 1 and GSP 1 (MANRev) gene specific primers were used to exponentially amplify the target DNA (Figure 3.17 A).

For the secondary PCR amplification reactions, the PCR product from the primary reaction was diluted (1/10, 1/100 and 1/500) to determine which dilution would be optimal for the secondary PCR reaction. The diluted primary PCR

product, SiteFinder 2 and the gene specific primers (GPS 2 or GSP 3) (Table 2.10) were used in the secondary reaction to exponentially amplify target DNA molecules (Figure 3.17 B). The primary PCR product was determined to amplify best when undiluted (Figure 3.17 B).

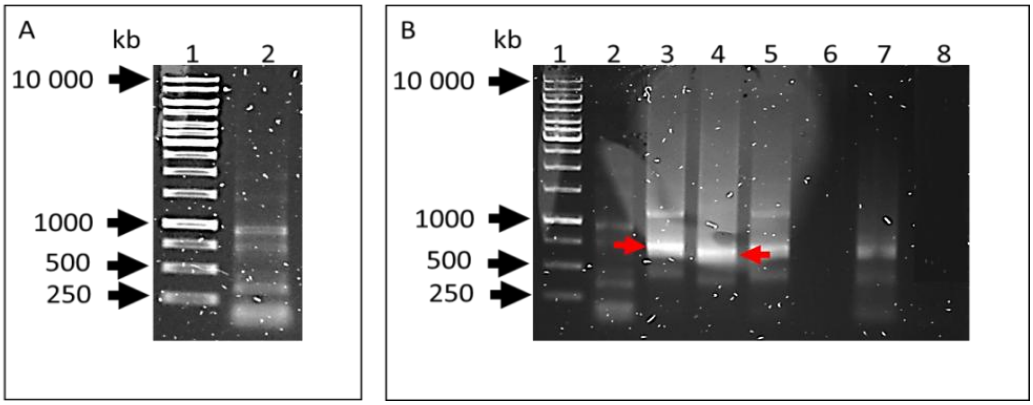


Figure 3.17 Analysis of SiteFinding PCR products from (A) the primary Nested PCR reaction and (B) the secondary Nested PCR reaction. (A) Lane 1, 1kb DNA Ladder and Lane 2, primary PCR product. (B) Lane 1, 1kb DNA Ladder; Lane 2, primary PCR product (positive control); Lane 3, GSP 2 undiluted template; Lane 4, GSP 3 undiluted template; Lane 5, GSP 2 1/100 diluted template; Lane 6, GSP 3 1/100 diluted template; Lane 7, GSP 2 1/1000 diluted template and Lane 8, GSP 3 1/1000 diluted template.

Bands approximately 650 bp in size (Figure 3.18 B) were cleaned from the gel (Figure 3.18 A) and cloned into the pGEM®-T easy vector. Transformed clones were screened for the presence of the 650 pb fragment (Figure 3.18 B) using SiteFinder 2 and GSP 2/ GSP 3 gene specific primers.

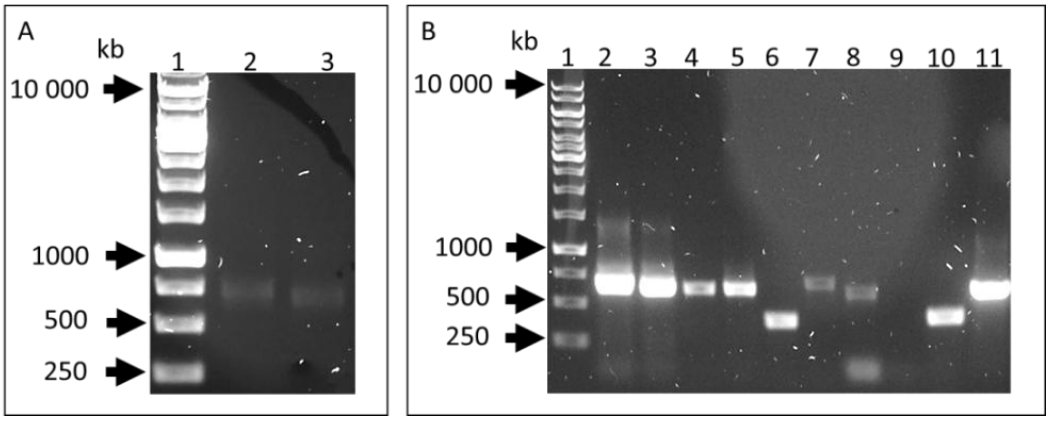


Figure 3.18 Agarose gel analysis of (A) amplicons cleaned from agarose gel (Figure 3.17 B) and (B) Colony PCR for confirming the presence of the 650 bp *man* gene fragment in transformed clones. (A) Lane 1, 1kb DNA Ladder; Lane 2, GSP 2 cleaned amplicon and Lane 3, GSP 3 cleaned amplicon. (B) Lane 1, 1kb DNA Ladder and Lane 2-11, Clones screened for the presence of the 650 bp *man* gene fragments.

Plasmid DNA was isolated from Clones 1 and 2 (Figure 3.18 B) and digested with *Eco*RI restriction endonuclease. Restriction digestion released the cloned 650 bp DNA fragment amplified from the 5' region of the *man* gene from the vector. The *man* gene fragment from Clone 1 was sequenced and sequencing generated a 700 bp long nucleotide sequence upstream of the *man* gene fragment (Figure 3.19).

Further chromosome walking experiments were carried out to obtain the complete *man* gene sequence as well as the rest of the up- and down-stream regulatory sequences. The complete *man* gene nucleotide sequence was 1135 bp long and the putative mannanase had a theoretical molecular weight of 40.16 kDa. Transcriptional elements flanking the *man* gene were identified by searching for conserved motifs identified from literature. To facilitate *in silico* analysis of the *man* gene regulatory sequences, the TRANSFAC database and programs (Match-1.0 Public and Patch 1.0 http://www.gene-regulation.com/pub/programs.html) were used. Match-1.0 is a program that uses positional weight matrices to predict the presence of transcription factor

binding sites (TFBS) in DNA sequences. A positional weight matrix (PWM) is a statistical representation of the binding pattern of a transcription factor estimated from known binding site sequences (Georgi & Schliep, 2006). Match-1.0 makes use of the positional weight matrices library from TRANSFAC® Public 6.0 whereas, the Patch 1.0 program uses patterns and binding sites from TRANSFAC® Public 6.0 to predict TFBS in DNA sequences. The possible transcriptional elements that were identified as well as their potential factor binding sites are shown in Figure 3.19. The sequences and functions of the transcriptional elements are listed in Table 3.2.

Two TATA boxes were predicted in positions 1220 and 2088 using the Patch 1.0 program, and a putative transcription initiation site (TAC) was identified at position 2121exactly 27 nucleotides downstream of the identified TATAAA box. This corresponds to the 25-30 nucleotide distance reported between the TSS and the TATA transcription elements in higher eukaryotes (Zhang & Dietrich, 2005). One CAAT box was also identified at position 1308, similar to promoter regions that have been studied in other *Ophiostoma* species. The CAAT box was located 240 bp adjacent to the transcription start site (TSS) (Pereira et al., 2000; Wu et al., 2006). In literature, the CAAT box has been demonstrated to play a role in increasing promoter strength (Bezhaniet al., 2001). Three heat shock transcription factors (HSF) were identified upstream of the TSS. HSF_02 and HSF_03 have been discovered in *S. cerevisiae* and *Kluyveromyces lactis* to be heat shock gene activators. They are therefore essential for growth as they mediate the expression of heat-shock genes in response to elevated temperatures or to other stress conditions such as glucose starvation (Brown et al., 2010) Other putative transcriptional factors were also identified. Most transcription factors that were found had homologues in fungal species and are described in Table 3.2 and shown in Figure 3.19.

Figure 3.19 Putative transcriptional factors identified in the 5' region of *man* gene sequence. The start of the predicted ORF is underlined in black and highlighted in turquoise. The two identified TATA boxes which are known to bind transcription factors such as the DNA polymerase II and the TFIID are in red bold text. The predicted open reading frame (ORF) is highlighted in yellow. Descriptions of the other identified transcription factors are given in Table 3.2

Table 3.2 A descriptions of the transcriptional elements identified in the 5' region of the *man* gene and their position on the DNA sequence.

	Motif ID	Sequence	Description	No.	Position

1.	CAP	TAC	Cap signal for transcription initiation.	1	2121
2.	TBP	TATAAA AATAAA	TATA box. A conserved motif found in a number of eukaryotes including *S.cerevisiae* and *S. pombe*. Activator of all three polymerases (Pol I, II and III) in *S. cerevisiae.*	2	2090, 1222
3.	GCN4	TGACTA AAGTCA	Activator of genes involved in protein and purine biosynthesis identified in *S. cerevisiae.*	2	627, 908
4.	HSF_03 HSF_02 HSF_02	GGAAAgggagtttct CctcttttctCTTCT cttctcctctTTTCT	An activator identified in *S. cerevisiae* and *K. lactis*. It is essential for growth, mediating expression of heat-shock genes in response to elevated temperatures or to other stress conditions such as glucose starvation.	3	110, 1706, 1725
5.	CP1	CCAAT	CAAT Box is a complex that binds regulatory elements found in the promoters of genes encoding plant cell wall degrading enzymes and transcriptional activators or factors modulating expression.	1	1310
6.	MIG1	TTTTA TATTC TTTTT TTTAT AATTA	Cre transcription factor that mediates glucose repression.	13	19, 44, 62, 270, 446, 665, 724, 1043, 1214, 1310, 1537, 1656, 1695
7.	GATA	GATA	*Are*A transcription factor. Regulates nitrogen catabolic enzymes during the use of secondary nitrogen sources.	9	257, 435, 565, 604, 615, 746, 793, 1188 ,1754
8.	CHA4	TCCGC GCGGA	Transcriptional element in *S. cerevisiae* that cctivates the CHA1 gene for L-serine (L-threonine) deaminase responsible for the utilisation of serine/threonine as nitrogen sources.	8	1100, 1383, 1807, 1822, 1972, 1990, 2008, 2026
9.	ADR1	TCTCC TTGGGG	Positive regulator of peroxisomal protein genes	9	249, 718, 757, 924,

			identified in *S. cerevisiae*. It is required in yeast for ADH2 activation and glycerol metabolism.		1019, 1420, 1679, 1688, 1718
10.	NIT2	TATCAT TATCTC	Activator of nitrogen-regulated genes; NIT2 can partially complement for lack of AREA in *A. nidulans*. NIT2 is also major nitrogen regulatory gene in *N. crassa*. It encodes a protein with a putative zinc finger DNA-binding domain.	2	67, 518
11.	Homeo box	AATTATT	Homeobox motif found within a number of genes. Involved in morphogenesis of fungi, animals and plants.	1	476
13.	ATG TAG	Start codon Stop codon	Start and stop codons indicating the beginning and the end, respectively, of the ORF.	1 1	2197 3331
14.	Poly A tail	AAAAAA	Protects the 3' end of the mRNA from exonuclease degradation.	1	3424

In Table 3.2 the column headed Motif ID indicates the conserved motif identified within the sequence, the column headed No. refers to the number of conserved motifs identified within the 5' promoter region and Position indicates the position of the conserved motif from the start of the nucleotide sequence.

CHAPTER FOUR

DISCUSSION

The expression of heterologous genes in filamentous fungal hosts is an alternative to natural protein production allowing high level production of recombinant proteins. In this study, the CMW fungal culture collection was screened for a fungal isolate with the natural ability to produce high levels of extracellular protein. The reason for screening was to exploit the isolate's promoter and regulatory regions flanking the gene encoding the highly expressed protein for application in the development of a host-vector expression system. In comparison to the other screened fungal isolates, *O. phasma* 20676 produced the highest amount of a single extracellular protein in both ME and MYE media. *O. phasma* 20676 belongs to the Ophiostomataceae family within the Ascomycota Phylum. The Ophiostomataceae are saprophytic polymorphic fungi that are found worldwide (Zipfel et al., 2006). *Ophiostoma* species utilize starch and soluble sugars as main carbon sources for growth (Seifert, 1993). Some fungi of the *Ophiostoma* species naturally secrete large amounts of specialized enzymes that aid in hydrolysis of cellulose and lignin into simple sugars. Examples of cellulolytic enzymes secreted by some *Ophiostoma* species are mannanases, amylases and laccases (Przybyl et al., 2006; Wu et al., 2006). These enzymes have been proven to be important for the establishment, spreading and colonization of *Ophiostoma* species in wood (Abraham & Breuil, 1996; Brush, 1999).

Two prominent bands of approximately 40 kDa (lower band) and 50 kDa (upper band) were detected on SDS-PAGE during screening for high level protein production. The double bands suggested that the protein could be an isoenzyme with varying molecular weight. However, this was impossible to

conclude from SDS-PAGE analysis and therefore both bands were excised from the gel and sent for LC-MS/MS peptide identification. Analysis of the amino acid peptide sequences generated from LC-MS/MS for both bands confirmed the bands to be the same protein. The protein produced by *O. phasma* 20676 corresponding to a size between 40-50 kDa was subsequently identified as a putative mannanase through LC-MS/MS and sequence based identification. Puchart et al. (2004) purified two extracellular endo-1, 4-beta mannanases (MAN I, 60 kDa and MAN II, 63 kDa) from a locust bean gum spent culture of *A. fumigates*. Both mannanases contained identical internal amino acid sequences that corresponded to a glycoside hydrolase family 5 protein. This data suggested that both mannanases were products of the same gene varying only in post translational modification.

Mannans make up the major constituents of the hemicellulose fraction in softwoods and a wide range of plant tissues (Rodriguez-Gacio Mdel et al., 2012; Silva et al., 2011). The major mannan-degrading enzymes have been identified to be β-mannanases, β-mannosidases and β-glucosidases (Chauhan et al., 2012). Endo-β-mannanases are involved in the hydrolysis and/or modification of mannan-based polysaccharides as catalysts for the hydrolysis of the β-1,4-linked backbone within different mannans (Wang et al., 2013). In the CAZy (carbohydrate-active enzymes) database, enzymes with this activity are classified into three different glycoside hydrolase families: GH5, GH26 and GH113 (Cantarel et al., 2009). Mannanases are ubiquitous in nature and are known to be produced by a variety of bacteria, fungi, actinomycetes, plants and animals (Chauhan et al., 2012). In a study to develop albino *Ophiostoma* strains for biopulping treatments for the paper and pulp industry (Farrell et al., 2004), a mannanase secreted by *O. picea* was isolated and characterized. Raffaello

et al. (2013) also identified a mannanase from the pine heartwood colonizing fungus *H. annosum* which had high level of similarity with mannanases from the wood degrading basidiomycete *Armillaria tabescens*, the dry rot fungus *Serpula lacrymans* var. *lacrymans* S7.9 and the white-rot fungus *Phanerochaete chrysosporium*. Mannanases are therefore thought to be common proteins secreted by wood colonizing filamentous fungi.

Characteristics targeted when screening filamentous fungi as possible production hosts for extracellular proteins include rapid growth rate, high level protein production capability and environmental growth parameters such as pH and temperature (Visser et al., 2011). The successful high level production of fungal proteins requires detailed knowledge of the growth characteristics and the physiology of the fungus of interest. For this reason, growth parameters for increased mannanase production were evaluated and results revealed that adjusting growth conditions of *O. phasma* 20676 significantly increased the level of mannanase production.

In submerged culture conditions, two morphological types can be identified; discrete pellets and freely dispersed mycelia. These two morphological forms determine the type of fungal product produced and depending on the desired product, the optimal morphology for a given bioprocess varies and cannot be generalized (Gibbs et al., 2000). For example, free mycelia are required for the production of penicillin from *Penicillium chrysogenum*, whereas pellets are required for the production of citric acid from *A. niger* (Wang et al., 2005). However, a study by Domingues et al. (2000) also showed that the size of inoculum (spore concentration) played an important role in fungal morphology, and showed that with *T. reesei,* high inoculum size (>10^7 spores/mL) led to the

presence of very small pellets with mostly freely dispersed mycelia, whereas a small inoculum size (< 10^5 spores/mL) produced mainly pellets. According to Qinnghe et al. (2004), the optimal spore concentration for most fungi is between 10^6 and 10^7 spores/mL; outside of this range, a decrease in enzyme productivity and activity is observed. In cases of very high or too low inoculum concentrations, the fungal metabolism and some enzyme activities are reduced. The disadvantage of having very high spore concentrations is that it can cause a significant decrease in the specific rate of oxygen consumption (Brown & Zainudeen, 1987). On the other hand, an insufficient spore concentration leads to poor mycelium production, and therefore low efficiency in the assimilation of carbon and nitrogen sources (Meyrath & Macintosh, 1963).

In this study, the inoculum size was standardized to investigate the effects of other growth parameters on mannanase production by *O. phasma* 20676. Protein production levels were similar between inoculum concentrations of 1 x 10^7spores/mL and 1 x 10^8 spores/mL. However the optimum spore inoculum concentration for high level mannanase production by *O. phasma* 20676 was determined to be 1 x 10^7 spores/mL.

Five different growth media tested in this study had an effect on mannanase production by *O. phasma* 20676. However, the optimal medium for high level mannanase production was MYE. According to Domingues et al., (2000), culture medium composition influences filamentous fungal morphology and protein production. The peptone in the MYE medium contains essential amino acids whereas the yeast extract is a good source of naturally occurring vitamin B complex. These two constituents have previously been reported to stimulate

both cell growth and enzyme production (Domingues et al., 2000). The synthesis of enzymes by filamentous fungi is subject to induction and/or catabolic repression (Sachslehner et al., 1998) and in the case of mannanase production by *O. phasma* 20676, the addition of a solid carbon substrate (locust bean gum) increased protein production. The addition of solid carbon substrates to media has not only been reported to optimize protein production but also enhances hyphal growth (Papagianni, 2004).

Steady and slow release of carbon from solid carbon substrates extends the process of product formation and reduces the effects of buildup of toxic medium constituents such as carbon dioxide and limiting metabolites, acids and alcohol (Blibech et al., 2011; Singh et al., 2004). Mannan substrates are divided into four groups (mannan, glucomannan, galactomannan and galactoglucomannan) depending on the backbone monomer composition and the presence of side chains (Wang et al., 2013). Galactomannan is often used as carbon source instead of galactoglucomannan, due to its commercial unavailability. Locust bean gum (LBG), and acacia guar gum (AGG) are both used as solid sources of carbon. They have distinct galactose to mannose ratios of 1:4 (AGG) and 1:2 (LBG), respectively (Klyosov et al., 2012; Moreira & Filho, 2008). According to a study by Malgas et al. (2015), mannanases displayed a higher affinity for the low galactose substituted locust bean gum rather than for the highly galactose decorated guar gum. This supports the findings in this study, where higher mannanase levels were obtained with LBG as carbon substrate, in comparison to AGG as carbon substrate. The lowest mannanase levels were obtained in media with glucose as the carbon source.

Neutral pH ranges are prevalent during mannanase production by bacteria (Mabrouk & Ahwany, 2008), whereas, mannanases from fungi are mostly produced in the acidic pH range (Blibech et al., 2010; Kote et al., 2009). The pH for mannanase production by *O. phasma* 20676 in this study was determined over a wide pH range; with the optimal pH for increased mannanase production between 5 and 8. pH drifts were observed in the different cultures from day 1 to day 5 after inoculation. Depending on the starting pH of the medium, the pH of cultures increased or decreased throughout the 5 day growth period, moving towards pH 8. According to Papagianni (2004), the initial pH of the culture medium and the extent of pH variations during fungal growth is greatly affected by the composition of the medium. The pH of media that is not sufficiently buffered tends to rise or drop depending on the constituent salts. Therefore, decreasing pH variations during growth is necessary as pH influences the concentration of dissolved bicarbonate which in return affects fungal growth and product formation. The observed pH fluctuations may have also been due to acid production from glucose utilization during the growth phase (Singh et al., 2004).

In general, mannanases are produced at temperatures ranging from 4 to 85°C (Kansoh & Nagieb, 2004; Politz et al., 2000) depending on the source organism. In a study to determine the role of various factors influencing host specificity of *Gondwanamyces* and *Ophiostoma* species found in the flower heads of *Protea* species in South Africa (Roets et al., 2012), temperature was monitored. Results showed that growth for both species differed significantly at different temperatures, although the maximum growth was observed at 25°C for both species. The fungal isolates screened in this study were obtained from the CMW fungal culture collection (Section 2.2) and isolates were cultured at

28°C on 2% (w/v) malt extract agar (MEA) (Zipfel et al., 2006). The two temperatures (25°C and 28°C) were therefore considered in the experiment to determine the effect of temperature on protein production by *O. phasma* 20676. Prominent mannanase bands were observed on SDS-PAGE when *O. phasma* 20676 was cultured at 25°C and 28°C. The intensely stained bands were indicative of high mannanase production levels.

One of the requirements for development of an efficient fungal expression system is a vector constituting strong transcription, translation and secretory signals flanking the gene encoding the highly expressed protein. The mannanase regulatory regions of the *O. phasma 20676 man* gene were identified for future use in a vector for efficient heterologous protein production. Putative transcriptional elements that may be involved in *man* gene expression were identified up- and down-stream of the gene. The promoter region up-stream of the 5' un-translated region (UTR) constituted of a CAP signal for transcription initiation, TATA-like boxes (TATAAA and AATAAA) and a CCAAT box that serves as a binding site of transcriptional elements involved in controlling promoter activity (Wu et al., 2006). These transcriptional elements have been previously identified in other *Ophiostoma* species (Robson, 2008). The mannanase coding region (ORF) extended from the ATG start codon and continued until terminated by the TAG translation stop codon (379 amino acids). The 3' un-translated region was followed by a polyadenylation signal (Poly A tail) which prevents degradation of the 3' end of mRNA by exoribonucleases (Lynch, 2006).

The identification of the highly expressed mannanase protein from *O. phasma* 20676 and the isolation and sequencing of the *man* gene with its regulatory

regions lays the foundation for developing a system for high level heterologous protein expression. The mannanase regulatory regions can be applied to a vector system for use in *O. phasma* 20676 as the expression host. Furthermore, an efficient transformation method has to be established for transformation of the recombinant vector into *O. phasma* 20676.

Appendix A

Table A.1 A list of all the isolates obtained from the CMW fungal culture collection and screened for high level protein production.

	Genus	Species	CMW	Host	Collected by
1	*Alternaria*	*alternata*	6076	Banana	Viljoen A
2	*Alternaria*	*alternata*	6134	Banana	Viljoen A
3	*Anthostomella*	*leucospermi*	22228	*leucospermum oleifolium*	Lee S
4	*Anthostomella*	*conorum*	20403	*Protea nerifolia*	Lee S
5	*Anthostomella*	*conorum*	20398	*Protea magnifica*	Lee S
6	*Anthostomella*	*cynaroides*	20405	*Protea cynaroides*	Lee S
7	*Aplosporella*	*sp*	38166	*Celtis africana*	Jami F, Gryzenhout M
8	*Aplosporella*	*sp*	38167	*Celtis africana*	Jami F, Gryzenhout M
9	*Aplosporella*	*sp*	38168	*Searsia lancea*	Jami F, Gryzenhout M
10	*Aplosporella*	*sp*	38169	*Searsia lancea*	Jami F, Gryzenhout M
11	*Armillaria*	*sp*	7207	*Protea sp*	Coetzee MPA
12	*Armillaria*	*sp*	11169	*Pinus taeda*	Roux J
13	*Armillaria*	*fuscipes*	3950	Letchwe	Coetzee MPA
14	*Armillaria*	*heimii*	3167	*Pinus elliottii*	Guillaumin / Ivory
15	*Arthrinium*	*phaeospermum*	18010	*Ischyrolepis cf gaudichaudiana*	Lee S
16	*Arthrinium*	*phaeospermum*	18061	*Ischyrolepsis subverticellata*	Lee S
17	*Arthrographis*	*cuboidea*	17142	Decaying timber	Smith A
18	*Aspergillus*	*versicolor*	24541	*Ischyrolepsis subverticellata*	De Meyer E
19	*Asperisporium*	*sp*	18011	*Ischyrolepis cf gaudichaudiana*	Lee S
20	*Aureobasidium*	*pullulans*	24544	Pine wood pole	De Meyer E
21	*Basidomycete*	*sp*	24546	Eucalyptus wood poles	De Meyer E
22	*Basidomycete*	*sp*	24547	Eucalyptus wood poles	De Meyer E
23	*Basidomycete*		24548	Eucalyptus wood poles	De Meyer E
24	*Batcheloromyces*	*leucadendri*	35691	*Leucadendron pondoense*	Marincowitz S
25	*Batcheloromyces*	*leucadendri*	35565	*Leucadendron spissifolium (Proteaceae)*	Marincowitz S
26	*Botryosphaeriaceous*	*sp*	2717	*Pterocarpus angolensis*	Mehl J
27	*Calonectria*	*pauciramosa*	30815	*Eucalyptus sp*	Crous PW
28	*Ceratocystis*	*omanenesis*	11056	Unknown	Unknown
29	*Ceratocystis*	*smalleyi*	14800	Unknown	Unknown
30	*Ceratocystis*	*savannae*	17300	*Acacia nigrescens*	Kamgan NG
31	*Ceratocystis*	*manginecans*	17570	Unknown	Unknown
32	*Ceratocystis*	*laricicola*	20928	Unknown	Unknown
33	*Ceratocystis*	*acaciivora*	22563	Unknown	Unknown
34	*Ceratocystis*	*decipiens*	30855	*Eucalyptus saligna*	Kamgan NG
35	*Chrysoporthe*	*austroafricana*	9343	*Tibouchina granulosa*	Roux J
36	*Circirotrichum*	*sp*	16688	*Sideroxylon ineme*	Lee S
37	*Circirotrichum*	*sp*	2153	*Eucalyptus leaves*	Crous PW
38	*Cladosporium*	*sp*	35732	*Nectaropetalum zuluense (Erythroxylaceae)*	Marincowitz S
39	*Cladosporium*	*cladosporioides*	23665	*Tipuana tipu*	Mehl J
40	*Cladosporium*	*delicatulum*	37700	*Syzygium cordatum*	Gryzenhout M
41	*Cladosporium*	*sp*	35733	*Nectaropetalum zuluense (Erythroxylaceae)*	Marincowitz S
42	*Coccomyces*	*lauraceus*	19972	*Protea burchellii*	Lee S
43	*Coccomyces*	Unknown	20397	*Protea nitida*	Lee S
44	*Coelomycete*	Unknown	315	*Widringtonia cedarbergensis*	Wingfiels MJ
45	*Coelomycete*	Unknown	316	*Widringtonia cedarbergensis*	Wingfiels MJ
46	*Coleroa*	*senniana*	19967	*Protea nitida*	Lee S
47	*Colletogloeopsis*	*zuluensis*	21216	Eucalyptus clone	Cortinas MN
48	*Colletotrichum*	*gloeosporoides*	19971	*Protea burchellii*	Lee S
49	*Colletotrichum*	*gloeosporioides*	22211	*Leucospermum sp*	Lee S
50	*Colletotrichum*	*truncatum*	23659	*Tipuana tipu*	Mehl J
51	*Colletotrichum*	*truncatum*	23660	*Tipuana tipu*	Mehl J

52	*Coniella*	*sp*	35752	*Heteropyxis natalensis (Myrtaceae)*	Marincowitz S
53	*Coniella*	*sp*	35753	*Heteropyxis natalensis (Myrtaceae)*	Marincowitz S
54	*Coniothyrium*	*ovatum*	17499	*Eucalyptus diversicolor*	Cortinas MN
55	*Coniothyrium*	*zuluense*	7541	*Eucalyptus grandis* clone & clone hybrid	Van Zyl L
56	*Coniothyrium*	*zuluense*	8912	*Eucalyptus nitens*	Roux J
57	*Coniothyrium*	*ovatum*	17500	*Eucalyptus diversicolor*	Cortinas MN
58	*Curreya*	*Proteae*	22122	*Protea laurifolia*	Lee S
59	*Cylidroclydioum*	*pauciramosum*	5683	*Podocarpus*	Unknown
60	*Cylindrocarpon*	*sp*	24212	*Pterocarpus angolensis*	Mehl J
61	*Cylindrocladium*	*scoparium*	2151	*Eucalyptus nitens*	Wingfiels MJ
62	*Cytospora*	*sp*	1505	*Eucalyptus nitens*	Wingfiels MJ
63	*Cytospora*	*sp*	1506	*Eucalyptus nitens*	Wingfiels MJ
64	*Diaporthe*	*cynaroides*	22190	*Protea cynaroides*	Lee S
65	*Diaporthe*	*perjuncta*	8597	*Vitis vinifera*	Mostert L
66	*Diaporthe*	*ambigua*	5289	Apple	Smit A
67	*Diatryphe*	Unknown	18260	*Thamnochortus spicigerus*	Lee S
68	*Diatryphe*	Unknown	18401	*Thamnochortus spicigerus*	Lee S
69	*Dibotryon*	*Unknown*	18303	*Elegia capensis*	Lee S
70	*Dichomera*	*eucalypti*	22723	*Tipuana tipu*	Mehl J
71	*Dictyochaeta*	*simplex*	18268	*Thamnochortus spicigerus*	Lee S
72	*Dinemasporium*	*strigosum*	17942	*Cannomois virgata*	Lee S
73	*Dinemasporium*	*strigosum*	18258	*Thamnochortus spicigerus*	Lee S
74	*Diplodia*	*pterocarpi sp.nov*	22721	*Pterocarpus angolensis*	Mehl J
75	*Diplodia*	*scrobiculata*	30227	*Pinus patula*	Bihon W
76	*Diplodia*	*africana*	31767	*Prunus persica*	Damm U
77	*Diplodia*	*pinea*	29135	*Pinus patula*	Wubetu B
78	*Diplodia*	*pinea*	323	*Pinus keysia*	Wingfiels MJ
79	*Diplodia*	*pinea*	29136	*Pinus pitula*	Wubetu B
80	*Diplodia*	*viriabilis*	25422	*Acacia mellifera*	van der Walt and Maris
81	*Diplodia*	*scrobiculata*	30455	*Acacia mearnsii*	Roux J
82	*Diplodia*	*pterocarpi sp.nov*	22635	*Pterocarpus angolensis*	Mehl J
83	*Diplodia*	*pinea*	322	*Pinus pinaster*	Wingfiels MJ
84	*Diplodia*	*teseivale*	2235	*Pinus patula*	Kemp GHJ
85	*Diversiporthe*	*metrosiderotis*	37321	*Metrosideros angustifolia*	Roux J
86	*Diversiporthe*	*metrosiderotis*	37322	*Metrosideros angustifolia*	Rouxj, Chen S
87	*Dothiorella*	*sp*	899	*Pinus taeda*	Unknown
88	*Dothiorella*	*oblonga*	25408	*Acacia mellifera*	van der Walt F.J.J. and Heath R.N.
89	*Drechslera*	*dematoidea*	1280	*Leucospermum cordifolium*	Von Broembsen S
90	*Drechslera*	*dematoidea*	6152	Banana	Viljoen A
91	*Endoxyla*	*Unknown*	35710	*Ficus sur (Moraceae)*	Marincowitz S
92	*Endoxyla*	*Unknown*	35709	*Ficus sur (Moraceae)*	Marincowitz S
93	*Epicoccum*	*Unknown*	24549	*Eucalyptus* wood poles	De Meyer E
94	*Eriospora*	*Filiform*	22253	*restionaceae*	Lee S
95	*Erythricium*	*salmonicolor*	7129	*Podocarpus henckellii*	Roux J
96	*Eupenicillium*	*sp*	2063	Pine bark	Wingfiels MJ
97	*Eutypa*	*consobrina*	22184	*leucospermum praecox*	Lee S
98	*Eutypa*	*consobrina*	22133	*Leucospermum canocarpodendron*	Lee S
99	*Eutypella*	*sp1*	20395	*Leucadendron salignum*	Lee S
100	*Everhartia*	*hymenuloides*	22242	*restionaceae*	Lee S
101	*Exophiala*	*heteromorpha*	24550	Pine wood pole	De Meyer E
102	*Exserolilum*	*rostratum*	1981	Pine seeds	Cilliers A
103	*Fganoderma*	*Unknown*	25879	Unknown (possibly a coral tree [Erytrina sp.]	Roux J
104	*Fusarium*	*circinatum*	24293	*Pinus pitula*	Coutinho TC
105	*Fusarium*	*ananatum*	28599	*Ananas comosus*	Van Wyk S
106	*Fusarium*	*crookwellense*	7013	Potato	Marasas WFO
107	*Ganoderma*	*sp*	10195	Karob	Roux J
108	*Ganoderma*	*Unknown*	14311	*Peltophosum africanum*	Roux J
109	*Ganoderma*	*Unknown*	25877	Salix sp	Coetzee MPA
110	*Ganoderma*	*Unknown*	25878	*Acacia sp.*	Coetzee MPA
111	*Ganoderma*	*Unknown*	25881	unknown	Coetzee M

112	*Ganoderma*	*Unknown*	25882	Unknow (possibly Willow [Salixi sp.])	Roux J
113	*Ganoderma*	*Unknown*	25883	African tulip tree (*Louis trichardt*)	Roux J
114	*Ganoderma*	*Unknown*	25884	Unknown	Roux J
115	*Ganoderma*	*Unknown*	25885	*Jacaranda mimosifolia*	Heath R
116	*Geotrichum*	*sp*	20883	*Pterocarpus augolensis*	Mehl J
117	*Geotrichum*	*sp*	20896	*Pterocarpus augolensis*	Mehl J
118	*Gibberella*	*sp*	19975	*Protea magnifica*	Lee S
119	*Gibberella*	*sp*	16670	*Protea magnifica*	Lee S
120	*Gibberella*	*Unknown*	18309	*Elegia capensis*	Lee S
121	*Gibberella*	*Unknown*	18334	*Chondropetalum tectorum*	Lee S
122	*Gibberella*	*Unknown*	18402	Eucalyptus	Lee S
123	*Gliocladium*	*roseum*	2060	Pine seeds	Wingfiels MJ
124	*Gliocladium*	*solani*	19966	*Protea nitida*	Lee S
125	*Gliocladium*	*sp*	2223	*Acacia melanoxylon*	Lamprecht S
126	*Glomerella sp*	*Unknown*	25208	*Mangifera indica*	Hinze B
127	*Glomerella sp*	*Unknown*	25213	*Mangifera indica*	Hinze B
128	*Mycosphaerella*	*nubilosa*	25970	*Eucalyptus nitens*	Perez G
129	*Mycosphaerella*	*nubilosa*	25971	*Eucalyptus nitens*	Perez G
130	*Ophiostoma*	*palmiculminatum*	23048	*Oodynichus sp.*	Roets F
131	*Ophiostoma*	*phasmae*	20676	*Protea laurifolia*	Roets F
132	*Ophiostoma*	*multiconvivator prov. Nov*	23060	*Protea longifolia*	Roets F
133	*Ophiostoma*	*africanum*	823	*Protea gaguedi*	Wingfiels MJ
134	*Ophiostoma*	*floccosum*	19359	*Pinus patula*	Zhou X
135	*Ophiostoma*	*phasmae*	20681	*Protea nerifolia*	Roets F
136	*Ophiostoma*	*phasmae*	20683	*Protea laurifolia*	Roets F
137	*Ophiostoma*	*phasmae*	20684	*Protea laurifolia*	Roets F
138	*Ophiostoma*	*phasmae*	20686	*Protea laurifolia*	Roets F
139	*Ophiostoma*	*phasmae*	23788	*Proctolaelaps vandenbergi* mite	Roets F
140	*Ophiostoma*	*phasmae*	20698	*Protea laurifolia*	Roets F

Appendix B

Media, buffers and solutions

Unless otherwise stated all media, buffers and solutions were sterilized by autoclaving at 121°C for 20 minutes. Heat liable substances were sterilized by filtration through 0.22 µL membrane filters (Millipore).

B.1 Media

B.1.1 Malt extract (ME)

Malt Extract	20 g
dH_20	1 L

B.1.2 Malt yeast extract (MYE)

Malt Extract	3 g
Yeast Extract	3 g
Glucose	10 g
Bacto-peptone	5 g
dH_20	1 L

B.1.3 Potato dextrose

Potato dextrose powder	24 g
dH_20	1 L

B.1.4 100X mineral stock solution

K_2HPO_4	100 g
$MgSO_4.7H_2O$	50 g
H_3BO_4	50 mg
$ZnSO_4.7H_2O$	40 mg
$FeCl_3.6H_2O$	30 mg
$CuSO_4.5H_2O$	4 mg
$Na_2MoO_4.2H_2O$	2 mg
Proline	1.1 g
dH20	1 L

B.1.5 100X stock calcium chloride solution

$CaCl_2.2H_2O$	1 g
dH_20	100 mL

B.1.6 1000X stock Pyroxidine-HCl

Pyroxidine-HCl	1 mg
dH_20	10 mL

(The stock solution was filter sterilised and stored in 500 µL aliquots at -20ºC.

B.1.7 Minimal (MM) medium

100X stock $CaCl_2$. $2H_20$	10 mL
1000X stock Pyroxidine-HCl	1 mL
100X mineral stock	1 L

B.1.8 Complete (CM) medium

(NH4)SO4	1.32 g
Yeast extract	5.0 g
Malt extract	5.0 g
Sucrose	20 g
dH_20	1 L
100X stock mineral solution	10 mL
100X stock $CaCl_2.2H_2O$	10 mL
1000X stock Pyroxidine-HCl	1 mL

B.1.9 Nutrient broth (NB)

Nutrient broth powder	16 g
dH_20	1 L

B.1.10 Luria-Bertani (LB) medium

Tryptone powder	10g
Yeast extract	5 g
NaCl	10 g
dH_20	1 L

B.1.11 Super Optimal Broth (SOB)

Tryptone powder	20 g
Yeast extract	5 g
NaCl	0.58 g
KCl	0.19 g
dH_20	1 L
After autoclaving add:	
$MgCl_2.6H_2O$ (1M)	10 mL
$MgSO_4.7H2O$ (1M)	10 mL

B.1.12 Super Optimal broth with Catabolite repression (SOC)

Glucose	3.60 g
SOB	1 L

Dissolve the chemical reagents in 1 L distilled water and sterilize by autoclaving at 121 ºC for 20 minutes. When supplementing with antibiotic, the autoclaved medium is cooled to approximately ~55°C and the appropriate volume of antibiotic stock is added.

For the preparation of solid medium, 2% (w/v) of agar is added before dissolved media is autoclaved.

B.2 Chemical reagents

All the chemicals used in this study, such as salts, sugars, organic solvents, and alcohols, were of analytical grade and were obtained from different suppliers. All buffers and solutions not described within the text are listed in Appendix A.

B.3 Buffers and solutions

B.3.1 SDS-PAGE electrophoresis

The SDS-PAGE gels constituted of:

4% Stacking gel
Per 5 ml:

0.5M Tris-HCL; pH 6.8	1.24 mL
30% Acrylamide	0.70 mL
10% SDS	50 µL
10% Ammonium persulfate (APS)	20 µL
dH_2O	3 mL
TEMED	5 µL

12% Resolving gel

Per 10 ml:

1.5M Tris-HCL; pH 8.8	2.5 mL
30% Acrylamide	3.7 mL
10% SDS	100 µL
10% APS	50 µL
dH_2O	3.4 mL
TEMED	5 µL

Buffers and solutions	Constituents	Purpose
1.5 M Tris-HCl (pH 8.8)	18.15% (w/v) Tris base, adjust pH with 0.1 M HCl	Stacking gel preparation
0.5 M Tris-HCL (pH 6.8)	6% (w/v) Tris base, adjust pH with 0.1 M HCl	Resolving gel preparation
SDS solution	10% (w/v) SDS	SDS-PAGE gel preparation
10% APS	10% (w/v) APS	SDS-PAGE gel preparation
2X Sample buffer	25% (v/v) 0.5 M Tris-HCl (pH 6.8), 20% (v/v) glycerol, 4% (w/v) SDS, 0.012% (w/v) bromophenol blue, 3.1% (w/v) DTT, or 4% (v/v) BME, store aliquotes at -80°C	SDS-PAGE gel loading buffer
6x Sample buffer	70% (v/v) 0.5 M Tris-HCl (pH 6.8), 30% (v/v) glycerol, 10% (w/v) SDS, 0.012% (w/v) bromophenol blue, 9.3% (w/v) DTT, or 6% (v/v) BME, store aliquotes at -80°C	SDS-PAGE gel loading buffer
1x Running buffer	0.25% (w/v) mM Tris-HCl, 2M glycine, 1% (w/v) SDS	SDS-PAGE gel electrophoresis
Coomassie stain	0.125% (w/v) Coomassie Brilliant Blue R250, 50% (v/v) methanol, 10% (v/v) glacial acetic acid	SDS-PAGE gel staining
Destaining solution	10% (w/v) methanol, 10% (v/v) glacial acetic acid	SDS-PAGE gel destaining

B.3.2 Protein concentration

Buffers and solutions	Constituents	Purpose
20% TCA solution	20% (w/v) TCA	Protein concentration

B.3.3 Competent cells preparation

Buffers and solutions	Constituents	Purpose
TFB-II buffer (pH 6.5)	10 m MOPS, 75 mM $CaCl_2$, 10 mM RbCl, 15% (v/v) glycerol solution. Adjust pH with 1M KOH and filter-sterilise (0.2 µm). Store at 4°C	Chemical competent cell preparation
TFB-I buffer (pH 5.8)	100 mM RbCl, 50 Mm $MnCl_2$, 10 mM $CaCl_2$, 30 mM potassium acetate, 15% (v/v) glycerol solution. Adjust pH with 1M acetic acid and filter sterilise (0.2 µm filter). Store at 4°C.	Chemical competent cell preparation
$MgSO_4$ (20mM)	Dissolve 50% (w/v) $MgSO_4.7H_20$ in nuclease free Milli-Q H_2O. Filter sterilise (0.2 µm filter) and store at - 20°C	Preparation of chemical competent cells

B.3.4 Transformation

Solution	Constituents	Purpose
IPTG (40 µg/mL)	Dissolve 238 mg of IPTG in 10 mL of Milli- Q H_2O .Filter-sterilize (0.22 µL filter) and store at -20°C	Blue-white screening
X-Gal (20 mg/mL)	Dissolve 2% (w/v) X-gal in DMSO. Store 500 µL aliquots in 1.5 mL tubes covered in foil and at -20°C	Blue-white screening

B.3.5 Storage

Buffer	Constituents	Purpose
TE buffer (Tris-EDTA)	1% (v/v) 1M Tris-HCl (pH 8.0), 0.2% (v/v) 0.5 M EDTA (pH 8.0), 99.2% (v/v) distilled water (dH_2O), autoclave to sterilise, aliquot and store at -20°C	DNA/RNA storage

B.3.6 Precipitation

Solution	Constituents	Purpose
NaOAC	Dissolved 40.8 NaOAC. $3H_20$ in 40 mL of dH_20 in a 100-200 mL beaker with stirring. Adjust pH value to 5 and make up volume to 100 mL with deionized water. Autoclave and store at room temperature	DNA precipitation

B.5 Chemicals and suppliers

Chemical	Supplier
1000x GelRed™	Biotium, USA
Acetone	Merck-Saarchem, South Africa
Acrylamide/bis-acrylamide, 30% solution	Sigma-Aldrich, USA
Agar	Difco, USA
Agarose	Lonza, USA
Ammonium nitrare (NH_4NO_3)	Sigma-Aldrich, USA
Ammonium persulphate (APS)	Sigma-Aldrich, USA
Bacto-peptone	Merck, Germany
Beta-mercaptoethanol (BME)	Sigma-Aldrich, USA
BigDye ®Terminator v3.1	Applied Biosystems, USA
Bromophenol blue	Sigma-Aldrich, USA
Calcium chloride ($CaCl_2$)	Merck-Saarchem, South Africa
Coomassie brilliant blue R250	Sigma-Aldrich, USA
Dextrose	Sigma-Aldrich, USA
D-glucose	Sigma-Aldrich, USA
Diaminoethane tetraacetic acid (EDTA)	Merck-Saarchem, South Africa
Dimethyl sulfoxide	Sigma-Aldrich, USA
Dithiothreitol (DTT)	Sigma-Aldrich, USA
Ethanol	Illovo Sugar, South Africa
Ferrous sulphate heptahydrate ($FeSO_4.7H_2O$)	Merck-Saarchem, South Africa
Glacial acetic acid	Merck-Saarchem, South Africa
Glycerol	Merck-Saarchem, South Africa
Glycine	Merck-Saarchem, South Africa
Hydrochloric acid (HCl)	Merck-Saarchem, South Africa
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Sigma-Aldrich, USA
LB Agar	Lab M Limited, UK
LB Broth	BioLab Merck, South Africa
Magnesium chloride hexahydrate ($MgCl_2.\ 6H_2O$)	Merck, Germany
Magnesium sulfate heptahydrate ($MgSO_4.\ 7H_2O$)	Merck, Germany
Magnesium sulphate anhydrous ($MgSO_4$)	Sigma-Aldrich, USA
Malt extract	Biolab Merck, South Africa
Manganese chloride ($MnCl_2$)	Merck-Saarchem, South Africa
Methanol	Merck-Saarchem, South Africa
MOPS	Merck-Saarchem, South Africa
Nutrient broth	BioLab Merck, South Africa
Potassium acetate	Merck-Saarchem, South Africa
Potassium chloride (KCl)	Sigma-Aldrich, USA
Potassium hydrogen phosphate, anhydrous (K_2PO_4)	Sigma-Aldrich, USA
Potassium hydroxide (KOH)	Merck, Germany
Potassium nitrate (KNO_3)	Merck, Germany
Potato dextrose broth	Difco, USA
Rubidium chloride (RbCl)	Merck-Saarchem, South Africa
Saccharose	Merck-Saarchem, South Africa
SeaKem LE Agarose	Lonza, USA

Chemical	Supplier
Sodium acetate ($NaAC.3H_20$)	Sigma-Aldrich, USA
Sodium chloride (NaCl)	Sigma-Aldrich, USA
Sodium hydroxide (NaOH)	Sigma-Aldrich, USA
Sodium lauryl sulfate (SDS)	Sigma-Aldrich, USA
Sodium nitrate ($NaNO_3$)	Sigma-Aldrich, USA
Sucrose	Merck-Saarchem, South Africa
SYBR® Safe	Life Technologies, USA
Trichloroacetic acid	Sigma-Aldrich, USA
Tris (hydromethyl) aminomethane (Tris base)	Merck, Germany
Tryptone powder	Merck, Germany
X-gal	Sigma-Aldrich, USA
Yeast extract	Difco, USA

REFERENCES

Abraham, L. D. & Breuil, C. (1996). Isolation and characterization of a subtilisin-like serine proteinase secreted by the sap-staining fungus *Ophiostoma piceae*. *Enzyme and Microbial Technology*, *18*, 133–140.

Abuodeh, R. O., Orbach, M. J., Alejandra, M. M., Das, A. & Galgiani, J. N. (2000). Genetic Transformation of *Coccidioides immitis* facilitated by *Agrobacterium tumefaciens*. *The Journal of Infectious Diseases*, *181*, 2106–2110.

Agbo, E. C. (Ed.). (2012). Microbial Expression Systems and Manufacturing from a Market and Economic Perspective. In *Innovations in Biotechnology* (p. 474). InTech .

Allgaier, S., Taylor, R. D., Brudnaya, Y., Jacobson, D. J. J., Cambareri, E. & Stuart, W. D. (2009). Vaccine production in *Neurospora crassa*. *Biologicals*, *37*, 128–132.

Almeida, J. R. M., Modig, T., Petersson, A., Hähn-Hägerdal, B., Lidén, G. & Gorwa-Grauslund, M. F. (2007). Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *Journal of Chemical Technology and Biotechnology*, *82*, 340–349.

Alvarado, P., Manjón, J. L., Matheny, P. B. & Esteve-Raventós, F. (2010). *Tubariomyces*, a new genus of *Inocybaceae* from the Mediterranean region. *Mycologia*, *102*, 1389–1397.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. D. & Struhl, K. (1987). *Current Protocols in Molecular Biology* (p. 1600). New York: John Wiley and Sons.

Bardiya, N. & Shiu, P. K. (2007). Cyclosporin A-resistance based gene placement system for *Neurospora crassa*. *Fungal Genetics and Biology*, *44*, 307–314.

Bellemare, D. R., Sanschagrin, M., Beaudoin, J., & Labbé, S. (2001). A novel copper-regulated promoter system for expression of heterologous proteins in *Schizosaccharomyces pombe. Gene, 273*, 191-198.

Bennet, J. (1998). Mycotechnology: the role of fungi in biotechnology. *Journal of Biotechnology*, *66*, 101–7.

Bezhani, S., Sherameti, I., Pfannschmidt, T. & Oelmüller, R. (2001). A repressor with similarities to prokaryotic and eukaryotic DNA helicases controls the assembly of the CAAT box binding complex at a photosynthesis gene promoter. *Journal of Biological Chemistry*, *278*, 23785–23789.

Bleichrodt, R. J., Veluw, G. J., Recter, B., Maruyama, J. I., Kitamoto, K., & Wösten, H. A. (2012). Hyphal heterogeneity in *Aspergillus oryzae* is the result of dynamic closure of septa by Woronin bodies. *Molecular microbiology*, 86, 1334-1344.

Blibech, M., Ellouz Ghorbel, R., Chaari, F., Dammak, I., Bhiri, F., Neifar, M. & Ellouz Chaabouni, S. (2011). Improved Mannanase Production from *Penicillium occitanis* by Fed-Batch Fermentation Using Acacia Seeds. *ISRN Microbiology*, *2011*, 938347.

Brandhorst, T. & Kenealy, W. R. (1995). Effects of leader sequences upon the heterologous expression of restriction in *Aspergillus nidulans* and *Aspergillus niger*. *Canadian Journal of Microbiology*, *41*, 601–611.

Broekhuijsen, M. P. (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*, 135–145.

Brown, A. J. P., Leach, M. D. & Nicholls, S. (2010). The relevance of heat shock regulation in fungal pathogens of humans. *Virulence*, *1*, 330–332.

Brown, D. E. & Zainudeen, M. A. (1987). Effect of inoculum size on the aeration pattern of batch cultures of a fungal microorganism. *Biotechnology and Bioengineering*, *20*, 1045–1061.

Brush, T. S., Chapman, R., Kurzman, R. & Williams, D. P. (1999). Purification and characterization of extracellular lipases from *Ophiostoma piliferum*. *Bioorganic and Medicinal Chemistry*, *7*, 2131-2138.

Bzymek, K. P., D'Souza, V. M., Chen, G., Campbell, H., Mitchell, A. & Holz, R. C. (2004). Function of the Signal Peptide and N- and C-terminal Propeptides in the leucine aminopeptidase from *Aeromonas proteolytica*. *Protein Expression and Purification*, *37*, 294–305.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Research*, *37*, D233–D238.

Carlquist, M., Maletti, M., Jornvall, H. & Mutt, V. (1984). A novel form of gastric inhibitory polypeptide (GIP) isolated from bovine intestine using a radioreceptor assay. fragmentation with staphylococcal protease results in GIP1-3 and GIP4-42, fragmentation with enterokinase in GIP1-16 and GIP17-42. *European Journal of Biochemistry*, *145*, 573–577.

Çelik, E., & Çalık, P. (2012). Production of recombinant proteins by yeast cells. *Biotechnology advances*, *30*, 1108-1118.

Chang, M. D. T. & Sheu, C. C. (2010). Provides combination of the modified signal peptide, increased copy number of the gene, and/or overexpressed Sec4p resulteing in high level secretion of proteins in *Pichia*. Google Patents.

Chatzifragkou, A., Makri, A., Belka, A., Bellou, S., Mavrou, M., Mastoridou, M., Mystrioti, P., Onjaro, G., Aggelis, G., Papanikolaou, S. (2011). Biotechnological conversions of biodiesel derived waste glycerol by yeast and fungal species. *Energy*, *36*, 1097–1108.

Chauhan, P. S., Puri, N., Sharma, P. & Gupta, N. (2012). Mannanases: Microbial sources, production, properties and potential biotechnological applications. *Applied Microbiology and Biotechnology*, *93*, 1817–1830.

Chen, R. (2012). Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnology Advances*, *30*, 1102–7.

Chiu, E., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. & Sheen, J. (1996). Engineered GFP as a vital reporter in plants. *Current Biology*, *6*, 325–330.

Chuan-Wei, J. & Terry, M. (2013). A novel selection marker for efficient DNA cloning and recombineering in *E. coli*. *PloS One*, *8*, e57075.

Colacino, E., Martinez, J., Lamaty, F., Patrikeeva, L. S., Khemchyan, L. L., Ananikov, V. P. & Beletskaya, I. P. (2012). PEG as an alternative reaction medium in metal-mediated transformations. *Coordination Chemistry Reviews*, *256*, 2893–2920.

Conesa, A., Punt, P. J., van Luijk, N. & van den Hondel, C. A. (2001). The secretion pathway in filamentous fungi: a biotechnological view. *Fungal Genetics and Biology, 33*, 155–71.

Cornett, C. A. & Reilly, P. J. (2003). Starch-binding domain shuffling in *Aspergillus niger* glucoamylase*. Protein Enginering, 16,* 521-529.

Craney, A., Hohenauer, T., Xu, Y., Navani, N. K., Li, Y. & Nodwell, J. (2007). A synthetic luxCDABE gene cluster optimized for expression in high-GC bacteria. *Nucleic Acids Research*, *35*, e46.

Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A. & Tsien, R. Y. (1995). Understanding, improving and using green fluorescent proteins. *Trends in Biochemical Sciences*, *20*, 448–455.

Culleton, H., Bouzid, O., A McKie, V., & P de Vries, R. (2014). New promoters to improve heterologous protein production in *Aspergillus vadensis*. *Current Biotechnology*, *3*, 244-251.

Cullen, D., Gray, G. L., Wilson, L. J., Hayenga, K. J., Lamsa, M. H. & Rey, M. W. (1987). Controlled expression and secretion of bovine chymosin in *Aspergillus nidulans*. *Nature Biotechnology*, *5*, 369–376.

De Boer, P., Bronkhof, J., Dukić, K., Kerkman, R., Touw, H., van den Berg, M. & Offringa, R. (2013). Efficient gene targeting in *Penicillium chrysogenum* using novel *Agrobacterium*-mediated transformation approaches. *Fungal Genetics and Biology*, *61*, 9–14.

De Groot, M. J., Bundock, P., Hooykaas, P. J. & Beijersbergen, A. G. (1998). *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nature Biotechnology*, *16*, 839–842.

Deb, P., Talukdar, S. A., Mohsina, K., Sarker, P. K. & Sayem, S. M. A. (2013). Production and partial characterization of extracellular amylase enzyme from *Bacillus amyloliquefaciens* P-001. *SpringerPlus*, *2*, 1–12.

Delic, M., Göngrich, R., Mattanovich, D. & Gasser, B. (2014). Engineering of protein folding and secretion-strategies to overcome bottlenecks for efficient production of recombinant proteins. *Antioxidants & Redox Signaling*, *21*, 414–37.

Demain, A. L. & Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances*, *27*, 297–306.

Desai, P. N., Shrivastava, N. & Padh, H. (2010). Production of heterologous proteins in plants: strategies for optimal expression. *Biotechnology Advances*, *28*, 427–35.

Deshpande, N., Wilkins, M. R., Packer, N., & Nevalainen, H. (2008). Protein glycosylation pathways in filamentous fungi. *Glycobiology*, *18*, 626-637.

Dhawan, S. & Kaur, J. (2007). Microbial mannanases: an overview of production and applications. *Critical Reviews in Biotechnology*, *27*, 197–216.

Domingues, F. C., Queiroz, J. A., Cabral, J. M. S. & Fonseca, L. P. (2000). The influence of culture conditions on mycelial structure and cellulase production by *Trichoderma reesei* Rut C-30. *Enzyme and Microbial Technology*, *26*, 394–401.

Dong, H. & Zhang, D. (2014). Current development in genetic engineering strategies of *Bacillus* species. *Microbial Cell Factories*, *13*, 13-63.

Dower, W., Miller, J. & Ragsdale, C. (1988). High efficiency trans-formation of *Escherichia coli* by high voltage electroporation. *Nucleic Acids Research*, *16*, 6127–6145.

Du Vigneaud, V., Ressler, C., Swan, J. M., Roberts, C. W. & Katsoyannis, P. G. (1954). The Synthesis of Oxytocin1. *Journal of American Chemical Society.*, *76*, 3115–3121.

Fadel, H. H. M., Mahmoud, M. G., Asker, M. M. S., & Lotfy, S. N. (2015). Characterization and evaluation of coconut aroma produced by

Trichoderma viride EMCC-107 in solid state fermentation on sugarcane bagasse. *Electronic Journal of Biotechnology*, *18*, 5-9.

Farrell, R. L., Thwaites, J. M., Su, Y., Wang, E. I., Ho, C. & Chang, H. (2004). Screening of fungi for removal of wood extractives. In *58th Annual Appita Proceedings* (pp. 1–10). Canberra, Australia.

Fernandez-Abalos, J. M., Fox, H., Pitt, C., Wells, B. & Doonan, J. H. (1998). Plant-adapted green fluorescent protein is a versatile vital reporter for gene expression, protein localization and mitosis in the filamentous fungus, *Aspergillus nidulans*. *Molecular Microbiology*, *27*, 121–130.

Fleißner, A. & Dersch, P. (2010). Expression and export: recombinant protein production systems for *Aspergillus*. *Applied Microbiology and Biotechnology*, *4*, 1255–1270.

Fowler, T., Berka, R. M. & Ward, M. (1990). Regulation of the gla*A* gene of *Aspergillus niger*. *Current Genetics*, *18*, 537–45.

Gasser, B. & Mattanovich, D. (2007). Antibody production with yeasts and filamentous fungi: on the road to large scale? *Biotechnology Letters*, *29*, 201–12.

Georgi, B., & Schliep, A. (2006). Context-specific independence mixture modeling for positional weight matrices. *Bioinformatics*, *22*, e166-e173.

Ghim, C., Lee, S. K., Takayama, S. & Mitchell, R. J. (2010). The art of reporter proteins in science : past , present and future applications. *BMB Reports, 43,* 451-460.

Gibb, G. D., Ordaz, D. E., & Strohl, W. R. (1989). Overproduction of extracellular protease activity by *Streptomyces* C5-A13 in fed-batch fermentation. *Applied Microbiology and Biotechnology*, *31*, 119-124.

Gibbs, P. A, Seviour, R. J. & Schmid, F. (2000). Growth of filamentous fungi in submerged culture: problems and possible solutions. *Critical Reviews in Biotechnology*, *20*, 17–48.

Glick, D., Barth, S. & Macleod, K. F. (2010). Autophagy : cellular and molecular mechanisms. *Journal of Pathology*, *221*, 3–12.

Goeddel, D. V, Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., … Riggs, A. D. (1979). Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proceedings of the National Academy of Sciences of the United States of America*, *76*, 106–10.

Gouka, R. J., Punt, P. J., & Van Den Hondel, C. A. M. J. J. (1997). Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Applied Microbiology and Biotechnology*, *47*, 1-11.

Gruber, F., Visser, J., Kubicek, C. P. & de Graaff, L. H. (1990). The development of a heterologous transformation system for the cellulolytic fungus *Trichoderma reesei* based on a pyrG-negative mutant strain. *Current Geneticts*, *18*, 71–6.

Güssow, D. & Clackson, T. (1989). Direct clone characterization from plaques and colonies by the polymerase chain reaction. *Nucleic Acids Research*, *17*, 4000-6.

Heim, R., Cubitt, A. B. & Tsien, R. Y. (1995). Improved green fluorescence. *Nature*, *373*, 663–664.

Hensing, M. C. M., Rouwenhorst, R. J., Heijnen, J. J., van Dijken, J. P. R. & Pronk, J. T. (1995). Physiological and technological aspects of large-scale heterologous-protein production with yeasts. *Antonie van Leeuwenhoek*, *67*, 261–279.

Hoffman, B. & Breuil, C. (2004). Disruption of the Subtilase Gene, albin1, in *Ophiostoma piliferum*. *Applied Environmental Microbiology*, *70*, 3898–3903.

Hu, W., Sillaots, S., Lemieux, S., Davison, J., Kauffman, S., Breton, A., ... & Roemer, T. (2007). Essential gene identification and drug target prioritization in *Aspergillus fumigatus*. *PLoS Pathogens*, 3, e24.

Huang, C.-J., Lin, H. & Yang, X. (2012). Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *Journal of Industrial Microbiology & Biotechnology*, *39*, 383–99.

Hunter, C. C., Siebert, K. S., Downes, D. J., Wong, K. H., Kruetzberger, S. D., Fraser, J. A., … Todd, R. B. (2014). Multiple Nuclear Localization Signals Mediate Nuclear Localization of the GATA Transcription Factor AreA. *Eukaryotic Cell*, *13*, 527–538.

Itakure, K., Hirose, T., Crea, R., Riggs, A., Heyneker, H. & Bolivar, F. (1977). Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science*, *198*, 1056–63.

Izard, J. W. & Kendall, D. A. (1994). Signal peptides : exquisitely designed transport promoters. *Microbiology Reviews*, *13*, 765–773.

Jalving, R., van de Vondervoort, P. J. Visser, J. & Schaap, P. J. (2000). Characterization of the kexin-like maturase of *Aspergillus niger*. *Applied and Environmental Microbiology*, *66*, 363–8.

James, A. L., Perry, J. D., Ford, M., Armstrong, L. & Gould, F. K. (1996). Evaluation of cyclohexenoesculetin-beta-D-galactoside and 8-hydroxyquinoline-beta-D-galactoside as substrates for the detection of beta-galactosidase. *Applied Environmental Microbiology*, *62*, 3868–3870.

Josephson, S. & Bishop, R. (1988). Secretion of peptides from *E.coli*: a production system for the pharmaceutical industry *Herpes simplex*. *Trends in Biotechnology*, *6*, 218–224.

Kamionka, M. (2011). Engineering of Therapeutic Proteins Production in *Escherichia coli*. *Current Pharmaceutical Biotechnology*, *2*, 268–274.

Kansoh, A. L. & Nagieb, Z. A. (2004). Xylanase and mannanase enzymes from *Streptomyces galbus* NR and their use in biobleaching of softwood kraft pulp. *Antonie van Leeuwenhoek*, *85*, 103–114.

Keen, H., Pickup, J. C., Bilous, R. W., Glynne, A., Viberti, G. C., Jarrett, R. J. & Marsden, R. (1980). Human insulin produced by recombinant DNA technology: Safety and hypoglycaemic potency in healthy men. *The Lancet*, *316*, 398–401.

Khow, O. & Suntrarachun, S. (2012). Strategies for production of active eukaryotic proteins in bacterial expression system. *Asian Pacific Journal of Tropical Biomedicine*, *2*, 159–62.

Kinghorn, J. R. & Hawkins, A. R. (1982). Cloning and expression in *Escherichia coli* " K-12 of the biosynthetic dehydroquinase function of the arom cluster gene from the eucaryote , *Aspergillus nidulans*. *Molecular & general genetics, 186,* 145–152.

Klyosov, A. A., Dotsenko, G. S., Hinz, S. W. A. & Sinitsyn, A. P. (2012). Structural features of - (1→4)-d-galactomannans of plant origin as a probe for -(1→4)-mannanase polymeric substrate specificity. *Carbohydrate Research*, *352*, 65–9.

Kosalková, K., García-Estrada, C., Barreiro, C., Flórez, M. G., Jami, M. S., Paniagua, M. A., & Martín, J. F. (2012). Casein phosphopeptides drastically increase the secretion of extracellular proteins in *Aspergillus awamori*. Proteomics studies reveal changes in the secretory pathway. *Microbial Cell Factories*, *11*, 1-15.

Kote, N. V., Patil, A. G. & Mulimani, V. H. (2009). Optimization of the production of thermostable endo- β -1,4 mannanase from a newly isolated *Aspergillus niger* gr. and *Aspergillus flavus* gr. *Applied Biochemistry and Biotechnology*, *152*, 213 – 223.

Kourtoglou, E., Anasontzis, G. E., Mamma, D., Topakas, E., Hatzinikolaou, D. G. & Christakopoulos, P. (2011). Constitutive expression, purification and characterization of a phosphoglucomutase from *Fusarium oxysporum*. *Enzyme and Microbial Technology*, *48*, 217–24.

Kück, U. & Hoff, B. (2010). New tools for the genetic manipulation of filamentous fungi. *Applied Microbiology and Biotechnology*, *86*, 51–62.

Kwon, M. J., Jørgensen, T. R., Nitsche, B. M., Arentshorst, M., Park, J., Ram, A. F. J. & Meyer, V. (2012). The transcriptomic fingerprint of glucoamylase over-expression in *Aspergillus niger*. *BMC Genomics*, *13*, 1471-2164.

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, *227*, 680–685.

Lang, C. & Looman, A. C. (1995). Efficient expression and secretion of *Aspergillus niger* polygalacturonase in *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology*, *44*, 147–156.

Larrondo, L. F., Colot, H. V., Baker, C. L., Loros, J. & Dunlap, J. C. (2009). Fungal functional genomics: tunable knockout-knock-in expression and tagging strategies. *Eukaryotic Cell*, *8*, 800–4.

Leroch, M., Mernke, D., Koppenhoefer, D., Schneider, P., Mosbach, A., Doehlemann, G. & Hahn, M. (2011). Living Colors in the gray mold pathogen *botrytis cinerea* : Codon-optimized genes encoding green fluorescent protein and mCherry , which exhibit bright fluorescence . *Applied and Environmental Microbiology*, *77*, 2887–2897.

Li, J., Wang, J., Wang, S., Xing, M., Yu, S., & Liu, G. (2012). Achieving efficient protein expression in *Trichoderma reesei* by using strong constitutive promoters. *Microbial Cell Factories*, *11*, 84-93.

Liao, X. G., Zhang, Y. J., Fan, Y. H., Ma, J. C., Zhou, Y. H., Jin, D., & Pei, Y. (2009). An ethanol inducible alc system for regulating gene expression in *Beauveria bassiana*. *World Journal of Microbiology and Biotechnology*, *25*, 2065-2069.

Liu, L., Liu, J. & Qiu, R. X. (2003). Improving heterologous gene expression in *Aspergillus niger* by introducing multiple copies of protein- binding sequence containing CCAAT to the promoter. *Letters in Applied Microbiology*, *36*, 358–361.

Loppnau, P., Tanguay, P. & Breuil, C. (2004). Isolation and disruption of the melanin pathway polyketide synthase gene of the softwood deep stain fungus *Ceratocystis resinifera*. *Fungal Genetics and Biology*, *41*, 33–41.

Lubertozzi, D. & Keasling, J. D. (2009). Developing *Aspergillus* as a host for heterologous expression. *Biotechnology Advances*, *27*, 53–75.

Lynch, M. (2006). The origins of eukaryotic gene structure. *Molecular Biology and Evolution*, *23*, 450–468.

Mabrouk, M. E. M., & Ahwany, A. M. D. (2008). Production of β-mannanase by *Bacillus amyloliquefaciens* 10A1 cultured on potato peels. *African Journal of Biotechnology*, *7*, 1123 – 1128.

Mach, R. L., Schindler, M. & Kubicek, C. P. (1994). Transformation of *Trichoderma reesei* based on hygromycin B resistance using homologous expression signals. *Current Geneticts*, *25*, 567–70.

Magaña-Ortíz, D., Coconi-Linares, N., Ortiz-Vazquez, E., Fernández, F., Loske, A. M. & Gómez-Lim, M. A. (2013). A novel and highly efficient method for genetic transformation of fungi employing shock waves. *Fungal Genetics and Biology, 56*, 9–16.

Mahmound, K. (2007). Recombinant Protein Production: Strategic technology and a vital research tool. *Research Journal of Cell and Molecular Biology*, *1*, 9–22.

Malgas, S., van Dyk, S. J. & Pletschke, B. I. (2015). β-Mannanase (Man26A) and α-galactosidase (Aga27A) synergism – A key factor for the hydrolysis of galactomannan substrates. *Enzyme and Microbial Technology*, *70*, 1–8.

Mao, W., Pan, R., & Freedman, D. (1992). High production of alkaline protease by *Bacillus licheniformis* in a fed-batch fermentation using a synthetic medium. *Journal of industrial microbiology*, *11*, 1-6.

Masani, M. Y. A., Noll, G. A, Parveez, G. K. A., Sambanthamurthi, R. & Prüfer, D. (2014). Efficient transformation of oil palm protoplasts by PEG-mediated transfection and DNA microinjection. *PloS One*, *9*, e96831.

Metz, B., & Kossen, N. W. F. (1977). The growth of molds in the form of pellets–a literature review. *Biotechnology and Bioengineering*, 19, 781-799.

Meyer, V., Wanka, F., van Gent, J., Arentshorst, M., van den Hondel, C. a M. J. & Ram, A. F. J. (2011). Fungal gene expression on demand: an inducible, tunable, and metabolism-independent expression system for *Aspergillus niger*. *Applied and Environmental Microbiology*, *77*, 2975–83.

Meyrath, J. & Macintosh, A. F. (1963). Size of inoculum and carbon metabolism in some *Aspergillus* species. *Journal of General Microbiology*, *33*, 47–56.

Moore, M. (2007). Genetic engineering of fungal cells. *Biotechnology*, *3*, 1-18.

Moralejo, F. J., Cardoza, R. E., Gutierrez, S., & Martin, J. F. (1999). Thaumatin production in *Aspergillus awamori* by use of expression cassetes with strong fungal promoters and high gene dosage. *Applied Environmental Microbiology*, *65*, 1168–1174.

Morán-Diez, M. E, Cardoza, R. E, Gutiérrez, S, Monte, E & Hermosa, R. (2010). TvDim1 of *Trichoderma virens* is involved in redox-processes and confers resistance to oxidative stresses. *Current Genetics, 56*, 63-73.

Moreira, L. R. & Filho, E. X. (2008). An overview of mannan structure and mannan-degrading enzyme systems. *Applied Microbiology and Biotechnology*, *79*, 165–178.

Mullins, E. D., Chen, X., Romaine, P., Raina, R., Geiser, D. M. & Kang, S. (2001). *Agrobacterium*-mediated transformation of *fusarium oxysporum*: An efficient tool for insertional mutagenesis and gene transfer. *Phytopathology*, *91*, 173–180.

Nevalainen, H. & Peterson, R. (2014). Making recombinant proteins in filamentous fungi- are we expecting too much? *Frontiers in Microbiology*, *5*, 75.

Nevalainen, K. M. H., Te'o, V. S. J. & Bergquist, P. L. (2005). Heterologous protein expression in filamentous fungi. *Trends in Biotechnology*, *23*, 468–74.

Newmann, H., Peak-Chew, S. Y, J.W C. (2008). Genetically encoding Ne-acetylllysine in recombinant proteins. *Nature Chemical Biology*, *4*, 232–4.

Nickoloff, J. A. & Miller, E. M. (1995). Electroporation Protocols for Microorganisms. In J. Nickoloff (Ed.), *Methods in Molecular Biology* (p. 372). Springer Science and Business media.

Nielsen, J. (2013). Production of biopharmaceutical proteins by yeast. *Advances Through Metabolic Engineering*, 10, 207–211.

Nothwehr, S. F. & Gordon, J. I. (1990). Targeting of proteins into the eukaryotic secretory pathway: signal peptide structure/function relationship. *Bioessays*, *12*, 479–484.

Olempska-Beer, Z. S ,Merker, R. I, Ditto, M. D & DiNovi, M. J. (2006). Food-processing enzymes from recombinant microorganisms—a review. *Regulatory Toxicology and Pharmacology*, *45*, 144–158.

Padh, H., Rai, M. & Padh, H. (2001). Expression systems for production of heterologous proteins. *Current science*, *12*, 18-19.

Palomerases, L. A. , Estrada-Mondaca, S. & Ramírez, O. T. (2004). Production of Recombinant Proteins: Challenges and Solutions in, *Recombinant Gene Expression: Reviews and Protocols*. (pp 100-125)Humana Press. New York.

Papagianni, M. (2004). Fungal morphology and metabolite production in submerged mycelial processes. *Biotechnology Advances*, *22*, 189–259.

Penttila, M., Nevalainen, H., Ratto, M., Salminen, E. & Knowles, J. (1987). A versatile transformation system for the cellulolytic filamentous fungi *Trichorderma reesei*. *Gene*, 61, 155–164.

Pereira, M., Felipe, M. S., Brigido, M. M., Soares, C. M. A. & Azevedo, M. O. (2000). Molecular cloning and characterization of a glucan synthase gene from the human pathogenic fungus *Paracoccidioides brasiliensis*. *Yeast*, *16*, 451–462.

Politz, O., Krah, M., Thomsen, K. K. & Borriss, R. R. (2000). A highly thermostable endo 1,4-beta-mannanase from the marine bacterium *Rhodotherus marinus*. *Applied Microbiology and Biotechnology*, *53*, 715–721.

Prabavathy, V. R., Mathivanan, N., Sagadevan, E., Murugesan, K. & Lalithakumari, D. (2006). Self-fusion of protoplasts enhances chitinase production and biocontrol activity in *Trichoderma harzianum*. *Bioresource Technology*, *97*, 2330–2334.

Przybyl, K., Dahm, H., Ciesielska, A. & Molinski, K. (2006). Cellulolytic activity and virulence of *Ophiostoma ulmi* and *O. novo-ulmi* isolates. *Forest Pathology*, *36,* 58 – 67.

Puchart, V., Vrsanska, M., Svoboda, P., Pohl, J., Ogel, Z. B. & Biely, P. (2004). Purification and characterization of two forms of endo-beta-1,4-mannanase from a thermotolerant fungus, *Aspergillus fumigatus* IMI 385708 (formerly *Thermomyces lanuginosus* IMI 158749). *Biochimica et Biophysica Acta*, *1674*, 239–50.

Punt, P. J., Oliver, R. P., Dingemanse, M. A., Pouweisa, P. H. & van den Handel, C. A. M. (1987). Transformation of *Aspergillus* based on the hygromycin-B resistance marker from *E. coli*. *Gene*, *56*, 117 – 124.

Punt, P. J. P., van Biezen, N., Conesa, A., Albers, A., Mangnus, J., van den Hondel, C., … van den Hondel, C. (2002). Filamentous fungi as cell factories for heterologous protein production. *Trends in Biotechnology*, *20*, 200–6.

Punt, P. J., Zegers, N. D., Busscher, M., Pouwels, P. H. & van den Hondel, C. A. (1991). Intracellular and extracellular production of proteins in *Aspergillus* under the control of expression signals of the highly expressed *Aspergillus nidulans* gpdA gene. *Journal of Biotechnology*, *17*, 19–33.

Qinnghe, C., Xiaoyu, Y., Tiangui, N., Cheng, J. & Qiugang, M. (2004). The screening of culture condition and properties of xylanase by white rot fungus *Pleurotus ostreatus*. *Processes in Biochemistry*, *39*, 1561–1566.

Qui, R., Zhu, X., Liu, L. & Tang, G. (2002). Detection of a protein AngCP, which binds specifically to the three upstream regions of glaA gene in *A. niger* T21. *Science China Series C Life Science*, *45*, 527–537.

Raffaello, T., Chen, H., Kohler, A. & Asiegbu, F. O. (2013). Transcriptomic profiles of *Heterobasidion annosum* under abiotic stresses and during saprotrophic growth in bark, sapwood and heartwood. *Environmental Microbiology*, *16*, 1654–1667.

Redkar, R. J., Herzog, R. W. & Singh, N. K. (1998). Transcriptional activation of the *Aspergillus nidulans* gpdA promoter by osmotic signals. *Applied Environmental Microbiology*, *64*, 2229–31.

Rivera, A. L., Magaña-Ortíz, D., Gómez-Lim, M., Fernández, F. & Loske, A. M. (2014). Physical methods for genetic transformation of fungi and yeast. *Physics of Life Reviews*, *11*, 184–203.

Robson, L. M. (2008). Comparative and Functional Analysis of Gene Expression in *Ophiostoma Species*. University of Waikato.

Rodriguez-Gacio Mdel, C., Iglesias-Fernandez, R. Carbonero, P. & Matilla, A. J. (2012). Softening-up mannan-rich cell walls. *Journal of Eperimental Botany*, *63*, 3976–3988.

Roets, F., Theron, N., Wingfield, M. J. & Dreyer, L. L. (2012). Biotic and abiotic constraints that facilitate host exclusivity of *Gondwanamyces* and *Ophiostoma* on *Protea*. *Fungal Biology*, *116*, 49–61.

Rollins, J. A., Wolpert, T. J., Johnson, K. B., Dickman, M. B., Ciuffetti, L. M., Lorang, J. M., … Ciuffetti, L. M. (2001). Green Fluorescent Protein is Lighting up Fungal Biology. *Applied and Environmental Microbiology*, *67*, 1987–1994.

Rotenstein, L. S., Ran, N., Shivers, J. P., Yarchoan, M., & Close, K. L. (2012). Opportunities and challenges for biosimilars : What ' s on the Horizon in the global insulin market ? *Clinical Diabetes*, *30*, 138–150.

Royer, J. C., Dewar, K., Hubbes, M. & Horgen, P. A. (1991). Analysis of a high frequency transformation system for *Ophiostoma ulmi*, the causal agent of Dutch elm disease. *Molecular & General Genetics* , *225*, 168–76.

Ruiz-Díez, B. (2002). Strategies for the transformation of filamentous fungi. *Journal of Applied Microbiology*, *92*, 189–95.

Sachslehner, A., Nidetzky, B., Kulbe, K. D. & Haltrich, D. (1998). Induction of Mannanase, Xylanase, and Endoglucanase Activities in *Sclerotium rolfsii*. *Applied and Environmental Microbiology*, *64*, 594–600.

Santamauro, F., Whiffin, F. M., Scott, R. J. & Chuck, C. J. (2014). Low-cost lipid production by an oleaginous yeast. *Biotechnology for Biofuels,7, 1-11.*

Saraswat, M., Musante, L., Ravidá, A., Shortt, B., Byrne, B. & Holthofer, H. (2013). Preparative purification of recombinant proteins: current status and future trends. *BioMed Research International*, *2013*, 312709.

Seifert, K. A. (1993). Sapstain of commercial lumber by species of *Ophiostoma* and *Cyratocystis*. In M. J. Wingfield, K. A. Seifert, & J. F. Webber (Eds.), *Ceratocystis and Ophistoma . Taxonomy, ecology and pathogenicity* (pp. 141–151). St. Paul, Minnesota: American Phytopathological Press.

Sharma, R., Katoch, M., Srivastava, P. S. & Qazi, G. N. (2009). Approaches for refining heterologous protein production in filamentous fungi. *World Journal of Microbiology and Biotechnology*, *25*, 2083–2094.

Shoji, J. Y., Maruyama, J. I., Arioka, M., & Kitamoto, K. (2005). Development of *Aspergillus oryzae* thiA promoter as a tool for molecular biological studies. *FEMS Microbiology Letters, 244*, 41-46.

Silva, G. B., Ionashiro, M., Carrara, T. B., Crivellari, A. C., Tine, M. A., Prado, J., … Buckeridge, M. S. (2011). Cell wall polysaccharides from fern leaves: evidence for a mannan-rich Type III cell wall in *Adiantum raddianum*. *Phytochemistry*, *72*, 2352–2360.

Singh, J., Vohra, R. M., & Sahoo, D. K. (2004). Enhanced production of alkaline proteases by Bacillus sphaericus using fed-batch culture. *Process Biochemistry*, *39*, 1093-1101.

Spellig, T., Bottin, A. & Kahmann, R. (1996). Green fluorescent protein (GFP) as a new vital marker in the phytopathogenic fungus *Ustilago maydis*. *Molecular and General Genetics*, *252*, 503–509.

Su, X., Schmitz, G., Zhang, M., Mackie, R. I., Cann, I. K. O. O., Xiaoyun, S., … Xiaoyun, S. (2012). Heterologous gene expression in filamentous fungi. *Advances in Applied Microbiology*, *81*, 1–61.

Tamayo-Ramos, J., & Orejas, M. (2014). Enhanced glycosyl hydrolase production in *Aspergillus nidulans* using transcription factor engineering approaches. *Biotechnology for Biofuels*, *7*, 103.

Tan, G., Gao, Y., Shi, M., Zhang, X., He, S., Chen, Z., & An, C. (2005). SiteFinding-PCR: a simple and efficient PCR method for chromosome walking. *Nucleic Acids Research*, *33*, e122.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, *72*, 211–22.

Tilburn, J., Scazzoccbio, C., Taylor, G. G., Zabicky-zissman, J. H., Robin, A. & Davies, R. W. (1983). Transformation by integration in *Aspergillus nidulans*. *Gene*, *26*, 205-221.

Tøttrup, H. V. & Carlsen, S. (1990). A process for the production of human proinsulin in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, *35*, 339–48.

Van den Hondel, C. A. M. J. J., Punt, P. J., & van Gorcom, R. F. (1991). Heterologous gene expression in filamentous fungi. *More gene manipulations in fungi*, 396-428

Verdoes, J. C., Punt, P. J. & Hondel, C. A. M. (1995). Molecular genetic strain improvement for the overproduction of fungal proteins by Filamentous fungi. *Applied Microbiology and Biotechnology*, *43*, 195–205.

Visser, H., Joosten, V., Punt, P. J., Gusakov, A. V, Olson, P. T., Joosten, R., … Wery, J. (2011). Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. *Industrial Biotechnology*, *7*, 214–224.

Vecht-Lifshitz, S. E., Magdassi, S., & Braun, S. (1990). Pellet formation and cellular aggregation in *Streptomyces tendae*. *Biotechnology and Bioengineering, 35*, 890-896.

Wang, B. & Xia, L. (2011). High efficient expression of cellobiase gene from *Aspergillus niger* in the cells of *Trichoderma reesei*. *Bioresource Technology*, *102*, 4568–4572.

Wang, C. Q., Eufemi, M., Turano, M. & Giartosio, A. (1996). Influence of the carbohydrate moiety on the stbility of glycoproteins. *Biochemistry*, *35*, 7299–7307.

Wang, G. Y., & Keasling, J. D. (2002). Amplification of HMG-CoA reductase production enhances carotenoid accumulation in *Neurospora crassa*. *Metabolic Engineering*, *4*, 193–201.

Wang, L., Ridgway, D., Gu, T., & Moo-Young, M. (2003). Effects of process parameters on heterologous protein production in *Aspergillus niger* fermentation. *Journal of Chemical Technology and Biotechnology*, *78*, 1259-1266.

Wang, L., Ridgway, D., Gu, T. & Moo-Young, M. (2005). Bioprocessing strategies to improve heterologous protein production in filamentous fungal fermentations. *Biotechnology Advances*, *23*, 115–29.

Wang, Y., Vilaplana, F., Brumer, H. & Aspeborg, H. (2013). Enzymatic characterization of a glycoside hydrolase family 5 subfamily 7 (GH5_7) mannanase from *Arabidopsis thaliana*. *Planta*, *7*, 1–13.

Ward, O. P. (2011). Production of recombinant proteins by filamentous fungi. *Biotechnology Advances*, *30*, 1119–39.

Ward, W. W. (1979). *Photochemical and Photobiological Reviews Vol. 4*. (K. C. Smith, Ed.) (pp. 1–57). Plenum Press.

Watts, C., Wickner, W., & Zimmermann, R. (1983). M13 procoat and a pre-immunoglobulin share processing specificity but use different membrane receptor mechanisms. *Proceedings of the National Academy of Sciences*, *80*, 2809-2813.

Weinhandl, K., Winkler, M., Glieder, A. & Camattari, A. (2014). Carbon source dependent promoters in yeasts. *Microbial Cell Factories, 13,* 1–17.

Wilde, C., Gold, N. D. D., Bawa, N., Tambor, J. H. M. H., Mougharbel, L., Storms, R. & Martin, V. J. (2012). Expression of a library of fungal β-

glucosidases in *Saccharomyces cerevisiae* for the development of a biomass fermenting strain. *Applied Microbiology and Biotechnology*, *95*, 647–59.

Wood, K. V. (1995). Marker proteins fro gene expression. *Current Opinion in Biotechnology*, *6*, 50–58.

Wu, C., Te'o, V. S. J., Farrell, R. L., Bergquist, P. L. & Nevalainen, K. M. H. (2006). Improvement of the secretion of extracellular proteins and isolation and characterization of the amylase I (amy1) gene from *Ophiostoma floccosum*. *Gene*, *384*, 96–103.

Xianjin, O., Chaoneng, J., Han Xueqing, Z., Zhao Xiaodong, R., Mao Yumin, L. & Bartlam Mark, W. L. (2006). Crystal Structures of Human Glycerol 3-phosphate Dehydrogenase 1 (GPD1). *Journal of Molecular Biology*, *3*, 858–869.

Yuan, X.-L., Goosen, C., Kools, H., van der Maarel, M. J. E. C., van den Hondel, C. M. J. J., Dijkhuizen, L. & Ram, A. F. J. (2006). Database mining and transcriptional analysis of genes encoding inulin-modifying enzymes of *Aspergillus niger*. *Microbiology*, *152*, 3061–73.

Zhang, Z. & Dietrich, F. S. (2005). Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5′ SAGE. *Nucleic Acids Research*, *33*, 2838–2851.

Zhang, J. (2010). Mammalian cell culture for biopharmaceutical production. *Manual of Industrial Microbiology and Biotechnology* (3rd edition). pp. 157–178.

Zhao, G., Yao, Y., Hou, L., Wang, C., & Cao, X. (2014). Draft genome sequence of *Aspergillus oryzae* 100-8, an increased acid protease production strain. *Genome Announcements, 2*, e00548-14.

Zhu, X., Qui, R., Liu, L. & Tang, G. (2004). Binding of trans -acting protein AngCP to the CCAAT-containing motifs in *Aspergillus niger* glaA promoter. *Progress in Natural Science*, *15*, 1–5.

Zipfel, R. D., de Beer, Z. W., Jacobs, K., Wingfield, B. D. & Wingfield, M. J. (2006). Multigene phylogenies define *Ceratocystiopsisi* and *Grosmannia* distinct from *Ophiostoma*. *Studies in Mycology,* 55, 75–97.